

DOCTORADO EN CIENCIA Y TECNOLOGÍA

---

Evaluado y acreditado por la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU).

*Resolución N° 1178/11. Calificación "B"*

---

MODELOS DE OPTIMIZACIÓN EN BIOLOGÍA:  
METODOLOGÍA DE EVALUACIÓN Y APLICACIONES

---

Trabajo de tesis para optar por el título de Doctor en Ciencia y Tecnología  
de la Universidad Nacional de General Sarmiento

Autor: SANTIAGO RAÚL DOYLE

Director DR. FERNANDO R. MOMO





DOCTORADO EN CIENCIA Y TECNOLOGÍA

Evaluado y acreditado por la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU). Resolución N° 1178/11. Calificación "B"

l. Alcance geográfico y/o temporal de la Tesis: El trabajo de Tesis posee un alcance internacional ya que trata sobre modelos matemáticos generales que pueden aplicarse a sistemas biológicos independientemente de su localización geográfica.

m. Temas tratados en la Tesis (palabras claves):

*optimización, modelos matemáticos, restricciones, inferencia, estadística, modelo probabilístico, teoría axiomática de la probabilidad, teoría de la medida.*

n. Resumen en español:

Aunque diversos modelos de optimización se encuentran en la base de gran parte de la biología teórica, no existe en la actualidad una metodología general para su evaluación empírica. Este trabajo de tesis busca establecer si es posible evaluar de forma general el sustento empírico de estos modelos mediante métodos estándar de inferencia estadística. En el Capítulo 1 se ofrece una breve introducción a la temática. En el Capítulo 2 se presenta una revisión bibliográfica que analiza cuál es la relevancia actual de las críticas planteadas a la aproximación teórico-metodológica de la optimización en la biología. En el Capítulo 3 se presenta una generalización de los modelos de optimización basada en la teoría axiomática de la probabilidad que permite una evaluación empírica exenta de las limitaciones de los métodos tradicionalmente utilizados. En el Capítulo 4 se presenta una primera aplicación de la metodología desarrollada a modelos de forrajeo óptimo. Finalmente, se presenta una discusión general de los aportes de este trabajo y sus potenciales contribuciones al estudio de procesos de optimización como principios generales de organización de los sistemas biológicos.

o. Resumen en portugués:

Embora diversos modelos de otimização encontram-se na base de uma grande parte da biologia teórica, não existe atualmente uma metodologia geral para sua avaliação empírica. Este trabalho de tese procura estabelecer se é possível avaliar de forma geral o sustento empírico desses modelos através de métodos padronizados de inferência estatística. Uma breve introdução ao assunto é apresentada no Capítulo 1. No Capítulo 2 é apresentado um trabalho de revisão bibliográfica com o objetivo de analisar qual é atualmente a relevância das críticas à aproximação teórico-metodológica da otimização na biologia. No Capítulo 3 é apresentada uma

DOCTORADO EN CIENCIA Y TECNOLOGÍA

Evaluado y acreditado por la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU). Resolución N° 1178/11. Calificación "B"

generalização dos modelos de otimização baseada na teoria axiomática da probabilidade que permite uma avaliação empírica isenta das limitações dos métodos tradicionalmente utilizados. No Capítulo 4 é apresentada uma primeira aplicação da metodologia desenvolvida a modelos de forragem ótimo. Finalmente, apresenta-se uma discussão geral dos aportes deste trabalho e suas potenciais contribuições ao estudo de processos de otimização como princípios gerais de organização dos sistemas biológicos.

p. Resumen en inglés:

Although a variety of optimization models lie at the foundations of theoretical biology, a general methodology for the empirical evaluation of these models is still lacking. This work seeks to establish whether it is possible to perform a general evaluation of the empirical support of optimization models using standard methods of statistical inference. Chapter 1 offers a general introduction to the subject. In Chapter 2 we present a bibliographic review with the objective to analyze which is the current relevance of the critics aimed at the theoretical and empirical optimization approach in biology. In Chapter 3 we present a generalization of optimization models based on axiomatic probability theory that allows an empirical evaluation without the limitations of traditionally used methods. In Chapter 4 we present the first application of the developed methodology to models of optimal foraging. Finally, we present a general discussion of the main contributions of this work and its potential implications to the study of optimization phenomena as general organization principles of biological systems.

q. Aprobado por (Apellidos y Nombres del Jurado):

Dra. Canziani, Graciela Ana  
Dr. Aparicio, Juan Pablo  
Dr. Crespo, José Emilio

Firma y aclaración de la firma del Presidente del Jurado:

Dra. Graciela Ana Canziani

Firma del autor de la tesis:

Santiago Raúl Doyle



*Modelos de optimización en biología:  
metodología de evaluación y aplicaciones*

Publicaciones:

No se han realizado a la fecha publicaciones que abarquen el contenido de este trabajo de Tesis.

Aportes originales:

En el Capítulo 2 se presenta una síntesis de las críticas que han sido realizadas a los modelos de optimización en biología, tanto en su fundamentación teórica como en su evaluación empírica, y un análisis de la prevalencia de deficiencias metodológicas en una selección de publicaciones recientes. El principal aporte de este trabajo de tesis se presenta en el Capítulo 3 y consiste en una formulación basada en la teoría axiomática de la probabilidad que generaliza los modelos de optimización y permite, bajo ciertas condiciones generales, una evaluación empírica objetiva mediante una inferencia estadística basada en la función de verosimilitud. En el Capítulo 4 se presenta una primer aplicación de la metodología desarrollada a la evaluación de modelos de optimización de forrajeo a través de una implementación basada en métodos de Monte Carlo.



*Dedicada a mi viejo,  
a mi madre,  
a Vero,  
y a León.*



## AGRADECIMIENTOS

---

A mi director de tesis, Dr. Fernando R. Momo, por haber hecho posible este trabajo al haberme introducido en el área de la modelación matemática cuando era todavía un estudiante de biología; por darme lugar de trabajo, trabajo a secas y la oportunidad de investigar en la temática que más me interesaba; y por mantener su paciencia y confianza hasta el final después de tantos años.

Al Dr. Gustavo Ferreyra por haber aceptado la co-dirección de la beca doctoral. Al Dr. Mariano Di Leo y Dra. Gabriela Romano, por haber hecho posible la gestión de una licencia de mis tareas de docencia universitaria, y sin la cual no hubiese podido completar este trabajo de tesis.

A todo el equipo técnico administrativo del Instituto de Ciencias de la UNGS por su ayuda con innumerables gestiones: Ana, Erika, Agustina, Mauro, Emilia, y en especial a Yamila, Romina y Sabrina.

A la Dra. Lilia Romanelli, por todo su apoyo en mis primeros años en la UNGS y todas las oportunidades que me dio. A Doris Abele, por su apoyo, cariño, las oportunidades brindadas, la buena onda y la paciencia. A Leo Saravia, en quien encontré un compañero de proyectos, modelos, código y otras locuras.

A mis amigos que me acompañaron todos estos años: Ana, Ceci, Pancho, Sandra, Gaby, Guille, Ruso, Fer y todos los lujanensis. A los “locales” por el aguante en el último tiempo: Eugenio, Cintia, Mariana y Cucho, a quién debo el resumen en portugués. A Lucas por todo el cariño y todo lo que nos hiciste reír, un abrazo enorme.

A mi familia por todo el apoyo de siempre, aunque nunca tuviesen claro de que se trataba exactamente el trabajo ni por qué nunca se terminaba: Ale, Vero, Gladys, Pity, Leo, Dani, Fran, Belu, Palito, Celeste, Mariana y Paula. A mis sobrinos peques por tanto amor y alegría: Manuco, Sabi, Saulo, Martu, Lucy y Gaspi (y ahora también Luca y Agustina!).

A mi mamá, por todo el amor, el cuidado, la ayuda, la paciencia y taaaantas cosas más. A mi viejo, a quién voy a querer, recordar y extrañar siempre, por darme tanto en el tiempo que compartimos.

A León, mi hijo, por tanto amor y felicidad, por enseñarnos tanto o más de lo que le enseñamos, y por esa luz que ojalá te siga acompañando siempre. A Vero, mi compañera de vida, mi amiga, mi apoyo, mi descanso, mi amor. Desde que nos conocimos y enamoramos, crecimos y compartimos todo: casa, hogar, felicidad, angustias, sueños, reflexiones (interminables!), pérdidas, viajes, e hicimos nuestra obra más linda.



## RESUMEN

---

Un modelo de optimización plantea que el estado observado de un sistema se produce porque dicho estado posee un valor máximo o mínimo de una función objetivo dentro de los estados posibles en que se podría encontrar el sistema. El ejemplo más emblemático de un fenómeno de optimización en la biología consiste en una población que se enriquece progresivamente en aquella variante fenotípica que posee una mayor aptitud biológica, y constituye el caso paradigmático de evolución por selección natural. Diversos modelos de optimización se encuentran en la base de una gran parte de la biología teórica del siglo XX, y han tenido en consecuencia un papel destacado en áreas de la biología tan diversas como la biología evolutiva, la ecología, la fisiología y, de forma más reciente, la biología de sistemas. Varios de los debates más intensos en la historia de la biología han estado relacionados de una manera u otra al concepto de optimización y, a pesar de que en reiteradas oportunidades han sido severamente criticados, el número de estudios que involucran modelos de optimización ha continuado en aumento. La persistencia en el uso de estos modelos puede atribuirse a que representan un abordaje general a uno de los interrogantes más importantes de la biología: en qué medida la estructura y dinámica de los sistemas biológicos deriva de principios generales de organización. A pesar de algunos avances recientes, no existe en la actualidad una metodología general para la evaluación empírica de modelos de optimización, y esto representa un impedimento para el avance de este campo del conocimiento. Este trabajo de tesis busca establecer si es posible evaluar de forma general el sustento empírico de un modelo de optimización mediante métodos estándar de inferencia estadística. Luego de una introducción general a la temática, en el Capítulo 2 se presenta un trabajo de revisión bibliográfica con el fin de analizar cuál es en la actualidad la relevancia de las críticas planteadas a la aproximación teórico-metodológica de la optimización en la biología. En primer lugar se identificaron críticas dirigidas tanto a la fundamentación teórica de los modelos como a los métodos utilizados en su evaluación empírica. Se encontró que existen posiciones con asiento en escuelas de pensamiento filosófico que brindan sustento a los modelos de optimización, aún en el caso de críticas a sus fundamentos teóricos en las que no se ha llegado a un consenso. En el resto de las críticas identificadas se encontró que existe un consenso acerca de que es efectivamente necesario corregir las falencias señaladas. Con el objetivo de evaluar cuál es el estado de las prácticas actuales en la formulación y evaluación de modelos de optimización se analizó una selección de publicaciones recientes sobre la modelización del metabolismo celular, un área en la que se han publicado un gran número de trabajos en los últimos años. Los resultados muestran que las deficiencias

en la evaluación de los modelos son todavía comunes en la actualidad. En el Capítulo 3 se presenta una metodología basada en la teoría axiomática de la probabilidad que permite una evaluación de modelos de optimización exenta de las limitaciones de los métodos tradicionalmente utilizados. La metodología propuesta tiene como punto de partida una generalización de los modelos de optimización y el planteo de un modelo nulo, lo que hace posible su expresión en la forma de un modelo probabilístico bajo ciertas condiciones generales y permite, en consecuencia, analizar de forma objetiva el soporte empírico de un modelo de optimización mediante una inferencia estadística. Al estar planteada de una manera general con asiento en la teoría matemática de la medida, la metodología desarrollada puede aplicarse a modelos de optimización de sistemas biológicos independientemente de su nivel de organización o de la complejidad del modelo matemático subyacente. En el Capítulo 4 se presenta una primera aplicación de la metodología desarrollada a modelos de forrajeo óptimo, uno de los campos de la biología donde se ha trabajado con modelos de optimización con mayor énfasis. Se reanalizaron los modelos planteados por Belovsky (1986b) para explicar la composición de la dieta de campo de 14 especies de herbívoros, donde los parámetros del modelo son constantes, y también una versión no determinista donde los parámetros poseen una distribución de probabilidad. Los resultados fueron, en líneas generales, concordantes con los de Belovsky (1986b), pero al ser producto de métodos estadísticos estándar permiten realizar una inferencia que carece de las deficiencias que varios autores señalan respecto del trabajo original. El análisis del modelo nulo mostró además que la dieta de una de las especies resulta incompatible con el modelo planteado, algo que no es mencionado en el trabajo original ni en otras publicaciones que también analizan los modelos de este estudio. Finalmente, se presenta una discusión general de los aportes de este trabajo y de las potenciales contribuciones al estudio de procesos de optimización como principios generales de organización de los sistemas biológicos.

## ÍNDICE GENERAL

---

1	INTRODUCCIÓN GENERAL	1
1.1	Relevancia de la optimización en las ciencias biológicas.....	1
1.2	Controversias y desafíos en la aplicación de optimización en biología.....	4
1.3	Definiciones.....	6
1.4	Objetivos.....	8
1.5	Estructura de la tesis.....	8
2	REVISIÓN DE FUNDAMENTOS Y EVALUACIÓN DE MODELOS DE OPTIMIZACIÓN	11
2.1	Introducción.....	12
2.2	Métodos.....	15
2.3	Críticas al uso de modelos de optimización.....	18
2.3.1	Críticas a la fundamentación teórica.....	18
2.3.2	Críticas a la evaluación del sustento empírico.....	25
2.4	Evaluación de modelos de optimización en aplicaciones recientes.....	28
2.5	Discusión.....	38
3	GENERALIZACIÓN Y EVALUACIÓN DE MODELOS DE OPTIMIZACIÓN	43
3.1	Introducción.....	43
3.2	El rol de las restricciones en los modelos de optimización.....	47
3.3	Generalización del problema de optimización.....	50
3.4	Introducción a la teoría de la medida.....	52
3.5	Formulación de un modelo nulo probabilístico.....	63
3.5.1	Definición de un espacio de medida estándar sobre el espacio muestral.....	64
3.5.2	Modelo nulo como una medida sobre el espacio muestral.....	65
3.5.3	Definición de una medida de probabilidad.....	73
3.6	Formulación de un modelo general probabilístico de optimización.....	81
3.6.1	Construcción de un modelo general de optimización.....	82
3.6.2	Definición de una medida sobre el espacio muestral.....	89
3.6.3	Definición de un espacio de probabilidad.....	93
3.7	Inferencia estadística.....	96
3.7.1	Hipótesis de optimización y predicciones.....	97
3.7.2	Función de verosimilitud.....	100
3.7.3	Cuantil de la función objetivo como variable aleatoria.....	104

3.7.4	Métodos de inferencia estadística . . . . .	107
3.8	Discusión . . . . .	115
4	ApLICACIÓN DE lA METODOlOGÍA GENERAL A MODElOs DE FORRAJEo óPTIMO . . . . .	119
4.1	Introducción . . . . .	120
4.2	Metodología . . . . .	122
4.3	Resultados . . . . .	127
4.4	Discusión . . . . .	136
5	discUSIÓN GENERAl . . . . .	141
A	ApÉNDICES . . . . .	145
A.1	Capítulo 4 . . . . .	145
A.1.1	Implementación metodología general . . . . .	145
A.1.2	Específicas para problemas de programación lineal . . . . .	167
A.1.3	Específicas del análisis de Belovsky (1986b) . . . . .	171
	BIBlIOGRAFÍA . . . . .	187

## ÍNDICE DE FIGURAS

---

Figura 1.1	Evolución de la publicación de artículos originales que aplican modelos matemáticos de optimización en el área del análisis de flujos metabólicos hasta el año 2017 inclusive.....	6
Figura 2.1	Evolución de la publicación de trabajos científicos que citan a Gould y Lewontin (1979) hasta el año 2017 inclusive.....	13
Figura 2.2	Ejemplos de figuras de algunos de los artículos analizados que son utilizadas en la evaluación subjetiva de predicciones de modelos de optimización. ....	33
Figura 4.1	Distribución de la composición de la dieta esperada de <i>Bison bison</i> en base al modelo nulo de acuerdo a las restricciones planteadas por Belovsky (1986b).....	126
Figura 4.2	Densidad de Radon-Nikodym de la composición de dieta predicha por la versión determinista del modelo nulo y de los modelos de optimización de máxima energía ingerida y mínimo tiempo de forrajeo para valores de $q$ de 0,5 y 0,95.....	128
Figura 4.3	Verosimilitud de los modelos de forrajeo óptimo para cada especie analizada en función del parámetro $q$ de la versión determinista de los modelos. ....	129
Figura 4.4	La variación de la verosimilitud de los modelos de forrajeo óptimo en función del parámetro $q$ con un mismo valor para el conjunto de las especies muestra un fuerte soporte a la hipótesis de máxima energía ingerida.....	131
Figura 4.5	Densidad de Radon-Nikodym de la composición de dieta de la versión no determinista del modelo nulo y los modelos de forrajeo óptimo para valores de $q$ de 0,5 y 0,95.....	134
Figura 4.6	Verosimilitud de los modelos de forrajeo óptimo para cada especie analizada en función del parámetro $q$ de la versión no determinista de los modelos. ....	135
Figura 4.7	Variación de la verosimilitud de la versión no determinista de los modelos de forrajeo óptimo en función del parámetro $q$ para el mismo grupo de especies analizado en la versión determinista. ....	136
Figura 4.8	Verosimilitud de la versión no determinista de los modelos de forrajeo óptimo en función del parámetro $q$ para la totalidad de las especies incluida <i>Dissosteira carolina</i> .....	138

## ÍNDICE DE TABLAS

---

Tabla 2.1	Datos bibliométricos de los 20 artículos analizados en el campo de análisis de flujos metabólicos publicados en el período 2007-2016; las citas recibidas se circunscriben al mismo período.....	29
Tabla 2.2	Breve reseña de los principales aportes de la selección de artículos analizados en el campo de análisis de flujos metabólicos en el período 2007-2016 .....	31
Tabla 2.3	Caracterización de los modelos de optimización planteados en los artículos analizados en el campo de análisis de flujos metabólicos publicados en período 2007-2016 .....	35
Tabla 2.4	Análisis de la metodología de evaluación empírica de modelos de optimización en artículos seleccionados en el campo del análisis de flujos metabólicos.....	37
Tabla 4.1	Descripción de los componentes de los modelos de optimización de maximización de la energía adquirida y minimización del tiempo de forrajeo.....	124
Tabla 4.2	Resultados de los métodos de inferencia estadística aplicados a la versión determinista del modelo de máxima ingesta de energía para cada especie de herbívoro de forma independiente.	130
Tabla 4.3	Resultados de los métodos de inferencia aplicados al reanálisis de modelos de forrajeo óptimo para el conjunto de las especies de herbívoros que cumplen con las restricciones planteadas por Belovsky (1986b). .....	132

## INTRODUCCIÓN GENERAL

---

### 1.1 RELEVANCIA DE LA OPTIMIZACIÓN EN LAS CIENCIAS BIOLÓGICAS

La mayoría de los científicos busca explicar además de describir y predecir regularidades empíricas, y con frecuencia dichas explicaciones involucran argumentos de optimización (Schoemaker 1991). En términos simples, se dice que existe optimización en un sistema dado cuando el estado observado posee un valor mínimo o máximo para un cierto criterio con respecto al conjunto de estados posibles en que se podría encontrar el sistema. La teoría de evolución por selección natural propuesta por Darwin (1859) constituye el ejemplo más emblemático de optimización en la biología. En términos modernos, para que en una población ocurra evolución por este mecanismo deben cumplirse tres condiciones: variabilidad fenotípica, heredabilidad y reproducción diferencial (Lewontin 1978). La condición de reproducción diferencial se verifica cuando existen diferencias en la aptitud biológica (es decir, en el valor promedio de descendientes que producen) entre individuos que poseen diferentes variantes fenotípicas. En caso cumplirse las tres condiciones mencionadas, la teoría predice que la población se enriquecerá, generación tras generación, en aquella variante fenotípica que posea una mayor aptitud biológica. Es clara la identidad entre una optimización y esta descripción de la evolución por selección natural: después de un número suficiente de generaciones, la variante fenotípica observada en la población será aquella que posea un valor máximo de la aptitud biológica respecto de las variantes fenotípicas preexistentes, que constituyen los estados alternativos posibles.

Dado que la selección natural es considerada con un amplio consenso como la principal fuerza determinista en la evolución biológica y es caracterizada como un proceso de optimización (e.g. Haldane 1957; Rosen 1967; Cody 1974; Yoshimura y Shields 1987; Parker y Maynard Smith 1990; Richardson 1994), este concepto ha desempeñado un rol central en múltiples áreas de la biología: diversos argumentos de optimización se encuentran en la base de una gran parte de la biología teórica. Si bien esta visión de la evolución por selección natural está expresada en la obra de Darwin (1859), su popularización puede atribuirse al impacto del trabajo pionero de Wright (1931) y especialmente de Fisher (1930) en genética de poblaciones. Fisher demostró por primera vez el célebre Teorema Fundamental de la selección natural, del que se obtiene como corolario que la aptitud biológica de una población bajo ciertas condiciones tenderá a

alcanzar un valor máximo si las condiciones ambientales se mantienen constantes <sup>1</sup>. Si bien Fisher se abstiene de darle a su teorema el estatus de un principio variacional a la manera tradicional de la física, él mismo es quien lo compara con la Segunda Ley de la termodinámica (Fisher 1930, p. 39). A partir de la segunda mitad del siglo XX el desarrollo de la matemática de la optimización en el ámbito de la ingeniería sirvió de inspiración y estímulo para la aplicación de modelos matemáticos basados en optimización en la biología, mientras que el avance de las ciencias de la computación permitió el abordaje de modelos progresivamente más complejos (Oster y Wilson 1979, capítulo 8).

El estudio de los sistemas biológicos mediante un abordaje basado en la optimización ha sido especialmente prolífico a un nivel de organización individual. En el campo de la biología evolutiva y la ecología, por ejemplo, los modelos matemáticos de optimización tienen un papel central en la teoría de forrajeo (Emlen 1966; MacArthur y Pianka 1966; Schoener 1969; Werner y Hall 1974; Charnov 1976a, 1976b; Pyke y col. 1977; Belovsky 1986b; Mangel y Clark 1986; Stephens y Krebs 1986; Werner y Anholt 1993; Robinson y Wilson 1998; Plath y Boersma 2001; Simpson y col. 2004; Plank y James 2008; Jensen y col. 2012) y en la teoría de evolución de historias de vida (Smith y Fretwell 1974; Brown y col. 1993; Charnov 1993; Koziowski y Weiner 1997; Cichon y Kozłowski 2000; Hatchwell y Komdeur 2000; Roff 2002; Kozłowski y col. 2004; Charnov y Ernest 2006; Öckinger y col. 2010; Williams y col. 2010; Flatt y Heyland 2011; Rytwinski y Fahrig 2012). En el campo de la fisiología, por otra parte, han sido utilizados para explicar la alometría existente entre tamaño corporal y tasa metabólica (West y col. 1997; Banavar y col. 1999; West y col. 1999; Dodds y col. 2001; Kozłowski y col. 2003; Glazier 2005), y en el análisis del metabolismo y productividad en vegetales (Horn 1971; King 1990; Makela y col. 2002; Anten 2005; Katul y col. 2010; Anten y During 2011; Valentine y Mäkelä 2012). De forma más reciente, han sido empleados en el análisis y predicción de redes metabólicas en el marco de la metodología de análisis de balance de flujo (Edwards y col. 2002; Kauffman y col. 2003; Price y col. 2003; Reed y Palsson 2003; Lee y col. 2006; Raman y Chandra 2009; Bull y Wang 2010; Orth y col. 2010; Velenich y Gore 2012; Harcombe y col. 2013). Los modelos matemáticos utilizados en la gran mayoría de estos trabajos representan o bien una hipótesis de maximización de la aptitud biológica, o bien de una propiedad que se considera una medida indirecta de ésta; por ejemplo, el uso de una hipótesis de maximización de la velocidad de adquisición de energía en el marco de la teoría de forrajeo óptimo (Parker y Maynard Smith 1990).

La aplicación de modelos de optimización también se ha dado en niveles de organización superiores al individual, y en este caso se pueden distinguir dos líneas de trabajo de acuerdo a la naturaleza de la propiedad que se hipotetiza sujeta a una op-

<sup>1</sup> Fisher (1930) demostró, específicamente, que la velocidad de cambio de la aptitud biológica en una población es igual al componente génico de su varianza para el caso particular de un gen con dos alelos

timización. Una línea de trabajo ha consistido en la aplicación de la selección natural más allá del nivel individual, y es más conocida por el debate clásico respecto de la selección de grupo y la selección de parentesco (Wynne-Edwards 1962; Hamilton 1964a, 1964b; Maynard Smith 1964; Williams 1966, 1972; Wilson 1975, 1983; Grafen 1984; Wilson y Sober 1994; Leigh 2010). En los últimos años ha tenido lugar una nueva síntesis que parece ampliar definitivamente el horizonte de aplicabilidad de la selección natural a niveles superiores al individual (e.g. Goodnight 2005; Traulsen y Nowak 2006; Godfrey-Smith 2008; Gardner y Grafen 2009; Okasha 2009; Massimo Pigliucci 2010). En este sentido, cabe señalar que en la última edición de su influyente libro sobre biología evolutiva Williams (1966) rechaza explícitamente la arraigada noción de que la selección a niveles superiores al individual debe ser descartada. Sin embargo, esta línea de investigación se encuentra todavía ligada a la controversia en torno de la selección de grupo (e.g. Eldakar y Wilson 2011), y es en esta temática en donde se enfocan la mayoría de los trabajos, salvo interesantes excepciones (e.g. Goodnight 2000; Swenson y col. 2000a; Swenson y col. 2000b; Whitham y col. 2003, 2006; Williams y Lenton 2007). Es probable que el debate en torno a la selección de grupo, sumado a la dificultad en la definición y estudio de unidades de variación y herencia en niveles superiores al individual, sean responsables de un desarrollo relativamente limitado en este campo.

De forma paralela, se ha trabajado con hipótesis de optimización basadas en fundamentos de índole termodinámica (e.g. Fath y col. 2001; Jørgensen y Fath 2004; Ray 2006; Yen y col. 2014; Chapman y col. 2015; Skene 2017). Este marco teórico fue popularizado por el trabajo pionero de Lotka (1922a, 1944), quién enunció la hipótesis de la máxima potencia como un principio de organización general de los sistemas biológicos, y es notorio que Lotka señalase que esta idea fue planteada originalmente nada menos que por Boltzmann (1886), quien sentó las bases de la mecánica estadística. Esta línea de investigación ha sido aplicada especialmente a un nivel de organización ecosistémico, y se han planteado en este marco varios principios de optimización alternativos (Prigogine y Wiame 1946; Margalef 1963; Jørgensen y Mejer 1979; Ulanowicz y Hannon 1987; Odum 1988; Schneider y Kay 1994). En los últimos años este enfoque ha despertado un renovado interés, particularmente en torno del principio de máxima producción de entropía (e.g. Martyushev y Seleznev 2006; Dewar 2010; Kleidon y col. 2010; Martyushev 2010; Meysman y Bruers 2010; Vallino 2010; Chapman y col. 2015; Vallino y Algar 2016). Aunque el enfoque termodinámico ha dominado los estudios basados en optimización a nivel ecosistémico, ha tenido un escaso contacto con la línea dominante desarrollada a nivel individual y de grupo en el marco de la biología evolutiva (algunas notables excepciones son Weber y col. 1989; Brown y col. 1993; Depew y Weber 1996; Weber y Depew 1996; Matthen y Ariew 2002; Hoelzer y col. 2006). Sin embargo, es importante destacar que Lotka (1922b) intentó tempranamente sentar las bases para una unificación de la teoría de la evolución por selección natural y de la termodinámica, y muchos autores han sostenido con posterioridad que un enfoque

integrado podría sentar las bases para hallar principios de organización comunes a todos los niveles de organización biológica (e.g. Weber y col. 1989; Hoelzer y col. 2006; Batten y col. 2008; Pennell y O'Connor 2017; Skene 2017).

## 1.2 CONTROVERSIAS Y DESAFÍOS EN LA APLICACIÓN DE OPTIMIZACIÓN EN BIOLOGÍA

Dada la estrecha relación entre optimización y selección natural, así como la variedad de temáticas que han sido abordadas en este contexto, no resulta extraño que varios de los debates más intensos en la historia de la biología hayan estado relacionados de una manera u otra al concepto de optimización. Se pueden mencionar, entre otros, el extenso debate entre Fisher y Wright en el marco de la teoría de genética de poblaciones (Wright 1929; Fisher 1930; Wright 1931; Fisher 1934; Fisher, Ford y col. 1950; Wright 1951, 1982), cuya influencia continúa hasta la actualidad (e.g. Coyne y col. 1997; Wade y Goodnight 1998; Orr 2005; Pigliucci 2008; Svensson y Calsbeek 2012; Hill 2014); la ya citada controversia en torno de la selección de grupo; y el debate existente respecto de la relación entre el tamaño corporal y la tasa metabólica (e.g. Agutter y Wheatley 2004; Kozłowski y Konarzewski 2004; Brown y col. 2005; O'Connor y col. 2007; Savage y col. 2008; Agutter y Tuszynski 2011; Glazier 2014; Ballesteros y col. 2018). Por este mismo motivo no es llamativo que, en reiteradas oportunidades, se hayan planteado críticas al uso de modelos de optimización en la biología. Especialmente a partir del trabajo clásico de Gould y Lewontin (1979), varios autores han cuestionado diferentes aspectos del uso de hipótesis de optimización (e.g. Levins y Lewontin 1985; Gray 1987; Pierce y Ollason 1987; Ward 1992; Khalil 1997; Pellis 2002; Fetzer 2010).

Quienes consideran que los principios de optimización son una herramienta fundamental para comprender a los sistemas biológicos sostienen asimismo que su utilidad debería ser evaluada de acuerdo a su poder predictivo (e.g. Parker y Maynard Smith 1990; Orzack y Sober 1994a; Gardner 2009; Bull y Wang 2010). Sin embargo, la metodología utilizada para evaluar hipótesis de optimización ha sido en particular objeto de severas críticas (e.g. Huggard 1994; Orzack y Sober 1994b; Owen-Smith 1996). Se han señalado numerosos problemas relacionados con aspectos básicos del método hipotético-deductivo en el trabajo con hipótesis de optimización. En muchos casos los trabajos de optimización en biología parecen orientados a mostrar un acuerdo entre datos empíricos y procesos de optimización antes que a poner a prueba una hipótesis de optimización. De hecho, en varios trabajos el principio de optimización representa un supuesto del trabajo antes que una hipótesis a poner a prueba (e.g. Gould y Lewontin 1979; Kamil 1983; Gray 1987). En el campo de los análisis de balance de flujos metabólicos, por ejemplo, Sánchez, Sáez y col. (2014) plantean que, si bien múltiples principios de optimización han sido propuestos para explicar el metabolismo celular,

típicamente los trabajos en esta área abordan un único principio de optimización en lugar de someter hipótesis alternativas a un contraste cuantitativo. Se puede agregar además que sería cauteloso no asumir que el metabolismo celular se encuentra bajo optimización sino que esto debería ser también evaluado de manera objetiva. Una situación similar existe respecto de las hipótesis de optimización termodinámica, donde un abundante cuerpo de desarrollo teórico y trabajos donde se interpretan datos experimentales contrastan con el bajo número de trabajos dedicados a una evaluación empírica.

Más allá de los cuestionamientos, el número de estudios que involucran hipótesis de optimización ha continuado en aumento y con aplicaciones a nuevos campos de la investigación en biología (Figura 1.1). Parker y Maynard Smith (1990) plantean que la optimización es una de las herramientas más poderosas y elegantes para la comprensión de la adaptación. Dekel y Alon (2005), por su parte, concluyen que un enfoque basado en principios de optimización podría contribuir a una descripción fundamental de la interacción entre el metabolismo, la evolución y el entorno. En suma, el uso ininterrumpido de modelos matemáticos de optimización puede atribuirse a que, a pesar de las críticas recibidas, este enfoque permite abordar uno de los interrogantes más importantes de la biología: en qué medida la estructura y dinámica de los sistemas biológicos deriva de principios generales de organización.

Las problemáticas previamente señaladas se pueden relacionar con la metodología comúnmente usada en trabajos de optimización con predicciones cuantitativas. De manera general, se parte de un modelo que asume que ciertas restricciones operan sobre el sistema bajo estudio y, a partir de un principio de optimización, se deriva mediante un método matemático un estado óptimo. La relativa facilidad con la que se obtiene de esta manera una predicción contrasta con la dificultad en su interpretación. Al carecer de un método estándar que permita hacer una inferencia estadística de los resultados hallados, diferentes autores utilizan criterios disímiles, lo que dificulta la comparación entre trabajos similares dada la dispersión de los métodos utilizados (Orzack y Sober 1996). En una revisión reciente, Yen y col. (2014) plantean que un punto clave para el avance de la aplicación de principios de optimización con fundamentos termodinámicos en el campo de la ecología es la generación de predicciones cuantitativas que permitan una validación empírica. A pesar de algunos avances recientes en el campo del análisis del metabolismo celular en el contexto de la biología de sistemas (e.g. Burgard y Maranas 2003; Pérez-Escudero y col. 2009; Wintermute y col. 2013; Zhao y col. 2016), no existe en la actualidad un consenso sobre una metodología general que permita contrastar hipótesis de optimización en el marco de una inferencia estadística formal. La falta de una metodología con tales características parece representar un impedimento para el avance del conocimiento sobre el grado en que los procesos de optimización influyen los sistemas biológicos.

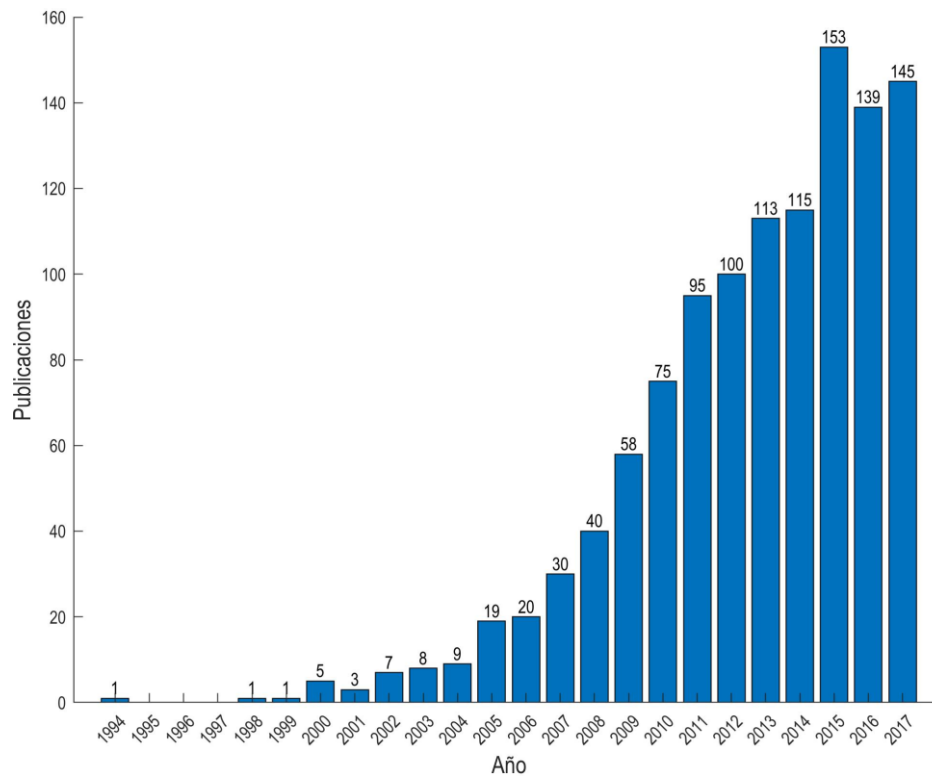


Figura 1.1: Evolución de la publicación de artículos originales en el área del análisis de los flujos metabólicos que aplican modelos matemáticos de optimización hasta el año 2017 inclusive. Datos obtenidos de la base de datos bibliográfica Web of Science (Reuters 2019) para la búsqueda “flux balance analysis” OR (“flux balance” OR “flux analysis”) AND (metabolism OR metabolic) AND “objective function”, lo que arrojó 1137 artículos originales en total. Para evaluar que las publicaciones resultantes sean efectivamente de la temática de interés se revisó el texto completo de 25 trabajos al azar; el 100 % de estos cumplió con el criterio de analizar y/o aplicar modelos matemáticos de optimización.

### 1.3 DEFINICIONES

En los textos científicos en lengua inglesa se han utilizado múltiples términos como sinónimos o para referirse a conceptos estrechamente relacionados: *optimization principles*, *optimality principles*, *extremum principles* y *extremal principles*, en orden decreciente de frecuencia de uso <sup>2</sup>. En el ámbito particular de la biología es común el uso de *optimality*; si bien el término *optimalidad* se ha usado con frecuencia en textos en español,

<sup>2</sup> Se realizaron las siguientes búsquedas en el servicio Google Scholar (<https://scholar.google.com>, 15/07/2017) para evaluar la frecuencia de los términos en la literatura científica, con los siguientes resultados: “optimization principle” OR “optimization principles” OR “optimisation principle” OR “optimisation principles”, 16600 resultados; “optimality principle” OR “optimality principles”, 13700 resultados; “extremum principle” OR “extremum principles”, 8360 resultados; “extremal principle” OR “extremal principles”, 2860 resultados.

constituye un anglicismo que no está reconocido oficialmente en la lengua española (Real Academia Española 2017). Se ha optado en este trabajo por denominar de manera colectiva como *principios de optimización* a cualquier postulado que afirme que los sistemas tienen una tendencia a alcanzar un estado en particular entre los muchos posibles, aquel donde una función alcanza un valor máximo o mínimo (i.e. óptimo). El término *optimización* se utilizará para designar de manera global el marco teórico y conjunto de herramientas metodológicas utilizadas en el estudio de sistemas materiales basados en principios de optimización. Asimismo, se ha optado por utilizar *hipótesis de optimización* para designar cualquier hipótesis de trabajo que se basa en un principio de optimización, ya sea este un componente fundamental de la hipótesis o un simple supuesto asociado. Se define como *modelo de optimización* la formalización de una hipótesis de optimización en la forma de un constructo matemático utilizado para el planteo general de problemas de optimización en el campo de la matemática y las ciencias de la computación, y del cual pueden derivarse predicciones cuantitativas.

Un modelo de optimización consta en su formulación típica de tres componentes: un espacio muestral, un conjunto de restricciones y una función objetivo. El dominio sobre el que están definidas las restricciones y la función objetivo incluye necesariamente al espacio muestral, pero de forma más general estas funciones pueden estar definidas además sobre un espacio de parámetros que *no* está sujeto a una optimización. Esta formulación es la que se utiliza con mayor frecuencia en el estudio de sistemas biológicos y es la que será tratada en este trabajo de tesis, pero cabe aclarar que existen generalizaciones de este modelo. Cuando sea necesario distinguir el modelo típico de otras formulaciones más generales, nos referiremos al primer caso como modelo de optimización *simple*. Ejemplos de generalizaciones de esta formulación son los modelos multi-objetivo, que poseen más de una función objetivo (Collette y Siarry 2013), y los modelos multi-nivel, que se caracterizan por un conjunto jerárquico de modelos de optimización simples donde el espacio muestral de funciones objetivo de menor jerarquía es parte del espacio de parámetros de las funciones objetivo de mayor jerarquía (Migdalas y col. 2013).

La solución a un problema de optimización se denomina *óptimo global*, y está formada por aquellos elementos del dominio que cumplen con las restricciones y en los que el valor de la función objetivo es óptimo, es decir, representa un máximo o mínimo entre los valores posibles. Se define como *subespacio de soluciones* el conjunto de elementos del dominio que satisfacen las restricciones planteadas para un modelo de optimización dado. Si bien el término *restricción* es de uso común en el lenguaje, se utilizará o bien de acuerdo a la definición precedente, o bien para referirse a factores que sin estar expresados en una forma matemática podrían ser formalizados de esta manera en el contexto de un modelo de optimización.

## 1.4 OBJETIVOS

El objetivo general de este trabajo de tesis es generar conocimiento sobre la aplicación y evaluación de modelos de optimización en las ciencias biológicas.

Los objetivos particulares de esta tesis son:

- Elaborar una síntesis de las críticas a la aproximación de la optimización en biología, evaluar si existe un consenso en cuanto a la validez de dichos cuestionamientos, y analizar su relevancia en las prácticas utilizadas en la actualidad en la formulación y evaluación empírica de los modelos de optimización.
- Establecer las bases de una metodología para la formulación y evaluación empírica de modelos de optimización que sea de aplicación general, carezca de deficiencias señaladas en los métodos comúnmente utilizados, y permita realizar una inferencia estadística basada en la función de verosimilitud.
- Aplicar la metodología desarrollada para evaluar modelos de optimización en el marco de la teoría de forrajeo sobre datos publicados, comparar los resultados de la inferencia con aquellos obtenidos previamente mediante otros métodos, y analizar los efectos que puede tener sobre las conclusiones de trabajos que plantean hipótesis de optimización.

## 1.5 ESTRUCTURA DE LA TESIS

Se describe a continuación la estructura general de este trabajo de tesis. En el Capítulo 2 se presenta una revisión de los principales cuestionamientos que ha recibido la aplicación de modelos de optimización en la biología. Se analizan en primer lugar aspectos de la fundamentación de las hipótesis de optimización en calidad de constructos teóricos. En segunda instancia se realiza una síntesis de las críticas dirigidas a la evaluación del valor empírico de los modelos de optimización. Finalmente se evalúa la relevancia de las críticas en la actualidad a través del análisis de trabajos recientes de referencia.

En el Capítulo 3 se presentan las bases de una metodología que permite realizar una evaluación objetiva de modelos de optimización basada en la generación de un modelo nulo y que carece de las falencias que han sido típicamente señaladas. La metodología presentada permite derivar predicciones en forma de distribuciones de probabilidad tanto para la hipótesis de optimización como para una hipótesis nula carente de optimización y, de esta forma, hace posible aplicar métodos estándar de inferencia estadística. Se aplica esta metodología al reanálisis de un caso de estudio del área de teoría de forrajeo óptimo con el fin de ejemplificar los pasos involucrados

para su uso y permitir una comparación con los métodos que han sido comúnmente utilizados.

En el Capítulo 4 se presenta una primer aplicación de la metodología desarrollada al análisis de modelos de composición de la dieta de herbívoros de Belovsky (1986b) planteados en el marco de la teoría de forrajeo, uno de las campos de la biología donde se ha trabajado con modelos de optimización con mayor énfasis.

Finalmente, en el Capítulo 5 se presenta una discusión general sobre la relevancia de los aportes del presente trabajo para el establecimiento de la evidencia empírica de principios de optimización planteados por distintos autores, el estudio del valor empírico de las idealizaciones de los modelos de optimización, y el análisis de las restricciones que podrían afectar la dinámica y estructura de los sistemas biológicos.



## UNA REVISIÓN CRÍTICA DE LA FUNDAMENTACIÓN TEÓRICA Y EVALUACIÓN DE MODELOS DE OPTIMIZACIÓN

---

### Resumen

Los modelos de optimización han tenido un rol importante en diversas áreas de la biología desde mediados del siglo XX y, al mismo tiempo, han sido objeto de una serie de críticas de diversa índole, especialmente a partir del trabajo clásico de Gould y Lewontin (1979). Si bien gran parte de los cuestionamientos apuntan a los fundamentos y racionalidad de los modelos en su calidad de constructos teóricos, un punto en el que tanto proponentes como críticos coinciden es en la necesidad de una evaluación empírica rigurosa. Sin embargo, la metodología usada para evaluar el ajuste de los modelos también ha sido severamente cuestionada, lo que dificulta arribar a conclusiones sólidas respecto de su poder predictivo.

En este trabajo intentamos abordar cuál es en la actualidad la relevancia de las críticas planteadas a la aproximación teórico-metodológica de la optimización en la biología. En primera instancia buscamos identificar mediante una revisión bibliográfica aspectos de los modelos de optimización que han sido criticados en la literatura, tanto a nivel de su fundamentación teórica como de la evaluación del sustento empírico. En cada caso evaluamos si existe un consenso respecto de la relevancia de los cuestionamientos, así como acerca de las prácticas que deberían seguirse para subsanar los problemas señalados. En una segunda instancia analizamos las prácticas utilizadas en la evaluación empírica de los modelos de optimización en una selección de 20 de los trabajos más citados en el campo del análisis del metabolismo celular, un área de aplicación reciente que cuenta con un gran número de publicaciones.

Encontramos que la fundamentación teórica de los modelos de optimización ha sido criticada respecto de los supuestos ontológicos utilizados, la exclusión de mecanismos causales, el tratamiento de la optimización como un supuesto antes que como una hipótesis de trabajo, y el uso de supuestos de optimización y de monotonía estricta. A nivel de la evaluación del sustento empírico identificamos las siguientes problemáticas: evaluación subjetiva de las predicciones, falta de hipótesis alternativas, y ausencia de hipótesis nula. Con respecto a las prácticas actuales en la evaluación empírica de los modelos, encontramos que la totalidad de artículos analizados adolecen de deficiencias. La mayoría de los trabajos realiza únicamente una evaluación subjetiva de las predicciones (65 %) y/o carece de hipótesis alternativas (60 %). Solo un artículo plantea una hipótesis nula respecto de la optimización, pero posee igualmente falencias en el análisis estadístico.

A pesar de que la aplicación de principios de optimización tiene una popularidad sostenida en la biología y continua en aumento, los resultados obtenidos sugieren que las críticas dirigidas a la evaluación empírica de los modelos mantienen su vigencia en la actualidad. Se espera que este trabajo de síntesis pueda servir a la hora de diseñar y

evaluar estudios que tengan como objetivo analizar el soporte empírico de modelos de optimización en sistemas biológicos.

## 2.1 INTRODUCCIÓN

Los modelos matemáticos basados en principios de optimización han tenido un rol importante en la biología desde mediados del siglo XX, y varios autores han planteado que el estudio de los sistemas biológicos mediante un enfoque de optimización representa el abordaje más prometedor para otorgar a la biología de poder predictivo (e.g. Rosen 1967; Parker y Maynard Smith 1990; Dekel y Alon 2005; Sutherland 2005). En la actualidad, el planteo de hipótesis de optimización en la forma de modelos matemáticos es una práctica que se utiliza no solo en áreas donde su uso se popularizó hace varias décadas, como en el caso de la teoría de forrajeo y la teoría de historias de vida, sino que su aplicación continúa extendiéndose a nuevos campos de la biología. De manera reciente, por ejemplo, se ha empleado en el análisis y predicción del metabolismo celular en base a información obtenida mediante técnicas de biología molecular en el marco de la llamada biología de sistemas (e.g. Edwards y col. 2002; Kauffman y col. 2003; Price y col. 2003; Reed y Palsson 2003; Lee y col. 2006; Raman y Chandra 2009; Bull y Wang 2010; Orth y col. 2010; Velenich y Gore 2012; Harcombe y col. 2013). A pesar de tener un uso cada vez más extendido, o quizás debido precisamente a este hecho, el uso de hipótesis de optimización en biología ha estado sujeto a una serie de críticas de diversa índole. Este aparente contraste invita a examinar cuál es la relación entre los aspectos de los modelos de optimización que han sido cuestionados y su uso en la actualidad.

Al examinar el origen de las críticas a los modelos de optimización, la influencia de Gould y Lewontin (1979) se destaca de forma notoria: desde su publicación ha resultado una referencia casi obligada no solo para quienes critican el uso de estos modelos, sino también para quienes intentan refutar dichos cuestionamientos. Los autores de esta publicación desarrollan un severo cuestionamiento del abordaje dominante en biología evolutiva, basado en la aplicación de modelos matemáticos en los que la selección natural se trata como un agente de optimización, y al que los autores denominan programa adaptacionista o simplemente adaptacionismo. Es importante destacar que las críticas al adaptacionismo preceden a dicho trabajo; Williams (1966), por ejemplo, afirma que gran parte de su libro está destinado a la crítica de lo que considera un uso injustificado del concepto de adaptación, y esta obra tuvo una gran influencia en el desarrollo de la biología evolutiva en su forma contemporánea. Sin embargo, el artículo de Gould y Lewontin tuvo una repercusión inmediata, probablemente debido a una combinación de factores que lo hacen difícil de ignorar: la popularidad de las explicaciones basadas en el concepto de optimización, la reconocida trayectoria de ambos

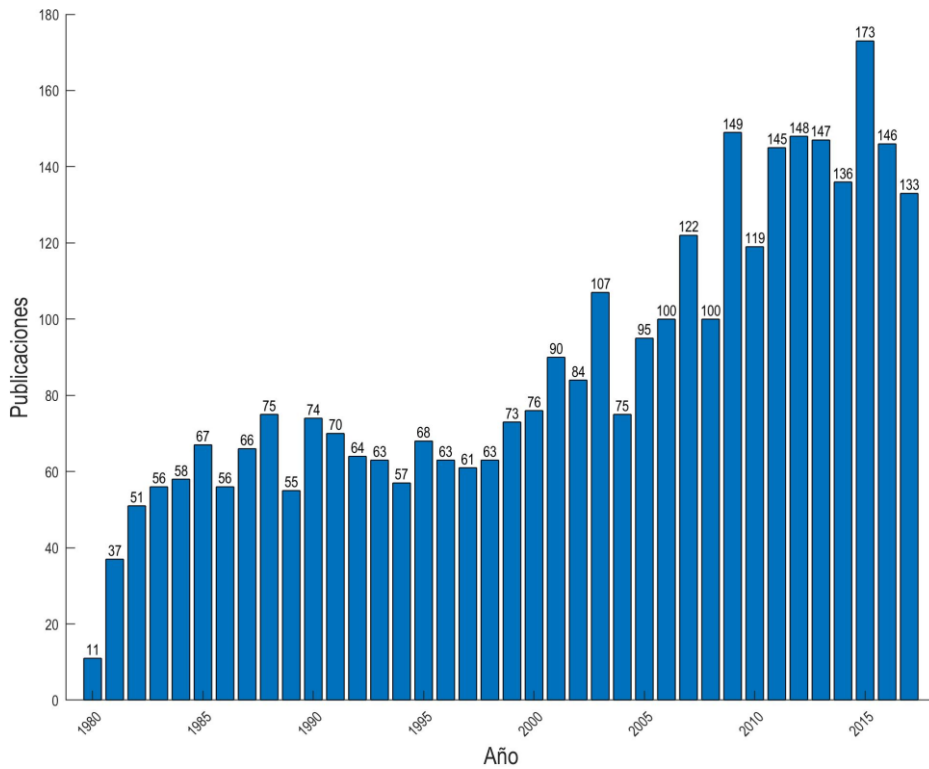


Figura 2.1: Evolución de la publicación de trabajos científicos que citan a Gould y Lewontin (1979) hasta el año 2017 inclusive. Datos obtenidos de la base de datos bibliográfica Web of Science (Reuters 2019).

autores y una redacción tan elegante como provocadora. Con el paso del tiempo, este trabajo se ha convertido en un clásico cuya influencia ha sido destacada en múltiples oportunidades (e.g. Beatty 1987; Parker y Maynard Smith 1990; Reeve y Sherman 1993; Lauder 1996; Arthur 2001; Andrews y col. 2002a; Stoltzfus 2012; Linde-Medina y Diogo 2014), y su impacto perdura hasta la actualidad (Figura 2.1).

A partir de la publicación de Gould y Lewontin (1979) el estudio de los sistemas biológicos a través de un enfoque de optimización ha sido objeto de un extenso debate. Las opiniones sobre la utilidad de esta aproximación han variado entre extremos: mientras que algunos aseguran que representa una de las herramientas más poderosas y elegantes para comprender fenómenos biológicos (Parker y Maynard Smith 1990), otros han llegado a sostener que debería ser abandonada por completo (Pierce y Ollason 1987). Una cantidad considerable de trabajos ha realizado un análisis crítico del uso de modelos de optimización bajo el formato de revisiones y artículos de opinión, así como parte de números especiales y libros dedicados al tema (e.g. Dupré 1987; Orzack y Sober 2001). Una parte de las críticas cuestionan la racionalidad y fundamentos de los modelos de optimización en sí, es decir, en su calidad de constructos teóricos,

independientemente de su poder predictivo. Un punto en el que tanto críticos como defensores del enfoque de optimización en la biología coinciden es en la necesidad de una evaluación rigurosa del sustento empírico de los modelos matemáticos formulados. Sin embargo, la metodología aplicada comúnmente en la evaluación de los modelos ha sido justamente uno de los puntos cuestionados en repetidas oportunidades (e.g. Gray 1987; Orzack y Sober 1994b; Olson y Arroyo-Santos 2015), lo que dificulta arribar a conclusiones respecto de su valor empírico. Este es un aspecto especialmente relevante dado que, más allá del sustento teórico que posea, si la hipótesis que subyace a un modelo matemático carece de poder predictivo al ser evaluada en repetidas ocasiones debería ser desechada, o sufrir al menos una profunda revisión. A la inversa, si se da el caso opuesto y se concluye que efectivamente un modelo posee sustento empírico esto no implica que su fundamentación teórica sea correcta. Sin embargo, las problemáticas asociadas a la evaluación empírica de los modelos de optimización no han sido en general abordadas con el mismo énfasis con el que se han tratado aspectos referidos a la fundamentación de las hipótesis en sí.

Es su gran mayoría, las críticas a los modelos de optimización se circunscriben a un campo particular de aplicación, como la teoría de forrajeo óptimo (e.g. Pyke 1984; Gray 1987; Pierce y Ollason 1987; Ward 1992, 1993; Sih y Christensen 2001), la fisiología vegetal (e.g. Anten y During 2011), la ecología del comportamiento (e.g. Mayhew 1997), o el análisis de metabolismo celular (e.g. Bull y Wang 2010). El disponer de análisis críticos enfocados en aplicaciones específicas no parece haber contribuido a la obtención de conclusiones generales que trasciendan un área de estudio específica o la naturaleza del sistema biológico bajo estudio. La situación descrita no solo dificulta el reconocimiento de problemáticas comunes, sino que también disminuye las chances de que nuevas herramientas metodológicas desarrolladas en un cierto ámbito de aplicación sean aplicadas al estudio de sistemas biológicos en otros campos de la biología.

Dado el sostenido interés en los modelos de optimización en la biología, creemos que es relevante explorar cuál es en la actualidad la relevancia de las críticas planteadas a esta aproximación teórico-metodológica a cuarenta años de publicado el trabajo de Gould y Lewontin. ¿Representan acaso las críticas realizadas verdaderos impedimentos para el avance del conocimiento, o se trata de asuntos con una importancia únicamente histórica, sin relevancia en la actualidad? ¿Es la popularidad actual del enfoque de optimización una indicación de que se han resuelto o superado los aspectos cuestionados, o que, por el contrario, estos han sido simplemente ignorados? El presente capítulo aborda estos interrogantes a través de una metodología de revisión bibliográfica de acuerdo a la estructura que se describe a continuación. La primer sección tiene como objetivos identificar aspectos de la fundamentación y evaluación empírica de modelos de optimización en biología que han sido criticados en la literatura, así como evaluar si en cada caso existe un consenso establecido acerca de la validez de las críticas y, si es este el caso, qué lineamientos o prácticas deberían seguirse para

subsanan las falencias señaladas. La segunda sección, por su parte, tiene como objetivo examinar cuáles son las prácticas utilizadas en la evaluación de modelos de optimización en artículos recientes en el campo del análisis del metabolismo celular, un ámbito donde la aplicación de estos modelos se ha convertido en una técnica estándar y que cuenta con un gran número de publicaciones en la última década.

## 2.2 MÉTODOS

Con el fin de encontrar publicaciones que realicen un análisis crítico del uso de modelos de optimización se siguió la metodología de búsqueda bibliográfica prospectiva y retrospectiva propuesta por Webster y Watson (2002), con el trabajo de Gould y Lewontin (1979) como punto de partida. Esta metodología combina la revisión de una selección inicial de trabajos con un proceso iterativo que consta de dos etapas: un paso prospectivo de búsqueda de trabajos que citen dichos artículos y un paso retrospectivo de revisión de trabajos relevantes citados. Se optó por utilizar este método debido a que la realización de una revisión bibliográfica exhaustiva no resulta factible dada la diversidad de áreas temáticas de la biología en las que se han utilizado modelos de optimización, así como a la naturaleza usualmente multidisciplinaria de dichos estudios. En los artículos encontrados se buscó identificar críticas que hayan sido tratadas por varios autores y cuya relevancia fuese independiente del campo disciplinar específico de aplicación del modelo de optimización, de forma tal que su análisis sea aplicable a la aplicación de modelos de optimización en general. A fin de organizar el tratamiento de la información y facilitar la lectura, cada crítica identificada se analiza en apartados independientes a los que se les intentó dar una estructura común que se describe a continuación. En primer lugar se presenta una breve introducción al aspecto de los estudios que se ha sido cuestionado. A continuación se realiza una síntesis de los cuestionamientos planteados al respecto y una breve descripción de su fundamentación. En caso de haberse encontrado en la revisión bibliográfica publicaciones donde se plantea un rechazo a las críticas, se exponen y analizan a su vez los argumentos de estos autores. Por último, se intenta establecer si se ha arribado a un consenso en el tema, y se evalúa si a partir de lo expuesto pueden derivarse recomendaciones concretas para el planteo y/o evaluación de las hipótesis de optimización.

Con el objetivo de analizar las prácticas utilizadas en la actualidad para la evaluación de modelos de optimización, se optó por realizar una revisión bibliográfica sistemática en el área del análisis del metabolismo celular, ya que esta es el área que más desarrollo ha tenido en los últimos 20 años. Se realizó una búsqueda de publicaciones en el período 2007 - 2016 utilizando el servicio bibliográfico *Web of Science* (Reuters, 2017 (accedido el 25 de julio de 2017)) de los siguientes términos: “flux balance analysis” OR (“flux balance” OR “flux analysis”) AND (metabolism OR metabolic)). A fin de

obtener una muestra acotada pero representativa de artículos influyentes se realizó el procedimiento que se describe a continuación. Los resultados de la búsqueda se ordenaron en función del número promedio de citas por año recibidas desde su publicación durante el mismo período (2007 a 2016 inclusive). El número de citas por año es una variable comúnmente utilizada como un indicador de la influencia de una publicación (e.g. Callaham y col. 2002; Uthman y col. 2013). Al utilizarse para comparar trabajos pertenecientes a una temática específica no fue necesario realizar normalización alguna (Hutchins y col. 2016). Se realizó una revisión a texto completo en orden decreciente de promedio citas por año con el objetivo de seleccionar los 20 artículos más influyentes que realizaron una evaluación de predicciones derivadas de uno o más modelos de optimización mediante algún tipo de comparación con datos experimentales. La definición de este criterio de selección se basó en la noción elemental de que el grado en que los datos experimentales dan respaldo a un modelo matemático refleja el apoyo relativo a las hipótesis asociadas (Johnson y Omland 2004). El criterio utilizado permitió incluir tanto a publicaciones que tienen un objetivo manifiesto de realizar una evaluación de una hipótesis de optimización en la forma de un modelo matemático como a aquellas donde la optimización es considerada como un supuesto, dado que en ambos casos la optimización se trata de una hipótesis asociada.

A continuación se describen brevemente los aspectos que se analizaron en cada uno de los trabajos seleccionados y el criterio utilizado en cada caso:

- Principales aportes: las contribuciones más importantes del trabajo de acuerdo a lo expresado por los autores.
- Tipo de modelo de optimización: se registró si se utilizaba un modelo simple que posee una formulación clásica con una única función objetivo, o una versión generalizada de este, como los modelos multi-objetivo (Collette y Siarry 2013) o multi-nivel (Migdalas y col. 2013).
- Justificación de la función objetivo utilizada: se analizó cómo se justifica la función objetivo utilizada en el modelo de optimización. En caso de no encontrarse una justificación explícita, se buscaron en el texto del artículo los términos *evolutio*<sup>1</sup>, *natural selection*, *adaptation* y *fitness* con el fin de evaluar si los autores asumen de forma implícita que la evolución por selección natural optimiza la función objetivo utilizada. Por último, en caso contrario, se asumió que la justificación es similar a la expuesta en las referencias del trabajo.
- Supuesto de optimización estricta: considera al óptimo global como única predicción del modelo de optimización.

<sup>1</sup> Se utilizó el modificador de búsqueda "\*" para incluir una conjunto de términos relacionados como *evolution*, *evolutionary*, etc..

- Supuesto de monotonía estricta: las predicciones se derivan de la optimización de una función objetivo diferente de la que se hipotetiza optimizada en última instancia, y con la que se asume posee una relación monótonamente creciente.
- Evaluación objetiva de predicciones: realiza una evaluación del ajuste de las predicciones mediante el planteo explícito de un criterio *a priori*. En el caso de utilizar métodos estadísticos se consideró que realizan una evaluación objetiva si el criterio de evaluación de los resultados es el utilizado de forma estándar aún si no era explicitado por los autores, por ejemplo un criterio de rechazo para una prueba de significación estadística del 5 %.
- Análisis de sensibilidad: realiza una evaluación de la respuesta de las predicciones de un modelo de optimización a variaciones en los parámetros y/o formulaciones alternativas de sus restricciones y/o función objetivo, siempre que éstas representen la incertidumbre asociada a una implementación de un misma hipótesis de optimización.
- Inferencia estadística: realiza una evaluación objetiva de modelos de optimización mediante algún tipo de método estadístico. En este caso se evaluó además si se utiliza una prueba de significación clásica (Chow 1998), una selección de modelos a través de un criterio derivado de la teoría de la información (Burnham y Anderson 2002), o un análisis en el marco de la estadística bayesiana (Ando 2010).
- Hipótesis alternativas: compara las predicciones de un modelo de optimización con otros modelos alternativos. En caso de plantearse en el trabajo múltiples hipótesis, se registró si se trataba de una formulación alternativa, como modelos dinámicos del sistema por ejemplo, o de modelos de optimización que diferían en la función objetivo y/o en las restricciones. En este último caso, no se considera que un modelo de optimización tenga restricciones alternativas si poseen una misma forma funcional pero cambian los parámetros, lo que sucede por ejemplo en caso de que se realicen predicciones en diferentes condiciones ambientales. En caso de formulaciones alternativas, se consideró que si éstas no alteran la lógica y fundamentación del modelo de optimización en su calidad de hipótesis sino que representan una forma de evaluar una incertidumbre en la correcta implementación de dicha hipótesis, no constituyen modelos alternativos.
- Hipótesis nula: compara la predicción de un modelo de optimización con la obtenida por un modelo alternativo pero que difiere únicamente por no incluir la hipótesis de optimización.

## 2.3 CRÍTICAS AL USO DE MODELOS DE OPTIMIZACIÓN

Las críticas a los modelos de optimización cubren diversos aspectos de la investigación científica. Por una cuestión tanto de organización como de jerarquía de los argumentos involucrados hemos agrupado el análisis de las críticas en dos grandes grupos, de acuerdo a si cuestionan fundamentos teóricos de los modelos de optimización o si se focalizan en su evaluación empírica. Como es esperable, en algunos casos se da un solapamiento en la fundamentación de diferente tipo de críticas. A fin de evitar una repetición innecesaria, los argumentos se presentan en el apartado en el que la temática surge en primera instancia y, en caso de ser necesario, son analizados posteriormente con una mayor profundidad.

### 2.3.1 *Críticas a la fundamentación teórica*

El análisis de críticas a la fundamentación de las hipótesis de optimización conduce invariablemente al abordaje de aspectos filosóficos de la práctica científica. Al igual que en la práctica dominante tanto en la biología como en las ciencias naturales en general, en los trabajos en que se utilizan modelos de optimización no se suelen incluir reflexiones acerca de los fundamentos del propio quehacer científico, menos aún referencias explícitas a cuestiones filosóficas. Como consecuencia, queda en general a cargo del lector interesado la tarea de reconstruir los fundamentos sobre en los que implícitamente se sostiene una hipótesis de optimización y el subsecuente modelo matemático. A este hecho se le agrega la variedad de posiciones que pueden asumirse en el marco de diferentes escuelas de pensamiento filosófico. Por lo tanto, no es extraño que los autores presenten puntos de vista que parecen incompatibles entre sí, o inconmensurables en el sentido planteado por Kuhn (1962). En este sentido, cabe aclarar que está fuera de los objetivos y alcances del presente trabajo realizar un análisis exhaustivo sobre aspectos filosóficos relacionados a las hipótesis de optimización. Se pretende presentar aquí, en cambio, una síntesis de los principales cuestionamientos al estudio de los sistemas biológicos en el marco de la optimización que involucran aspectos a un nivel ontológico y epistemológico.

#### 2.3.1.1 *Supuestos ontológicos*

En el aspecto más fundamental, los modelos de optimización han sido criticados a nivel ontológico. En el ámbito de la filosofía de la ciencia, la ontología de una teoría científica es definida como la clase de entidades teóricas y de procesos que reconoce (Depew y Weber 1996). En otros términos, las teorías científicas realizan dos tipos de compromisos ontológicos en la explicación de la experiencia fenomenológica: acerca

de qué entidades existen, y respecto de cuáles son las fuerzas que actúan sobre dichas entidades (Mitchell 2003).

Las críticas que invocan argumentos ontológicos se han centrado en cuestionar las entidades a las que apelan este tipo de explicaciones antes que a los procesos involucrados, probablemente debido al consenso que existe en la biología acerca de la selección natural como un agente de optimización (e.g. Haldane 1957; Rosen 1967; Cody 1974; Yoshimura y Shields 1987; Parker y Maynard Smith 1990; Richardson 1994). Esta temática se ha tratado de forma extensa dentro de la de la biología evolutiva en el marco del debate sobre las unidades de selección. En su artículo titulado justamente “Las unidades de la selección”, Lewontin (1970) plantea que la teoría de evolución propuesta por Darwin es un esqueleto lógico que resulta un poderoso sistema predictivo para cambios en todos los niveles de organización. En palabras de Wilson (2007), la teoría de evolución por selección natural es general en el sentido de que las condiciones que se consideran necesarias, variación heredable en el *fitness* asociado a un carácter dentro de una población, son en sí neutrales respecto a sobre qué tipo de entidades opera la selección natural. La discusión respecto a las unidades de selección ha sido abordada extensamente en torno a la selección de grupo (Wynne-Edwards 1962; Hamilton 1964a, 1964b; Maynard Smith 1964; Williams 1966; Wilson 1975, 1983; Grafen 1984; Wilson y Sober 1994). De forma sintética, la posición que se impuso considera que los procesos selectivos tienen lugar a nivel de los organismos individuales. Esta fue la postura sostenida por Williams y llevada a un extremo por Dawkins (1976), dando lugar al llamado neo-darwinismo contemporáneo (Gould 1984). Khalil (1997) plantea que la ortodoxia neo-darwinista es reduccionista a nivel ontológico ya que se encuentra anclada en la tradición naturalista, y como consecuencia afirma que unidades de nivel jerárquico superior al individual no pueden ser objeto de optimización.

La postura dominante en biología evolutiva no ha sido compartida en todas las áreas de la biología. En claro contraste, el planteo de optimización en unidades de nivel jerárquico superior al individuo no ha sido objeto de controversia alguna desde un enfoque termodinámico, sino que los artículos en este ámbito usualmente plantean una optimización a nivel ecosistémico (e.g. Cai y col. 2006; Meysman y Bruers 2007; Vallino 2010). El origen de esta postura se puede rastrear en la influencia del trabajo de Odum y Pinkerton (1955) en el ámbito de la ecología, autores influidos a su vez por Lotka (1922a, 1922b), para quién las unidades sujetas a procesos selectivos podían ser de cualquier índole.

De manera reciente varios autores han instado a la adopción de una visión pluralista en el estudio de los sistemas biológicos. Desde un enfoque anclado en la filosofía de la ciencia, Mitchell (2003) argumenta que la complejidad del objeto de estudio de la ciencias biológicas, así como las limitaciones de las representaciones utilizadas, requiere de un modelo de ciencia integrador que permita ontologías pluralistas. En los últimos años ha tomado forma una nueva síntesis evolutiva que parece ampliar definitivamente

te el horizonte de aplicabilidad de la selección natural respecto del neodarwinismo ortodoxo (e.g. Wilson y Sober 1994; Goodnight 2005; Traulsen y Nowak 2006; Godfrey-Smith 2008; Gardner y Grafen 2009; Okasha 2009; Massimo Pigliucci 2010). Si bien la mayoría de los trabajos sobre optimización en niveles superiores al individual trata el caso particular de la selección de grupo, existen también algunos ejemplos en el marco de la biología evolutiva a nivel comunitario y especialmente a nivel ecosistémico (e.g. Goodnight 2000; Swenson y col. 2000a; Swenson y col. 2000b; Whitham y col. 2003, 2006; Williams y Lenton 2007).

Sería arriesgado decir que se ha alcanzado un consenso amplio en la biología respecto de la posibilidad de optimización en niveles superiores al individual, especialmente cuando esta posición parece seguir dominando en los hechos, incluso cuando parte de sus fundamentos en torno a la selección de grupo hayan sido virtualmente descartados (Eldakar y Wilson 2011). Sí podemos afirmar, en cambio, que no existen motivos que justifiquen descartar *a priori* modelos de optimización en niveles de organización por encima del individual esgrimiendo argumentos ontológicos .

### 2.3.1.2 Exclusión de mecanismos causales

Los modelos de optimización suelen apelar a idealizaciones sobre el sistema bajo estudio en lugar de representar de manera explícita una compleja dinámica subyacente (Potochnik 2010). Richardson (1994) denomina a un modelo como *dinámicamente correcto* si refleja los mecanismos causales involucrados en la producción del resultado. Varios autores han criticado el uso de la optimización en sistemas biológicos por este motivo, amparándose en la noción de que una explicación científica válida debe representar, al menos en parte, aquellos mecanismos que dieron lugar al estado actual del sistema bajo estudio.

En el ámbito de la biología evolutiva, en particular, los modelos de optimización han sido criticados por ignorar mecanismos genéticos (Parker y Maynard Smith 1990). Estas críticas provienen en general de la genética de poblaciones, donde el consenso mayoritario sostiene que no posible ignorar los mecanismos de la herencia y expresión génica (Birch 2016). Por un lado, diferentes trayectorias genotípicas pueden llevar a un mismo fenotipo actual, y este ha sido uno de los intereses fundamentales de la biología evolutiva. Un modelo de optimización típico no puede aportar información sobre trayectorias genéticas, y por lo tanto carece de interés para aquellos interesados en este aspecto de la dinámica evolutiva. En palabras de Levins y Lewontin (1985), al carecer de una descripción de la dinámica, estos modelos simplemente asumen que la evolución lleva el sistema a su óptimo. Sin embargo, trabajos clásicos de la genética cuantitativa demostraron que poblaciones sujetas a un proceso de selección natural no necesariamente alcanzan una máxima aptitud biológica promedio. La probabilidad de fijación de un alelo depende tanto de la ventaja selectiva que este otorgue como de

una dinámica estocástica (Kimura 1962), e incluso la fijación puede verse directamente impedida por fenómenos genéticos como la epistasis (Moran 1964). Aún cuando la dinámica de un sistema sea tal que asegure que eventualmente se alcanzará un óptimo global, el tiempo de evolución podría no ser suficiente para alcanzar dicho estado. Este es un punto especialmente importante en el análisis de experimentos manipulativos, donde no puede asumirse que los tiempos de evolución y/o la intensidad del proceso de optimización en el sistema sean suficientes para alcanzar un estado óptimo (Bull y Wang 2010).

En contraste, desde el campo de la filosofía de la ciencia se han elaborado argumentos que afirman lo contrario. En el ámbito de la biología evolutiva, Potochnik (2007) sostiene que la aproximación de la optimización tiene un rol explicativo asegurado cuando el interés se centra en los efectos de la selección natural sobre el fenotipo, y que la inclusión de mecanismos causales tales como la transmisión de la información genética puede conducir a una pérdida de generalidad. En un marco más amplio, Rice (2012, 2015) afirma que la independencia de relaciones causales no es un defecto, sino una virtud de los modelos de optimización. De acuerdo a Rice, la idealización y abstracción propias de estos modelos pueden jugar un rol importante en las explicaciones científicas justamente porque elimina o abstrae los mecanismos causales de un sistema con el fin de capturar las características dominantes de un fenómeno. Por ejemplo, las condiciones iniciales y la trayectoria causal del sistema pueden no ser importantes para comprender por qué ha ocurrido lo que se pretende explicar si diferentes historias causales hubieran llevado a un mismo resultado. De igual manera, Irvine (2015) considera que una característica destacable de las explicaciones que apelan a la optimización es su robustez a las condiciones iniciales. Estos argumentos son similares a los señalados respecto al uso de modelos matemáticos en la física (Batterman 2000, 2001), y se enmarcan en un creciente interés por las así llamadas explicaciones no-causales (Pincock 2015; Lange 2017; Reutlinger 2017a, 2017b). Por lo tanto, desde el punto de vista de la filosofía de la ciencia, los modelos de optimización como explicaciones de fenómenos biológicos no solo poseen sustento aún cuando omiten una descripción de los mecanismos causales involucrados, sino que esto puede ser visto como una fortaleza de esta aproximación.

El uso de una idealización tal como la exclusión de mecanismos causales representa un supuesto de la mayoría de los modelos de optimización. Orzack y Sober (1994b) apuntan que la racionalidad de focalizarse en el fenotipo está sostenida en la noción de que la selección natural superará cualquier restricción genética o del desarrollo en el caso de caracteres que afecten de manera significativa la aptitud biológica; esto no es otra cosa que un supuesto de la hipótesis de trabajo. Parker y Maynard Smith (1990) plantean que, salvo que posean defectos formales, los modelos de optimización no pueden calificarse de erróneos, sino que pueden resultar inapropiados para describir el sistema bajo estudio si están basados en supuestos equivocados. Esta no es otra

que la posición dominante respecto del uso de modelos matemáticos en la biología: por definición se considera que los modelos son aproximaciones que pueden resultar útiles para representar algunos aspectos de la realidad. Si se establece que un modelo posee sustento empírico, esto indica únicamente que los supuestos en los que se basa representan una aproximación aceptable de la realidad. Desde esta óptica se puede afirmar que la exclusión de mecanismos causales, en su calidad de idealización de una realidad más compleja, debería ser evaluada con el mismo criterio que cualquier otro supuesto de un modelo matemático: en función de su poder predictivo.

### 2.3.1.3 Optimización como supuesto

Una de las principales críticas a los trabajos que utilizan modelos de optimización en la biología evolutiva es que tratan a la optimización como un supuesto antes que como una hipótesis de trabajo. Es en este aspecto que el cuestionamiento al programa adaptacionista planteado por Gould y Lewontin orilla la denuncia, algo que ha sido reforzado y ampliado por otros autores (e.g. Kamil 1983; Gray 1987).

Potochnik (2009) sostiene que el uso general de los modelos de optimización difiere del que se les ha dado típicamente en la biología evolutiva. Esta autora plantea que el debate en torno al adaptacionismo se ha centrado en un uso particular de estos modelos: demostrar que el estado actual de un carácter fenotípico es ante todo una consecuencia de un proceso de evolución por selección natural. Este tipo de uso se corresponde con lo que Godfrey-Smith (2001) denomina *adaptacionismo empírico*. Orzack y Sober (1994b) señalan que esta escuela de la biología evolutiva tiene la tesis de que la selección natural es una explicación suficiente para la evolución de un carácter fenotípico bajo estudio, del que se espera que su estado actual sea localmente óptimo. Es esta concepción de la biología evolutiva a la que dirigieron sus críticas Gould y Lewontin (1979), y es en este caso que la utilidad de los modelos de optimización queda ligada a la validez de la tesis adaptacionista.

Bajo la concepción del adaptacionismo empírico, la falta de soporte empírico de un modelo justifica la formulación de otro modelo de optimización en lugar de considerar alternativas no adaptativas, lo que podría interpretarse como una falla de la tesis adaptacionista en sí. Uno de los defensores más notorios y extremos de esta postura resulta Mayr (1983), que afirma que solo cuando todos los intentos de explicar los fenómenos biológicos a través de la selección natural han fallado se deberían buscar alternativas. Pérez-Escudero y col. (2009) destacan que cuando se encuentra que el estado observado difiere del óptimo global, típicamente se busca aplicar un nuevo modelo de optimización que provea un mejor ajuste en lugar de explorar otras alternativas. Richardson (1994) afirma que este tipo de reajuste de los modelos no es objetable en sí, pero que de esta forma se encontrará eventualmente algún modelo de optimización que pueda ajustar los datos que se pretenden explicar. En una misma línea, Schoemaker (1991)

considera que un punto clave resulta a qué nivel se da la reformulación de un modelo, y cuántos rechazos son necesarios antes de que la teoría o paradigma subyacente sea cuestionado.

Potochnik (2009) distingue un uso alternativo de los modelos de optimización cuando estos son utilizados con el fin de representar el rol de la selección natural, pero sin afirmar que esta sea la única fuerza que influencia el resultado del proceso evolutivo. Tal uso se corresponde con lo que Godfrey-Smith (2001) denomina *adaptacionismo metodológico*, marco en el que un modelo puede servir para dilucidar el rol de factores no selectivos relacionados con el desarrollo o la genética. Milo y Last (2012) afirman que los estudios basados en la optimización deberían tener como objetivo identificar las fuerzas que dan forma y restringen a los sistemas biológicos, antes que preguntarse si su biología es óptima. Es claro que el uso de modelos de optimización en este contexto resulta un caso más general que su uso en el marco del adaptacionismo empírico, ya que la utilidad de los modelos resulta independiente de la valoración del principio de optimización.

En la actualidad existe una virtual unanimidad en que debe considerarse a la optimización como una hipótesis de trabajo antes que como un supuesto (e.g. Reeve y Sherman 1993; Mayhew 1997; Watt 2000; Makela y col. 2002; Anten y During 2011; Milo y Last 2012; Birch 2016). Podemos concluir que la utilidad de un modelo de optimización, en caso de poseer sustento empírico, radica en aportar información no solo sobre la acción del principio de optimización postulado, sino también sobre las restricciones que operan sobre el sistema bajo estudio, así como sobre el efecto de las idealizaciones en las que está basado dicho modelo.

#### 2.3.1.4 *Supuesto de optimización estricta*

Cuando se considera al óptimo global como única predicción de un modelo de optimización se asume un supuesto de optimización en sentido estricto. Este supuesto descansa en la presunción de que, en caso de existir una optimización, la dinámica del proceso asegura que el sistema bajo estudio haya alcanzado el óptimo global. Sin embargo, en general se carece de un conocimiento sobre la dinámica del sistema que dé sustento a este supuesto. Los trabajos que aplican modelos de optimización asumen típicamente una optimización estricta, ya sea de forma explícita o implícita.

Varios autores han resaltado que los sistemas biológicos pueden encontrarse en un estado subóptimo aún cuando efectivamente exista un fenómeno de optimización, y que esta posibilidad debería ser explícitamente considerada como una hipótesis alternativa (e.g. Gladstein y col. 1991; Carlin y Gladstein 1992; Ward 1992, 1993; Callebaut 1998; Bull y Wang 2010). Como se expuso previamente (Sección 2.3.1.2), a diferencia de lo que usualmente se asume, la dinámica de un sistema bajo optimización podría ser tal que impidiese alcanzar el estado óptimo global. Por lo tanto, un sistema podría

encontrarse en estados subóptimos *aún* si efectivamente se halla sujeto a una optimización.

En el ámbito de la teoría de forrajeo, varios autores han abogado por una relajación del supuesto de optimización estricta apelando al concepto de suficiencia. La noción de suficiencia fue introducida por Simon (1956) en el contexto de su teoría de racionalidad limitada de los agentes económicos en oposición al supuesto de toma de decisiones óptimas de los modelos económicos clásicos, trabajo por el que fue reconocido con el Premio Nobel de Economía en 1978. Apelando a este marco teórico, Ward (1992, 1993) plantea que junto con un modelo de optimización estricta deberían evaluarse alternativas subóptimas, ya que si dichas opciones no son consideradas el valor del modelo de optimización estricta resultará desconocido. Callebaut (1998) argumenta que la optimización estricta debería verse como un caso límite de la suficiencia, y que a través de esta aproximación la potencia real y limitaciones de las hipótesis de optimización pueden ser valoradas en un contexto más amplio.

Como ya se señaló en relación a la exclusión de mecanismos causales, los supuestos de un modelo matemático, en su calidad de idealización de una realidad más compleja, deberían evaluarse en función del soporte empírico. Para evaluar la validez del supuesto de optimización estricta, en particular, es necesario construir modelos alternativos que consideren optimización pero que relajen este supuesto, de manera de comparar las predicciones de ambos modelos. Por este motivo, el caso típico de considerar al óptimo global como única predicción, sin tener en cuenta configuraciones subóptimas, impide la evaluación empírica de dicho supuesto. Se puede concluir, por lo tanto, que la construcción de modelos en los que se relaje el supuesto de optimización permite su evaluación, así como una eventual generalización de la hipótesis de optimización planteada.

#### 2.3.1.5 *Supuesto de monotonía estricta*

El estudio de sistemas biológicos en condiciones de campo o laboratorio suele verse afectado por limitaciones en la adquisición de datos. En los estudios basados en la optimización puede suceder que existan limitantes que impidan el cómputo de la función objetivo del modelo planteado. El caso más común se da con hipótesis fundamentadas en la selección natural: cuantificar la aptitud biológica requiere un seguimiento de la descendencia muy trabajoso y poco factible en la mayoría de los casos (Parker y Maynard Smith 1990). Es por este motivo que en diferentes campos de la biología se utiliza a veces una función objetivo sucedánea en lugar de aquella que se hipotetiza realmente optimizada. Por ejemplo, en el marco de la teoría de forrajeo es común utilizar la energía asimilada como función objetivo en lugar de la aptitud biológica (e.g. Sih y Christensen 2001; Stephens y col. 2007), mientras que en el análisis del metabolismo celular con frecuencia se utiliza la velocidad de crecimiento individual (e.g. Kauffman

y col. 2003; Schuetz y col. 2007). El uso de una función objetivo sucedánea se justifica en el supuesto de que existe una relación de monotonía con la verdadera función objetivo. Para los ejemplos citados, asumir que la velocidad de crecimiento o de adquisición de energía guardan una relación creciente con la aptitud biológica tiene una justificación trivial si se considera el rango inferior de valores. Un caso típico es que se considere al óptimo global del modelo como única predicción. En tal caso, se asume que los óptimos globales de ambas funciones objetivo coinciden, y esto significa que se asume una relación de monotonía *estricta* cuya justificación dista de ser elemental o evidente.

Como sucede con todo supuesto, que se asuma una relación de monotonía estricta no significa que ésta se dé efectivamente en la realidad. Gray (1987) plantea que no existe un sustento alguno que permita simplemente asumir dicha relación en el caso de la teoría de forrajeo. En la misma línea, Ward (1992) destaca la necesidad de considerar modelos alternativos que no asuman una relación de monotonía estricta, y menciona casos de estudio donde se ha efectivamente descrito que a partir de un umbral mínimo de energía ingerida no se observan diferencias en el valor de aptitud biológica (Nonacs y Dill 1990). Algo similar puede suceder en el análisis del metabolismo celular: una mayor velocidad de crecimiento podría no estar ligada con una mayor aptitud biológica a partir de cierto umbral. En este punto se aplica nuevamente lo antes mencionado: los supuestos de un modelo deberían ser evaluados en función de su poder predictivo. En este caso el supuesto es independiente de la hipótesis de optimización, y puede ser por lo tanto evaluado separadamente estudiando la relación que efectivamente existe entre la función objetiva sucedánea y la del principio de optimización en la que se basa el modelo. Al igual que con la optimización estricta, también es posible su evaluación empírica si se comparan las predicciones de un modelo que asume monotonía estricta con otro modelo más general en el que este supuesto ha sido relajado.

### 2.3.2 *Críticas a la evaluación del sustento empírico*

En esta sección nos focalizamos en cuestionamientos dirigidos a la evaluación empírica de los modelos de optimización. Si bien la metodología de evaluación de un modelo es independiente de su fundamentación teórica, se verá que, en los hechos, el origen de las prácticas cuestionadas al analizar el soporte empírico se relacionan con aspectos vistos en la sección previa.

#### 2.3.2.1 *Evaluación subjetiva de las predicciones*

Una de las principales críticas que han recibido los estudios sobre optimización en biología es que, en muchos casos, un extenso análisis cuantitativo culmina con un

evaluación cualitativa de las predicciones del modelo (e.g. Cashdan 1983; Pyke 1984; Orzack y Sober 1993, 1994b; Olson y Arroyo-Santos 2015). Orzack y Sober (1994b) señalan que los análisis cualitativos típicamente carecen de un criterio previo acerca de qué diferencia entre predicciones y observaciones debería tomarse como una falta de congruencia, lo cual resulta particularmente problemático porque puede llevar a una aceptación acrítica de las hipótesis de optimización. En este sentido, la gran discrepancia en el criterio que utilizan los investigadores para determinar cuándo un modelo de optimización es aceptado o rechazado socava la relevancia de estos análisis en su conjunto (Orzack y Sober 1996). Por su parte, Olson y Arroyo-Santos (2015) sostienen que la falta de objetividad de los análisis cualitativos constituye un defecto grave en el contraste de modelos de optimización. Una práctica poco estricta a la hora de evaluar modelos matemáticos no es exclusiva del campo de la optimización, sino común en toda la biología, donde la noción de que la naturaleza es compleja parece otorgar licencia para confirmar la validez de un modelo siempre que una predicción sea *cercana* a los datos (Orzack 2012).

La necesidad y relevancia de realizar una evaluación de los modelos de optimización mediante pruebas cuantitativas ha sido señalada en múltiples oportunidades (e.g. King 1987; Mayhew 1997). Orzack y Sober (1994a) plantean que un análisis cuantitativo mediante criterios estadísticos estándar debería ser un requerimiento para el contraste de modelos de optimización. En la misma línea, Sober (2008) destaca la importancia de utilizar inferencia estadística a la hora de evaluar el ajuste de las predicciones de modelos de optimización. Varios autores coinciden con este criterio y recomiendan el uso de herramientas que permitan realizar una selección de modelos, ya sea en el marco de la máxima verosimilitud (e.g. Greeff y Newman 2011), el criterio de información de Akaike proveniente de la teoría de información (e.g. Ward 1992), o mediante el uso de estadística bayesiana (e.g. Kitcher 1987). Para la aplicación de cualesquiera de estos métodos es un requisito la definición de una función de verosimilitud (Burnham y Anderson 2002; Ando 2010). Sin embargo, varios autores subrayan que típicamente las predicciones realizadas en base a hipótesis de optimización son puntuales en lugar de probabilísticas (e.g. Cashdan 1983), haciendo por lo tanto imposible la derivación de una función de verosimilitud. Desde una visión general sobre el tema, Cashdan (1983) sostiene que es justamente la falta de predicciones probabilísticas lo que imposibilita una evaluación objetiva de las hipótesis de optimización, ya que al carecer de varianza no es posible ponderar diferencias entre lo predicho y lo observado.

### 2.3.2.2 *Ausencia de hipótesis alternativas*

La imposibilidad de aplicar una estrategia de selección de modelos mencionada anteriormente está relacionada a su vez con una problemática más general, que es la ausencia de hipótesis alternativas a la de optimización, un punto que también ha

sido criticado de manera recurrente (e.g. King 1987; Pierce y Ollason 1987; Gladstein y col. 1991; Carlin y Gladstein 1992; Ward 1993; Pellis 2002). En el contexto de la teoría de forrajeo, Ward (1993) señala que, ante la ausencia de hipótesis alternativas, una congruencia entre datos observados y predicciones del modelo no resulta de utilidad, dado que otros modelos podrían producir similares predicciones sin hacer uso de una hipótesis de optimización. Desde una óptica complementaria, Andrews y col. (2002b) resaltan que el uso de múltiples hipótesis no sólo es necesario para poner a prueba a una hipótesis de optimización, sino que la consideración de ésta puede ser igualmente necesaria para evaluar explicaciones alternativas. En campos como la fisiología (Anten y During 2011) y el metabolismo celular (Raman y Chandra 2009; Sánchez, Sáez y col. 2014) se ha destacado la necesidad de trabajar con múltiples hipótesis con el fin de comparar el soporte empírico de diferentes principios de optimización alternativos.

La ventaja de considerar múltiples hipótesis de trabajo fue planteada a fines del siglo XIX por Chamberlin (1890) y luego popularizada por Platt (1964). Es importante destacar que la ausencia de hipótesis alternativas no es una crítica exclusiva a los estudios sobre optimización, sino que su relevancia ha sido resaltada en el contexto de las ciencias biológicas en general (e.g. Johnson y Omland 2004; Elliott y Brook 2007). En suma, existe un amplio consenso en la necesidad de considerar modelos alternativos a la hora de evaluar un modelo de optimización. Lógicamente, la comparación con otros modelos implica una mayor carga de trabajo que la evaluación de un único modelo de optimización, pero no parecen existir razones que justifiquen obviar este requisito.

### 2.3.2.3 *Ausencia de una hipótesis nula*

Al destacar la importancia del uso de hipótesis alternativas, tanto Johnson y Omland (2004) como Elliott y Brook (2007) plantean que el uso estándar de una hipótesis nula en el estudio de sistemas biológicos no es suficiente. Sin embargo, varios autores han criticado a los trabajos sobre optimización porque no suelen incluir siquiera una hipótesis nula (e.g. Ward 1993; Garland Jr y Carter 1994; Watt 2000; Andrews y col. 2002a; LaMunyon y Shackelford 2002; Orzack 2002; Sober 2008). Es notorio que incluso autores que destacan las ventajas de trabajar con hipótesis de optimización alternativas omitan mención alguna sobre la necesidad de considerar una hipótesis nula (e.g. Raman y Chandra 2009; Anten y During 2011; Sánchez, Sáez y col. 2014). En este contexto se entiende por *nula* a una hipótesis que se diferencia de la de optimización únicamente porque supone que la optimización es inexistente. En términos más específicos, la hipótesis nula debería poseer el mismo conjunto de restricciones que la de optimización, pero carecer de referencia alguna a la función objetivo.

La ausencia de una hipótesis nula ha sido justificada en el marco del adaptacionismo, argumentando que una hipótesis de optimización puede servir como hipótesis nula amparándose en la supuesta ubicuidad de la selección natural. Tal como se dis-

cutió anteriormente (Sección 2.3.1.3), en este punto existe un consenso general en la necesidad de considerar a la optimización como una hipótesis de trabajo a ser evaluada en función de su poder predictivo antes que como un supuesto. De forma un tanto contradictoria, Parker y Maynard Smith (1990) rechazan el trato de la optimización como un supuesto pero al mismo tiempo sostienen su uso como hipótesis nula, a la vez que reconocen la dificultad que esto conlleva dado que las técnicas estadísticas clásicas permiten el rechazo de una hipótesis nula antes que su confirmación. Garland Jr y Carter (1994) sostienen que considerar una hipótesis de optimización como una hipótesis nula es contrario a la manera que los modelos nulos son utilizados típicamente tanto en la biología a la hora de realizar una inferencia estadística. Watt (2000) coincide con este punto de vista, y agrega que el modelo nulo es justamente aquel carente del efecto bajo estudio, y en consecuencia un modelo de optimización no cumple con este requisito si el objetivo del trabajo es evaluar el soporte a una hipótesis de optimización.

La importancia de trabajar con una hipótesis nula radica en que solo así puede evaluarse si el poder predictivo del modelo depende realmente de la hipótesis de optimización en sí. Al generar predicciones basándose en un argumento de optimización se utilizan modelos que incluyen necesariamente conocimiento sobre su estructura y/o dinámica, tanto en el planteo mismo del problema como en la formulación de las restricciones y en el valor de los parámetros. Es por esta razón que cuando se presenta un acuerdo entre observaciones y predicciones cuantitativas debería mostrarse que el poder predictivo depende de la optimización y no es un producto exclusivo de las restricciones planteadas en el modelo. Por lo tanto, se debería considerar como requisito metodológico indispensable una comparación del poder predictivo de un modelo de optimización con el de un modelo nulo.

## 2.4 EVALUACIÓN DE MODELOS DE OPTIMIZACIÓN EN APLICACIONES RECIENTES

En esta sección describimos en primer lugar la selección de artículos analizados, para luego abordar aspectos relacionados con el planteo de los modelos de optimización como la función objetivo y supuestos asociados, y por último analizar las prácticas utilizadas en su evaluación empírica.

La búsqueda bibliográfica en el campo del análisis de balance de flujos metabólicos arrojó en un total de 2.483 artículos publicados en el período 2007 - 2016. Los 20 artículos más citados que realizan una evaluación empírica de modelos de optimización se ubicaron dentro de los 81 trabajos más citados; en la Tabla 2.1 se presentan datos bibliométricos de los trabajos analizados. El número de citas por año de los artículos seleccionados varió entre 59.2 y 12.7, y fueron publicados tanto en revistas de interés general (*Science*, *Nature*, *PNAS* y *PLOS One*) como de campos específicos (*Molecular Systems Biology*, *PLOS Computational Biology*, etc.).

Tabla 2.1: Datos bibliométricos de los 20 artículos analizados en el campo de análisis de flujos metabólicos publicados en el período 2007-2016; las citas recibidas se circunscriben al mismo período.

Artículo	Posición en selección	Citas por año	Citas totales	Publicación
Feist y col. (2007)	4	59.2	592	<i>Molecular Systems Biology</i>
Fan y col. (2014)	7	52.3	157	<i>Nature</i>
Schuetz y col. (2007)	11	32.2	322	<i>Molecular Systems Biology</i>
Schuetz y col. (2012)	12	31.2	156	<i>Science</i>
Machado y Herrgård (2014)	22	22.3	67	<i>PLOS Computational Biology</i>
Harcombe y col. (2014)	27-29	20.0	60	<i>Cell Reports</i>
Nogales y col. (2012)	27-29	20.0	100	<i>PNAS</i>
Chang y col. (2011)	31	19.2	115	<i>Molecular Systems Biology</i>
Chandrasekaran y Price (2010)	32	18.7	131	<i>PNAS</i>
Lewis y col. (2010)	33	18.4	129	<i>Molecular Systems Biology</i>
Colijn y col. (2009)	39	17.0	136	<i>PLOS Computational Biology</i>
Stolyar y col. (2007)	44	16.5	165	<i>Molecular Systems Biology</i>
Poolman y col. (2009)	47	16.1	129	<i>Plant Physiology</i>
Zomorodi y Maranas (2012)	55	14.8	74	<i>PLOS Computational Biology</i>
Selvarasu y col. (2012)	66	13.6	68	<i>Biotechnology and Bioengineering</i>
Wintermute y Silver (2010)	67	13.6	95	<i>Molecular Systems Biology</i>
Covert y col. (2008)	72	13.1	118	<i>Bioinformatics</i>
Saha y col. (2011)	73-78	13.0	78	<i>Plos One</i>
Vazquez y col. (2010)	73-78	13.0	91	<i>BMC Systems Biology</i>
Oberhardt y col. (2008)	79-81	12.7	114	<i>Journal of Bacteriology</i>

En la tabla Tabla 2.2 se presenta una breve descripción de los principales aportes de los artículos analizados que da cuenta de la variedad de temáticas y modelos biológicos abordados. Una porción de los artículos se centra en la evaluación de una metodología original para la construcción de modelos del flujo metabólico (Covert y col. 2008; Colijn y col. 2009; Chandrasekaran y Price 2010; Lewis y col. 2010; Wintermute y Silver 2010; Selvarasu y col. 2012; Zomorodi y Maranas 2012; Harcombe y col. 2014), o bien analizan metodologías propuestas por terceros (Machado y Herrgård 2014). Otros artículos utilizan modelos de optimización con el fin de realizar una primer estimación de los flujos metabólicos en una especie dada (Oberhardt y col. 2008; Poolman y col. 2009; Chang y col. 2011; Saha y col. 2011), una actualización de un modelo previo mediante

mejoras en la metodología (Feist y col. 2007; Nogales y col. 2012), y en el caso de Stolyar y col. (2007) el primer modelo a escala genómica de una comunidad microbiana. De hecho, un 20 % de los trabajos consisten en análisis de comunidades mediante modelos de optimización (Stolyar y col. 2007; Wintermute y Silver 2010; Zomorodi y Maranas 2012; Harcombe y col. 2014). Finalmente, solo un 15 % de los artículos plantean como un objetivo explícito poner a prueba una hipótesis de optimización (Schuetz y col. 2007; Vazquez y col. 2010; Schuetz y col. 2012).

Los modelos de optimización planteados en el 80 % de los artículos analizados son de tipo simple, mientras que un 15 % es de tipo multi-nivel, y Schuetz y col. (2012) presenta el único modelo multi-objetivo, donde múltiples funciones objetivo son optimizadas en un mismo espacio muestral (Tabla 2.3). En la mayoría de los artículos analizados se utiliza la maximización del crecimiento celular como función objetivo: en el 50 % de los casos es la única función usada, mientras que en un 25 % del total se incluyen además modelos con otras funciones objetivo (Tabla 2.3). La minimización de la sumatoria de flujos metabólicos, por su parte, es utilizada como función objetivo en un 25 % de los artículos. En ambos casos, la justificación de la función objetivo es su (hipotético) vínculo con la aptitud biológica. El máximo crecimiento celular se hipotetiza asociado a una máxima velocidad de reproducción en organismos unicelulares (Feist y Palsson 2010), mientras que la mínima sumatoria de flujos metabólicos se asume relacionada a una máxima eficiencia enzimática en el crecimiento celular y de esta forma a una máxima aptitud biológica (Bonarius y col. 1996). Otra función objetivo utilizada es la máxima producción de ATP, que también se asume representa una adaptación (Ebenhöh y Heinrich 2001). En consecuencia, el fundamento de la función objetivo utilizada en los modelos de optimización planteados se basa en el 90 % en los artículos en una hipótesis de evolución por selección natural. Sin embargo, cabe destacar que solo cinco de estos dieciocho trabajos plantean una justificación de la función objetivo de forma explícita (Schuetz y col. 2007; Oberhardt y col. 2008; Lewis y col. 2010; Zomorodi y Maranas 2012; Harcombe y col. 2014). Si bien en algunos trabajos se da cuenta de un marco teórico evolutivo, lo que podría permitir al lector inferir un vínculo entre la función objetivo y la evolución por selección natural (Feist y col. 2007; Nogales y col. 2012; Schuetz y col. 2012), la mayoría de los artículos omiten mención alguna al razonamiento que sostiene la elección de la función objetivo. Finalmente, los únicos dos artículos en que el uso de la función objetivo *no* está basado en una hipótesis de evolución por selección natural presentan una justificación explícita (Poolman y col. 2009; Wintermute y Silver 2010).

Ninguno de los artículos analizados hace referencia alguna a los supuestos de optimización estricta o de monotonía estricta, típicamente utilizados en los modelos de optimización. Los modelos planteados en el 85 % de los artículos asumen una optimización estricta, donde la predicción del modelo está compuesta por los estados óptimos únicamente (Sección 2.3.1.4). Los trabajos en donde este supuesto es relajado plantean

Tabla 2.2: Breve reseña de los principales aportes de la selección de artículos analizados en el campo de análisis de flujos metabólicos en el período 2007-2016.

Artículo	Principales aportes
Feist y col. (2007)	Reconstrucción de red metabólica a nivel de genoma de <i>Escherichia coli</i> y primer análisis de consistencia termodinámica
Fan y col. (2014)	Establecimiento de nueva vía metabólica en la producción de NADPH en células humanas mediante combinación de técnicas analíticas y modelización matemática
Schuetz y col. (2007)	Evaluación sistemática de múltiples modelos de optimización para predecir flujos metabólicos en <i>E. coli</i>
Schuetz y col. (2012)	Muestran que <i>E. coli</i> opera de forma cercana a una superficie óptima de Pareto en un espacio de tres dimensiones
Machado y Herrgård (2014)	Evaluación de diferentes métodos para integrar información del transcriptoma en modelos metabólicos de <i>E. coli</i> y <i>Saccharomyces cerevisiae</i>
Harcombe y col. (2014)	Entorno de modelización que permite predecir la dinámica de comunidades microbianas de un bajo número de especies
Nogales y col. (2012)	Análisis del proceso fotosintético en el alga <i>Synechocystis sp.</i> mediante modelo de la red metabólica a escala del genoma
Chang y col. (2011)	Reconstrucción de red metabólica de <i>Chlamydomonas reinhardtii</i> a escala genómica y predicción del crecimiento en función de fuente lumínica
Chandrasekaran y Price (2010)	Método que integra de forma automatizada datos de redes de transcripción regulatoria en modelos de la red metabólica
Lewis y col. (2010)	Evaluación del soporte de modelo de red metabólica en <i>E. coli</i> en la predicción de información del transcriptoma y proteoma
Colijn y col. (2009)	Método de modelización del metabolismo celular que utiliza y define nuevas restricciones en base a datos de expresión génica
Stolyar y col. (2007)	Primer modelo metabólico multi-específico basado en datos de secuenciación de <i>Desulfovibrio vulgaris</i> y <i>Methanococcus maripaludis</i>
Poolman y col. (2009)	Construcción y análisis de modelo metabólico a escala genómica de <i>Arabidopsis thaliana</i>
Zomorodi y Maranas (2012)	Entorno de modelización de comunidades microbianas mediante análisis de flujo con optimización mutli-nivel y multi-objetivo
Selvarasu y col. (2012)	Análisis de la fisiología de líneas celulares de mamíferos mediante integración de datos de cultivo y perfiles metabólicos en un modelo

Tabla 2.2 Breve reseña de los principales aportes de la selección de artículos analizados en el campo de análisis de flujos metabólicos en el período 2007-2016 (continuación).

Artículo	Principales aportes
Wintermute y Silver (2010)	Construcción de modelo que permite predecir el surgimiento de interacciones metabólicas complejas en comunidades microbianas
Covert y col. (2008)	Método de modelización que combina análisis de balance de flujos con regulación por lógica booleana y ecuaciones diferenciales ordinarias
Saha y col. (2011)	Primer modelo metabólico a escala genómica de <i>Zea mays</i> y su análisis bajo diferentes condiciones fisiológicas
Vazquez y col. (2010)	Análisis de alta velocidad de glicólisis en células cancerosas aún en presencia de oxígeno mediante modelo metabólico de producción de ATP
Oberhardt y col. (2008)	Primer modelo del metabolismo a escala genómica de <i>Pseudomonas aeruginosa</i>

modelos del metabolismo celular cuyas predicciones consideran el conjunto de estados que poseen una velocidad de crecimiento celular que resulta, como mínimo, un cierto porcentaje del óptimo global (Lewis y col. 2010; Zomorodi y Maranas 2012; Machado y Herrgård 2014).

El supuesto de monotonía estricta, por su parte, solo puede darse cuando la función objetivo del modelo se justifica por una relación supuesta con la aptitud biológica, lo que, como antes se mencionó, sucede en el 90 % de los artículos analizados. En consonancia con lo antes expuesto, en ninguno de estos trabajos se discute cuál es la relación específica entre dichas variables. En 15 de los 18 trabajos se asume una optimización estricta, y esto permite concluir que los modelos planteados asumen una monotonía estricta entre la función objetivo y la aptitud biológica, ya que en caso contrario la predicción del modelo debería incluir estados subóptimos (Sección 2.3.1.5). Con respecto a los tres artículos en donde se relaja el supuesto de optimización estricta, en ninguno encontramos una mención explícita a la relación entre la aptitud biológica y la velocidad de crecimiento, la función objetivo utilizada en dichos trabajos. Zomorodi y Maranas (2012) y Lewis y col. (2010) hacen mención a estados subóptimos para referirse a aquellos donde la velocidad de crecimiento es menor a la del óptimo global, pero dado que no discuten el sentido que otorgan a *óptimo* en este contexto no es posible establecer si asumen o no de hecho una monotonía estricta.

El análisis de la metodología usada en la evaluación de los modelos de optimización analizados se presenta en la Tabla 2.4. Encontramos que el 65 % de los artículos evalúa de forma únicamente subjetiva las predicciones de los modelos planteados, es decir, no

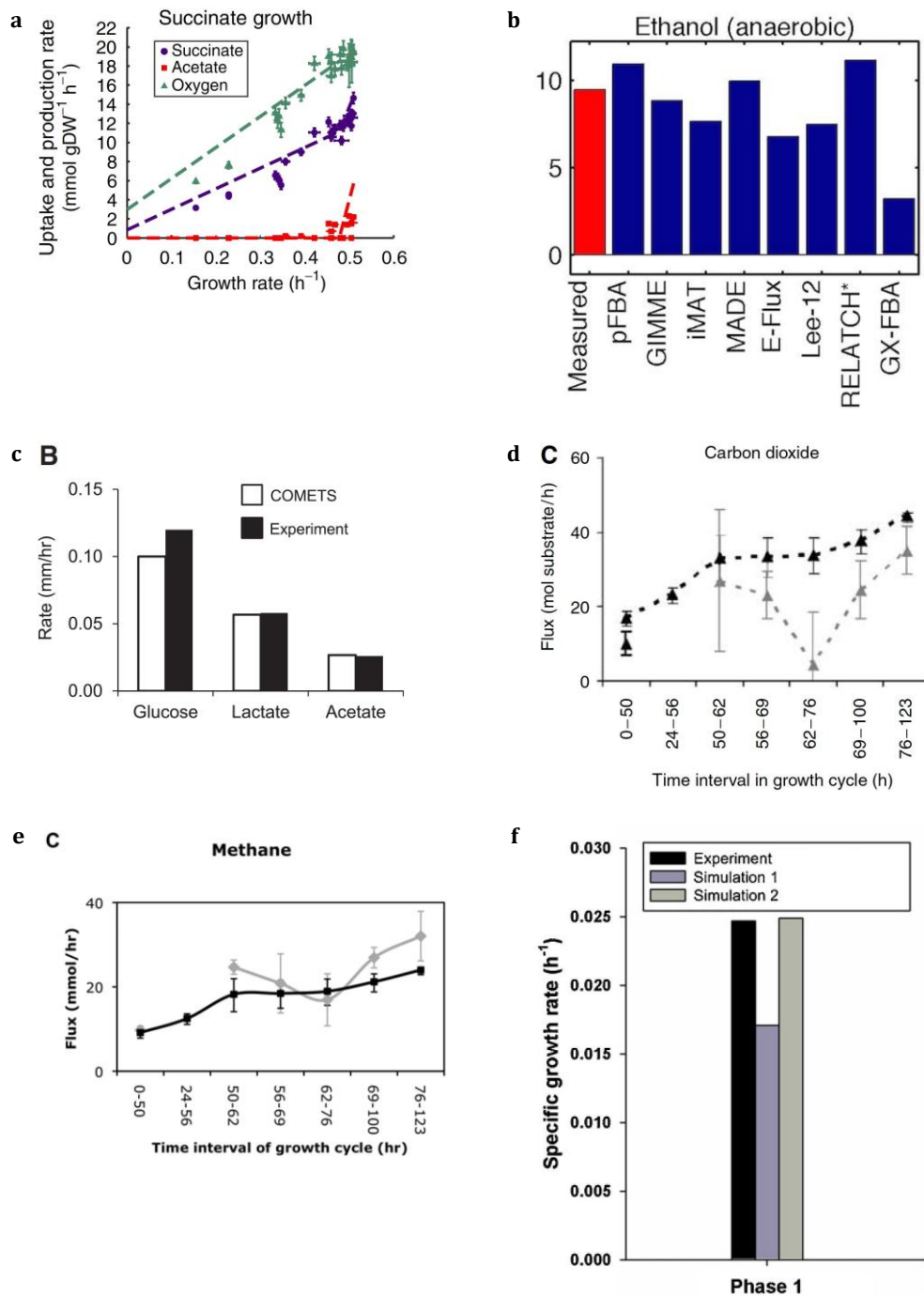


Figura 2.2: Ejemplos de figuras de algunos de los artículos analizados que son utilizadas en la evaluación subjetiva de predicciones de modelos de optimización. ¿En qué casos los resultados brindan soporte empírico a la hipótesis de optimización planteada? (a) Reproducido de Fan y col. (2014). (b) Reproducido de Machado y Herrgård (2014). El valor observado corresponde a la barra de la extrema izquierda mientras que el resto de los valores son predicciones de modelos de optimización a excepción de “iMAT”, “Lee-12”b y “REALTECH\*”, cuya formulación no involucra una función objetivo. (c) Reproducido de Harcombe y col. (2014). La leyenda “COMETS” corresponde al modelo de optimización. (d) Reproducido de Stolyar y col. (2007). La predicción del modelo de optimización en color negro, mientras los valores experimentales se indican con un color gris claro. (e) Reproducido de Zomorodi y Maranas (2012). La predicción del modelo de optimización en color negro, mientras los valores experimentales se indican con un color gris claro. (f) Reproducido de Selvarasu y col. (2012). La leyenda “Simulation 1” y “Simulation 2” corresponden a modelos de optimización que varían únicamente en la fórmula de la biomasa celular utilizada en la función objetivo de maximización del crecimiento celular.

explicitan el criterio con el que evalúan el ajuste de los modelos a los datos experimentales. Este tipo de evaluación se realiza en la mayoría de los casos en la forma de una comparación gráfica; en la Figura 2.2 presentamos algunos ejemplos extraídos de los artículos analizados. Feist y col. (2007) comparan la relación predicha por el modelo de optimización entre el crecimiento celular y el consumo y producción de compuestos en cultivos de *E. coli* bajo distintas condiciones experimentales (Figura 2.2a). Si bien se observa un aparente ajuste entre el modelo y los valores observados, al mismo tiempo varios puntos caen fuera de la predicción del modelo: no es claro qué alejamiento de lo predicho representaría una falta de ajuste. En el caso la Figura 2.2b de Machado y Herrgård (2014) esto es más notorio todavía, ya que se comparan las predicciones de varios modelos alternativos cuya distancia al valor observado es variable. Parece una obviedad que mientras menor sea la distancia entre predicción y valor observado existe un mayor soporte de un cierto modelo, pero ¿qué distancia es indicativa en dicha figura de que un modelo no provee un ajuste adecuado? Poolman y col. (2009) y Saha y col. (2011) prescinden incluso de un recurso gráfico y realizan únicamente unas comparaciones puntuales entre valores predichos y observados en el cuerpo del texto, lo que limita aún más la capacidad de evaluar los modelos de optimización que presentan.

Aquellos artículos que evalúan de forma objetiva el soporte empírico de los modelos de optimización lo hacen a través de métodos estadísticos. En todos los casos, la inferencia está basada en una estadística de tipo frecuentista en la forma de una prueba de significación, ya sea a través de un análisis de correlación lineal entre observaciones y predicciones de variables continuas (Chandrasekaran y Price 2010; Wintermute y Silver 2010), análisis de tabla de contingencia en el caso de variables discretas (Nogales y col. 2012), o mediante la implementación de un análisis *ad hoc* basado en un método de Monte Carlo (Lewis y col. 2010; Chang y col. 2011; Schuetz y col. 2012). Como excepción, Schuetz y col. (2007), explicitan *a priori* una región de rechazo para la diferencia entre predicciones y datos experimentales.

Un 60 % de los artículos plantea un único modelo de optimización del sistema bajo estudio, y por lo tanto carece de hipótesis alternativas. Entre aquellos trabajos que presentan modelos alternativos, todos evalúan múltiples modelos de optimización, que se diferencian en las restricciones (Covert y col. 2008; Chandrasekaran y Price 2010; Lewis y col. 2010), la función objetivo (Colijn y col. 2009; Schuetz y col. 2012; Fan y col. 2014), o ambas (Schuetz y col. 2007; Machado y Herrgård 2014). En contraste, solo tres artículos evalúan modelos alternativos a los de optimización. Machado y Herrgård (2014) incluyen modelos cuya formulación se basa en el planteo de restricciones pero carecen de una función objetivo, o en el caso de incluirla se utiliza para maximizar el ajuste a los flujos metabólicos a datos de expresión génica. Covert y col. (2008), por su parte, incluye un modelo basado en ecuaciones diferenciales ordinarias, y se trata del único caso que analiza un modelo que carece de restricciones. Por último, Schuetz

Tabla 2.3: Caracterización de los modelos de optimización planteados en los artículos analizados en el campo de análisis de flujos metabólicos publicados en período 2007-2016.

Artículo	Modelo de optimización	Función objetivo	Optimización estricta	Monotonía estricta
Feist y col. (2007)	Simple	Máximo crecimiento	C	C
Fan y col. (2014)	Simple	Mínima suma de flujos Máximo crecimiento	C	C
Schuetz y col. (2007)	Simple	Máximo crecimiento 10 funciones alternativas	C	C
Schuetz y col. (2012)	Multi-objetivo	Máximo crecimiento 53 funciones alternativas	C	C
Machado y Herrgård (2014)	Simple	Máximo crecimiento	×	?
Harcombe y col. (2014)	Simple	Máximo crecimiento	C	C
Nogales y col. (2012)	Simple	Máximo crecimiento	C	C
Chang y col. (2011)	Simple	Máximo crecimiento	C	C
Chandrasekaran y Price (2010)	Simple	Máximo crecimiento	C	C
Lewis y col. (2010)	Multi-nivel	Máximo crecimiento con mínima suma de flujos	×	?
Colijn y col. (2009)	Simple	Máximo crecimiento Máx. producción de ác. micolico	C	C
Stolyar y col. (2007)	Multi-nivel	Máx. crecimiento de <i>D. vulgaris</i> con máx. crecimiento de <i>M. maripaludis</i>	C	C
Poolman y col. (2009)	Simple	Mínima suma de flujos	C	-
Zomorodi y Maranas (2012)	Multi-nivel	Máx. biomasa de cada especie con máx. biomasa de la comunidad	×	?
Selvarasu y col. (2012)	Simple	Máximo crecimiento	C	C
Wintermute y Silver (2010)	Simple	Mínimo ajuste metabólico	C	-
Covert y col. (2008)	Simple	Máximo crecimiento	C	C
Saha y col. (2011)	Simple	Máximo crecimiento	C	C
Vazquez y col. (2010)	Simple	Máxima producción de ATP	C	C
Oberhardt y col. (2008)	Simple	Máximo crecimiento	C	C

y col. (2012) consideran un modelo alternativo que comparte con el de optimización las restricciones pero carece de función objetivo.

De lo antes expuesto se desprende que solo uno de entre veinte trabajos analizados compara predicciones de un modelo de optimización con una hipótesis nula. El que varios trabajos utilicen pruebas de significación estadística pero, a la vez, carezcan de una hipótesis nula con respecto a la optimización puede parecer contradictorio, dado que toda prueba de esta naturaleza está basada en un modelo nulo por definición. La explicación de este hecho reside en cómo son construidas las pruebas estadísticas, ya que en todos los casos la hipótesis nula *no* corresponde a una ausencia de optimización. En el caso del análisis de correlación, por ejemplo, Chandrasekaran y Price (2010) evalúan la capacidad de un modelo de optimización de predecir la velocidad de crecimiento de diferentes fenotipos de *E. coli* en condiciones aeróbicas y anaeróbicas. La hipótesis nula de la prueba estadística utilizada corresponde a la ausencia de una relación lineal entre la tasa de crecimiento observada y predicha, pero no a la relación que se esperaría en ausencia de la hipótesis de máximo crecimiento celular. Nogales y col. (2012) utilizan un modelo de optimización para predecir si un gen es esencial para el crecimiento de *Synechocystis sp.*, y evalúan los resultados mediante un análisis de tabla de contingencia. La hipótesis nula consiste, en este caso, en que existe una relación aleatoria entre la predicción del modelo acerca de cuales genes son esenciales y aquello que se observa en el laboratorio, pero nada indica que un modelo que carezca de optimización necesariamente predice una ausencia de relación. En lugar de utilizar una prueba estadística estándar, Chang y col. (2011) construyen un análisis *ad hoc*. Pero este análisis tiene un objetivo diferente que el de evaluar que se esperaría si no se da una optimización: la hipótesis nula corresponde a una aleatorización de la longitud de onda de la luz incidente, un parámetro del modelo, y esto les permite evaluar si el espectro de la radiación incidente afecta el ajuste del modelo. Lewis y col. (2010) consideran varias hipótesis nulas construidas por aleatorización que les permiten evaluar si datos de regulación génica se parecen a las predicciones del modelo de optimización más de lo esperable por efecto del azar. De esta manera, en ninguno de los casos las hipótesis consideradas nulas lo son en relación a la optimización propiamente dicha. En todos los casos la hipótesis nula se construye suponiendo que la relación entre las variables relevantes son al azar, pero nada indica que esta es la situación que se esperaría de un modelo que carezca de optimización pero preserve las restricciones en su formulación.

El trabajo de Schuetz y col. (2012), por lo tanto, es el único entre los analizados que cumple con todos los requerimientos en la evaluación empírica de modelos de optimización establecidos en la sección anterior. Sin embargo, en la metodología utilizada en este trabajo se pueden identificar igualmente falencias. Los autores evalúan un modelo de optimización mediante una prueba de significación en la que se compara la distancia al óptimo teórico de datos experimentales con las predicciones derivadas de la

Tabla 2.4: Análisis de la metodología de evaluación empírica de modelos de optimización en artículos seleccionados en el campo del análisis de flujos metabólicos.

Artículo	Evaluación objetiva	Inferencia estadística	Hipótesis alternativas	Hipótesis nula
Feist y col. (2007)	×	×	×	×
Fan y col. (2014)	×	×	<b>C</b>	×
Schuetz y col. (2007)	<b>C</b>	×	<b>C</b>	×
Schuetz y col. (2012)	<b>C</b>	<b>C</b>	<b>C</b>	<b>C</b>
Machado y Herrgård (2014)	×	×	<b>C</b>	×
Harcombe y col. (2014)	×	×	×	×
Nogales y col. (2012)	<b>C</b>	<b>C</b>	×	×
Chang y col. (2011)	<b>C</b>	<b>C</b>	×	×
Chandrasekaran y Price (2010)	<b>C</b>	<b>C</b>	<b>C</b>	×
Lewis y col. (2010)	<b>C</b>	<b>C</b>	<b>C</b>	×
Colijn y col. (2009)	×	×	<b>C</b>	×
Stolyar y col. (2007)	×	×	×	×
Poolman y col. (2009)	×	×	×	×
Zomorodi y Maranas (2012)	×	×	×	×
Selvarasu y col. (2012)	×	×	×	×
Wintermute y Silver (2010)	<b>C</b>	<b>C</b>	×	×
Covert y col. (2008)	×	×	<b>C</b>	×
Saha y col. (2011)	×	×	×	×
Vazquez y col. (2010)	×	×	×	×
Oberhardt y col. (2008)	×	×	×	×

hipótesis nula. Ahora bien, el modelo evaluado con dicha metodología no fue establecido *a priori*, sino que es seleccionado de entre un total de 151,686 modelos de optimización alternativos <sup>2</sup>. Más allá de la prueba de significación utilizada, la probabilidad de rechazar una hipótesis nula siendo verdadera aumenta cuando se realizan múltiples pruebas de significación, lo que se conoce como el problema de las comparaciones múltiples (Benjamini 2010). Cuando se utiliza una gran cantidad de modelos alternativos, especialmente, es necesario aplicar algún tipo de metodología que tenga en cuenta las múltiples comparaciones para evitar obtener resultados espurios, por ejemplo aquellas que permiten controlar la tasa de rechazos de hipótesis nula verdaderas (Yekutieli y Benjamini 1999; Pike 2011). En consecuencia, una correcta inferencia estadística en

este contexto requeriría la aplicación de un método de corrección por pruebas múltiples, lo que conduce al interrogante acerca de si los resultados de Schuetz y col. (2012) representan una evidencia sobre la existencia de optimización en *E. coli* o son, por el contrario, un artefacto producto de una metodología estadística deficiente.

## 2.5 DISCUSIÓN

Al comienzo de este trabajo nos planteamos como interrogante si las críticas a la aproximación de la optimización son todavía relevantes hoy en día en la biología, o si por el contrario revisten en la actualidad una importancia únicamente histórica. La metodología de revisión bibliográfica nos permitió identificar un conjunto de críticas de diversa índole dirigidas tanto a la fundamentación teórica de los modelos de optimización como a la evaluación de su soporte empírico. En la mayoría de las críticas analizadas en este trabajo encontramos un consenso acerca de la validez e importancia de los argumentos involucrados, lo que nos permite concluir que son efectivamente relevantes y no deberían ser ignoradas.

Al analizar los cuestionamientos a los fundamentos teóricos de los modelos de optimización, en ningún caso encontramos un consenso que plantee que deban ser destruidos de la investigación científica de la biología. En relación a las críticas motivadas por supuestos ontológicos y por exclusión de mecanismos causales se han planteado posiciones contrapuestas entre las que parece difícil que se pueda arribar a un consenso. ¿Poseen entonces los modelos de optimización sustento desde un punto de vista filosófico? Más allá de las controversias, en ambos casos existen argumentos con asiento en escuelas de pensamiento filosófico que brindan efectivamente sustento a estos modelos. La exclusión de mecanismos causales, de hecho, es identificada en el marco de la filosofía de la ciencia como una ventaja: no solo es posible plantear explicaciones en términos de un principio de optimización aún cuando el modelo utilizado ignore la dinámica del sistema, sino que el hecho de no realizar suposiciones adicionales podría tener como resultados la producción de explicaciones más robustas.

Con respecto a los demás cuestionamientos a los modelos de optimización en su calidad de constructos teóricos, encontramos que existe un consenso mayoritario acerca de su relevancia. El tratamiento de la optimización como un supuesto antes que como una hipótesis de trabajo es una práctica de difícil justificación en el estado actual de la biología, así como de las ciencias fácticas en general. Por su parte, los supuestos

2 A partir de un conjunto de 54 funciones objetivo, los autores seleccionaron 5 en base al ajuste a datos experimentales. Evaluaron luego todos los modelos de optimización múltiple de Pareto basados en combinaciones de a dos funciones de entre las cinco seleccionadas. Al no obtener en ninguno de los casos un ajuste satisfactorio, evaluaron entonces todas las combinaciones de *tres* funciones objetivo previamente seleccionadas, y encontraron que una de estas sí presenta un buen ajuste. Este fue finalmente el modelo de optimización evaluado. Dado que se partió de 54 funciones objetivo, la cantidad de combinaciones de 2 y/o 3 elementos que podrían haber sido evaluadas resulta de:  $\binom{54}{2} + \binom{54}{3} = 151,686$ .

de optimización y monotonía estricta presentes en la gran mayoría de los modelos de optimización representan idealizaciones sobre las que no puede afirmarse *a priori* que resulten erradas. Hecha esta aclaración, resultarían una limitante en la formulación de los modelos en caso de que efectivamente exista una optimización pero no se cumpla alguno de los supuestos antes mencionados. Por lo tanto, la relajación de dichos supuestos y la consiguiente generalización de los modelos representaría una ventaja sin perjuicio alguno. El análisis de estas críticas nos permite concluir, a su vez, que si se desea trabajar con modelos de optimización es importante tener conciencia de aquellos supuestos que les otorgan sustento. Asimismo, el postulado de optimización sobre el sistema bajo estudio *no* se debe encontrar entre dichos supuestos sino que debe constituir una hipótesis de trabajo que se someta a evaluación. A su vez, la evaluación del poder predictivo de modelos alternativos aporta información sobre la validez empírica de las idealizaciones en las que están basados, lo que resalta la importancia de una evaluación empírica rigurosa desde el punto de vista de su fundamentación teórica.

Las críticas a la evaluación empírica de los modelos de optimización gozan de un consenso aún mayor que reconoce la validez de las falencias señaladas y, como consecuencia, la necesidad de que la metodología de evaluación cumpla con ciertos estándares. Una evaluación de las predicciones establecida *a priori* y que considere hipótesis alternativas parece ser un requisito mínimo e indispensable para evaluar objetivamente si un trabajo brinda o no soporte empírico a un modelo de optimización. Existe también un consenso en que la forma menos arbitraria de realizar una evaluación objetiva implica el uso de algún tipo de inferencia estadística estándar, así como en la ventaja de que dicha inferencia esté basada en la función de verosimilitud, ya sea mediante el uso de métodos de máxima verosimilitud, selección de modelos mediante el criterio de Akaike, o estadística bayesiana. Si se desea evaluar el soporte empírico de la hipótesis de optimización en sí, es necesario contar además con una hipótesis nula que difiera del modelo de optimización únicamente en que no considera una función objetivo.

El segundo interrogante que nos planteamos al inicio de este trabajo es si la popularidad actual de los modelos de optimización indica que los aspectos cuestionados se han resuelto o superado, o si, por el contrario, han sido simplemente ignorados. Teniendo en cuenta el consenso sobre los cuestionamientos a la metodología de evaluación empírica, sumado al hecho de que en ningún caso se trata de planteos recientes, sería razonable esperar que artículos recientes en un área con un importante desarrollo cumplan los requisitos anteriormente descriptos. Sin embargo, en la mayoría de los artículos que analizamos en el campo del análisis de metabolismo celular encontramos deficiencias metodológicas en la evaluación empírica de los modelos. Si bien representan una muestra acotada, los artículos analizados son los más citados en un área que ha tenido un crecimiento exponencial en la última década. El hecho de tratarse de los trabajos más citados en su ámbito disciplinar no necesariamente implica que la metodología empleada sea superior a la de otros artículos similares pero menos citados. No

obstante, el número de citas es ampliamente utilizado en análisis bibliométricos como una medida del impacto de un artículo científico (e.g. Callahan y col. 2002; Uthman y col. 2013). Parece prudente afirmar que los artículos analizados han servido de referencia para otros estudios sobre el metabolismo celular. En virtud de esto, entendemos que el conjunto de prácticas metodológicas que siguen dichos artículos representa efectivamente una muestra de lo que en este ámbito se considera como prácticas aceptadas. En conclusión, entendemos que las deficiencias metodológicas encontradas en estos artículos indican que los cuestionamientos a la evaluación empírica de los modelos de optimización mantienen su vigencia en la actualidad, al menos en el campo particular del análisis del metabolismo celular.

En este punto es posible preguntarse cuál es la causa de que los trabajos en optimización continúen sufriendo de deficiencias metodológicas que han sido señaladas hace tiempo y sobre las que existe un amplio consenso. A fin de abordar este interrogante, consideraremos tres alternativas que cubren en cierta manera el conjunto de posibilidades. Podría suceder que los autores de los trabajos que poseen falencias en la evaluación de las hipótesis desconozcan las críticas y recomendaciones sobre este asunto. En caso de estar al tanto de los argumentos involucrados, es posible que los investigadores desestimen las críticas por no considerarlas pertinentes, o bien que aún cuando las consideren relevantes existan limitaciones metodológicas que les impidan atenderlas. Sin la pretensión de realizar un examen exhaustivo, analizaremos a continuación algunos factores que podrían contribuir a perpetuar defectos en la evaluación de hipótesis en cada uno de estos escenarios.

Dada la abundante bibliografía sobre esta temática, un investigador que realice una búsqueda de una profundidad suficiente encontrará sin dudas la existencia de cuestionamientos al uso de modelos de optimización. Podemos señalar, sin embargo, algunos aspectos que podrían contribuir a brindar un panorama diferente en caso de realizar una revisión bibliográfica acotada. Un porcentaje mayoritario de los artículos que plantean críticas a los modelos de optimización se han centrado en aspectos filosóficos, mientras que las cuestiones metodológicas han sido abordadas con menos énfasis. Es posible que ante una exploración superficial de la bibliografía aquellos interesados en aplicaciones concretas de la optimización puedan considerar que las críticas a la aplicación de optimización en biología tratan únicamente aspectos filosóficos poco dados a una formalización matemática y sin consecuencias para su trabajo concreto. El disponer de análisis típicamente enfocados de acuerdo al campo de aplicación tampoco facilita encontrar información relevante.

Como segunda posibilidad, consideramos que los investigadores podrían elegir ignorar las críticas por considerarlas carentes de sustento. Este es un aspecto difícil de analizar, ya que los trabajos sobre optimización no suelen tomar una posición explícita respecto de tales cuestiones. Cuando se profundiza en el análisis de las críticas a la metodología de evaluación empírica parece evidente su relación con cuestiones de

nivel filosófico. En particular, un punto que observamos de forma recurrente es la asociación entre deficiencias en las prácticas de evaluación empírica con el asumir que el principio de optimización del modelo opera de hecho en el sistema bajo estudio. Esto es especialmente notorio en la ausencia de hipótesis nula, pero también se relaciona con la falta de hipótesis alternativas en general, y a su vez con el problema más amplio de una evaluación subjetiva de las hipótesis. Sin embargo, ya se ha mencionado en varias oportunidades que existe un virtual acuerdo en tratar a la optimización como una hipótesis de trabajo; lo mismo puede decirse acerca de la importancia de las críticas a la evaluación empírica de los modelos. Por lo tanto, en este punto creemos que una mayor difusión sobre la relevancia de las deficiencias señaladas contribuiría a que un mayor número de autores realizase un análisis crítico sobre la forma en que evalúa modelos de optimización.

Como tercera alternativa, existe la posibilidad de que los autores, aún reconociendo la relevancia de los cuestionamientos en la evaluación de hipótesis, incurran en deficiencias metodológicas debido a una incapacidad para cumplir con las exigencias planteadas. En esta instancia, creemos que puede echar luz al respecto un análisis de trabajos basados en principios de optimización con fundamento en la termodinámica. En esta área de estudio se han planteado múltiples principios candidatos que derivan de una variedad de consideraciones teóricas, pero que carecen de un sustento empírico concluyente, así como de una aceptación general en la biología. Los trabajos que evalúan este tipo de hipótesis de optimización no representan una excepción, sino que poseen el mismo tipo de deficiencias metodológicas señaladas anteriormente al momento de evaluar las hipótesis (e.g. Jesus y col. 2012). En base a lo antes expuesto, parece justificado descartar que los autores de dichos estudios consideren como un supuesto el principio de optimización termodinámico sobre el que investigan, o que ésta sea la causa de una poca atención a las cuestiones metodológicas. De hecho, el objetivo explícito de estos trabajos suele ser brindar soporte empírico al principio de optimización en cuestión. Esto representa una diferencia fundamental con lo que sucede al utilizar hipótesis de optimización sustentadas en la selección natural, donde el propósito de los trabajos puede variar pero nunca está dirigido a probar la existencia del principio de optimización. Como se ha visto, los estudios realizados en el marco de la biología evolutiva pueden incluir como objetivo obtener información sobre las restricciones que operan sobre el sistema así como la validez de las idealizaciones del modelo de optimización, pero nunca involucran una evaluación de la selección natural en sí debido a su aceptación universal en la biología. Cabe preguntarse cómo es que los estudios empíricos sobre optimización con fundamentos termodinámicos, en los que se esperaría una especial atención a la metodología de acuerdo a sus propios objetivos, adolecen igualmente de problemas en este aspecto. La situación descrita refuerza la idea de que las deficiencias señaladas pueden tener origen en la falta de una metodolo-

gía que permita evaluar hipótesis de optimización cumpliendo con los requerimientos ya mencionados.

En este punto esperamos que los argumentos expuestos sirvan para establecer que, sin lugar a dudas, es fundamental un cuidado de la metodología de evaluación empírica de los modelos de optimización. Ahora bien, si se han publicado trabajos que no cumplen con los estándares señalados a la hora de analizar el sustento empírico de tales modelos, esto sugiere que el problema es más complejo que el desconocimiento o desinterés de los autores y/o limitaciones metodológicas. Creemos que, antes que ignorarlas, sería provechoso reconocer las limitaciones metodológicas que poseen los estudios que evalúan modelos de optimización: esto no solo permitiría un análisis más realista de sus resultados y conclusiones sino que contribuiría, a su vez, a identificar más claramente obstáculos al avance del conocimiento en esta área de trabajo.

Por último, en base a la revisión realizada cabe señalar que entre los autores que han criticado aspectos de la evaluación empírica de los modelos de optimización predominan aquellos que se limitan a señalar falencias antes que presentar posibles soluciones que permitan subsanar dichas deficiencias. Esta situación no debería resultar llamativa, dado que es bien sabido que así como un análisis crítico de los métodos utilizados en una rama de la ciencia resulta valioso y necesario, encontrar una solución práctica a los problemas señalados suele representar un desafío de otra índole. En el caso particular de la optimización, la dificultad a nivel metodológico podría tener origen en los diferentes objetivos de la biología y la ingeniería, de donde provienen las técnicas matemáticas y computacionales aplicadas típicamente en para abordar problemas de optimización. En palabras de Oster y Wilson (1979), mientras que la ingeniería busca diseñar una maquinaria o un proceso de la forma más eficiente posible, la biología busca inferir los principios que rigen la organización de los sistemas biológicos. En vista del enfoque antagónico, no es extraño que el uso de técnicas generadas en el marco de la ingeniería no permitan cumplir con los estándares de evidencia utilizados en las ciencias biológicas.

## GENERALIZACIÓN Y EVALUACIÓN DE MODELOS DE OPTIMIZACIÓN

---

### Resumen

Los modelos de optimización son utilizados en muchas áreas de la biología para explicar diversos fenómenos en sistemas de diferentes niveles de organización pero, sin embargo, su evaluación empírica suele poseer numerosas deficiencias. No hay disponible a la fecha una metodología que permita una evaluación de estos modelos libre de tales falencias y de aplicación general, y esto representa un obstáculo para establecer el poder predictivo de las hipótesis de optimización de forma objetiva.

En este trabajo presentamos una generalización de los modelos de optimización típicamente utilizados que posibilita su expresión en la forma de un modelo probabilístico bajo ciertas condiciones y, en consecuencia, hace posible analizar su soporte empírico mediante métodos estadísticos basados en la verosimilitud. En el planteo de la formulación general tomamos como punto de partida la noción de que el espacio muestral y las restricciones de un modelo de optimización constituyen un modelo nulo, lo que sumado a una relajación de los supuestos de optimización y monotonía estricta permite una generalización de los modelos típicamente utilizados. Con este fin introducimos dos tipos de funciones denominadas de transferencia y de normalización que permiten, respectivamente, la relajación de los supuestos y la aplicación a modelos no deterministas. La formulación resultante incluye como casos particulares tanto a la típica hipótesis de optimización estricta como a una hipótesis nula, así como a un abanico de situaciones intermedias. El desarrollo matemático requerido para la formalización de estas ideas se basa en la teoría axiomática de la probabilidad, y en consecuencia es válido para un modelo de optimización arbitrario siempre que cumpla con las condiciones establecidas.

Al permitir el uso de una inferencia estadística estándar, el método propuesto carece de las limitaciones de los métodos habitualmente utilizados. La generalidad de la formulación matemática, a su vez, permite la evaluación de hipótesis de optimización más generales y sobre modelos de sistemas biológicos de complejidad arbitraria. Por lo tanto, se espera que la aplicación de lo aquí presentado contribuya a responder el interrogante de hasta qué punto las propiedades de los sistemas biológicos pueden ser predichas en base a principios de optimización.

### 3.1 INTRODUCCIÓN

A pesar del uso extendido de modelos de optimización en diversos campos de la biología, todavía se mantienen vigentes en la actualidad interrogantes respecto de su real poder predictivo. De acuerdo a lo tratado en el Capítulo 2, uno de los pocos

puntos en que proponentes y críticos de los modelos de optimización coinciden es en la importancia de una evaluación rigurosa del soporte empírico de dichos modelos. Sin embargo, justamente la metodología utilizada para evaluar el poder predictivo ha sido objeto de numerosas críticas.

El procedimiento más extendido para evaluar modelos de optimización consiste en una comparación de los óptimos globales predichos con los valores experimentales. Dicha comparación se da a veces basada únicamente en la forma de una comparación visual, o bien acompañada en algunos casos por de una prueba de significación. Ahora bien, la evaluación gráfica carece de un criterio de rechazo establecido *a priori*, y en consecuencia constituye una forma de evaluación subjetiva. Aún cuando se realiza una prueba de significación, por ejemplo en la forma de un análisis de correlación, la hipótesis nula de la prueba de significación suele corresponder a una ausencia de correlación, *no* a una ausencia de optimización (Sección 2.3.2.1).

En vista de lo expuesto, cabe considerar qué atributos debería tener una metodología general para la evaluación de hipótesis de optimización. Una de las conclusiones del Capítulo 2 es que el poder predictivo de un modelo de optimización solo puede ser evaluado en relación al de otros modelos alternativos. Si es un objetivo del trabajo establecer además el soporte empírico de la optimización en sí, de forma separada al resto de los componentes del modelo, es necesario que se incluya al menos una hipótesis nula que carezca del componente de optimización. La metodología debería permitir una evaluación objetiva de la hipótesis de optimización mediante la aplicación de algún tipo de inferencia estadística. Para esto, debería ser capaz de derivar predicciones en forma de distribuciones de probabilidad al menos en el caso de la hipótesis nula, para aplicar entonces una inferencia a través de una prueba de significación estadística clásica. Para permitir la aplicación de una inferencia en el marco de la teoría de información o de la estadística bayesiana es necesario además que se pueda obtener una predicción probabilística a partir de los modelos de optimización que se desean evaluar. La utilidad de una metodología con tales características se incrementaría si hace posible además una relajación de los supuestos de optimización estricta y monotonía estricta, que son utilizados comúnmente en los estudios sobre optimización. Por último, para que una metodología sea de aplicación general y permita así realizar análisis comparativos entre diferentes sistemas, su formulación debería estar basada en elementos comunes a cualquier modelo de optimización. El consenso existente respecto de las deficiencias metodológicas, así como en la identificación de requisitos que se deberían cumplir para evitarlas, contrasta con la ausencia de una metodología estándar que se encuentre libre de dichas falencias.

En el marco del desarrollo que la biología de sistemas experimentó a partir del año 2000, se han propuesto métodos alternativos para la evaluación de modelos de optimización que dan respuesta a algunos de los requisitos planteados. Burgard y Maranas (2003) proponen un método que, de acuerdo con los autores, permite evaluar mode-

los de optimización alternativos. Otros trabajos han considerado de manera explícita estados subóptimos en las predicciones de los modelos como una forma de relajar los supuesto de optimización y/o monotonía estricta. Pérez-Escudero y col. (2009) parten de la noción de que en un sistema sujeto a optimización las variables que poseen un menor impacto en la función objetivo tienen una mayor probabilidad de desviarse del valor predicho por la configuración óptima. Partiendo de este concepto los autores evalúan hipótesis de optimización en dos sistemas experimentales a partir del cómputo de estados subóptimos, y por lo tanto esta metodología generaliza el tipo de predicciones que pueden derivarse a partir de un modelo de optimización. Wintermute y col. (2013) estudian el efecto de considerar estados subóptimos en el análisis y predicción del metabolismo celular, y resaltan que esto permite relajar al mismo tiempo los supuestos de optimización estricta y de monotonía estricta. Estas propuesta adolecen, sin embargo, de limitaciones ya mencionadas. La metodología propuesta por Burgard y Maranas (2003) parte del supuesto de que el sistema bajo estudio se encuentra sujeto a *algún* principio de optimización, y no se establece *a priori* en qué casos se debe concluir que ninguna de las hipótesis de optimización planteada es adecuada. De forma más general, ninguno de los métodos permite comparar las predicciones derivadas del modelo de optimización planteado con una hipótesis nula construida suponiendo ausencia de optimización. Por lo tanto, si bien los métodos descritos brindan efectivamente una respuesta a alguna de las deficiencias típicas en la evaluación de los modelos, no ofrecen una solución al conjunto de los problemas señalados. Asimismo, en ninguno de los casos queda establecido que la metodología de evaluación sea correcta desde un punto de vista matemático formal.

Un aspecto de la evaluación de los modelos de optimización que a nuestro entendimiento no ha sido abordado con un énfasis suficiente es cómo debe formularse la hipótesis nula, lo que consideramos resulta crítico para el desarrollo de una metodología general. En las pruebas de significación la hipótesis nula es derivada a partir de un modelo estadístico, y dada su formulación es mutuamente excluyente con la hipótesis alternativa. Ambas hipótesis derivan en general de un mismo modelo estadístico, y lo usual es que en una hipótesis nula algunos de los parámetros tomen un valor fijo incluido en el modelo alternativo, que representa un caso más general. Muchas veces la hipótesis nula recibe el nombre de modelo nulo para indicar que en dicho modelo está ausente el efecto sobre el que se desea realizar una inferencia. La relevancia del uso de modelos nulos es bien conocida en varias áreas de la biología (Gotelli y Graves 1996; Gotelli 2001; Gotelli y Ulrich 2012). La inclusión de modelos nulos puede modificar de forma determinante las conclusiones a las que se arriba, tal como atestigua un ejemplo en el ámbito de los modelos de optimización en biología que describimos a continuación. Un conjunto de trabajos en el marco de teoría de historia de vida predicen la existencia de relaciones aproximadamente constantes entre parámetros vitales basándose en modelos de optimización (e.g. Charnov 1993; Charnov y col. 2001). Estas

relaciones se han denominado rasgos invariantes de la historia de vida, y su existencia ha sido reportada en una variedad de taxones animales (e.g. Jensen 1996; Mangel 1996; Collin 2006). El método típicamente utilizado para evaluar estos modelos de optimización ha sido un análisis de correlación lineal entre los parámetros cuya relación se presume invariante. Ahora bien, de acuerdo a lo antes expuesto este método carece de referencia a una hipótesis nula respecto de la optimización, y por lo tanto no cumple los requerimientos para una evaluación rigurosa del modelo en cuestión. Buston y col. (2004) utilizaron un modelo nulo para cuestionar los resultados de Allsop y West (2003). Posteriormente, y de forma más general, Nee y col. 2005 mostraron que un modelo nulo que supone relaciones aleatorias entre parámetros vitales genera predicciones similares a la hipótesis de optimización, desacreditando así el sustento empírico de este campo de investigación. Sin embargo, la capacidad de los modelos nulos empleados en estos trabajos puede ser a su vez cuestionada. Dichos modelos son “nulos” solo en relación a la correlación esperada entre los parámetros de la historia de vida, *no* en cuanto a que constituyen una hipótesis nula respecto de la optimización, lo que dio lugar a un debate respecto de la forma en que se debe generar el modelo nulo (Savage y col. 2006; Linde y Palmer 2008). Más allá de detalles específicos de un cierto modelo, es claro que un modelo nulo debería ser similar en todas sus características respecto al modelo de mayor generalidad para el que actúa como “nulo”, a excepción de el o los mecanismos de los que desea evaluar el soporte empírico.

El objetivo del presente capítulo es desarrollar una metodología general con un sustento matemático formal que permita una evaluación objetiva del soporte empírico de modelos de optimización típicamente utilizados mediante el uso de inferencia estadística. Como resultado de este trabajo presentamos una metodología basada en la teoría axiomática de la probabilidad que permite evaluar modelos de optimización bajo ciertas condiciones generales y provee una solución a la totalidad de las críticas realizadas a los métodos comúnmente utilizados.

La idea central de la que partimos en este capítulo es que si buscamos construir un modelo nulo apropiado para evaluar modelos de optimización es necesario incluir el espacio muestral y las restricciones que operan sobre el sistema bajo estudio; una vez formulado el modelo nulo este puede ser generalizado en un modelo de optimización mediante la inclusión de la función objetivo. En la primer sección del trabajo abordamos el rol que tienen las restricciones en cualquier modelo de optimización de un sistema biológico. En la siguiente sección presentamos una generalización de los modelos de optimización típicos basada en una relajación de los supuestos de optimización estricta y monotonía estricta, y donde la formulación resultante incluye como casos particulares tanto la típica hipótesis de optimización estricta como una hipótesis nula, así como un abanico de situaciones intermedias. Dado que el desarrollo matemático requerido para la formalización de estas ideas se basa en la teoría axiomática de la probabilidad, se presenta una breve introducción a algunos de los principales concep-

tos de la teoría matemática de la medida. A continuación establecemos cómo y bajo qué condiciones es posible derivar de forma general un modelo nulo probabilístico a partir de un problema de optimización, y cómo mediante la definición de funciones de normalización y de transferencia es posible formular un modelo general en la forma de un espacio de probabilidad. En la última sección mostramos cómo realizar una inferencia estadística a partir del modelo general probabilístico de optimización presentado. Finalmente, se discuten las limitaciones y ventajas de la metodología propuesta y los aportes que podría brindar al estudio de problemas de optimización en biología.

### 3.2 EL ROL DE LAS RESTRICCIONES EN LOS MODELOS DE OPTIMIZACIÓN

Si bien un modelo matemático de optimización puede plantearse sin restricción alguna, cuando el objeto de estudio se trata de un sistema experimental siempre existirá alguna restricción a los valores que pueden tomar las variables de estado del sistema. Las restricciones sobre un sistema biológico pueden surgir a partir de una variedad de fenómenos biológicos, químicos y/o físicos, e imponen condiciones sobre las propiedades del sistema. Pueden originarse como consecuencia exclusiva de la estructura y funcionamiento de un sistema, es decir, a partir de características intrínsecas del sistema, independientemente de su entorno. Ejemplos de tales restricciones son la máxima tasa metabólica sostenida que puede soportar un mamífero sin que aumente su temperatura (Speakman y McQueenie 1996; Bacigalupe y Bozinovic 2002), la tasa de crecimiento máxima que un individuo puede alcanzar en condiciones óptimas (Werner y Griebeler 2014), o la estequiometría en la composición química que caracteriza a una especie (Raubenheimer y Simpson 2004; Frost y col. 2005). Las restricciones pueden derivar asimismo de la relación entre un sistema y su entorno, y en este caso influirán factores extrínsecos al sistema. En un ecosistema, por ejemplo, la máxima biomasa posible dependerá de los nutrientes disponibles en el ambiente, mientras que su productividad máxima dependerá de la radiación solar incidente en el caso un sistema basado en la fotosíntesis (Berger y col. 2006; Wang y col. 2014).

Desde un punto de vista matemático se pueden plantear, en general, dos tipos de restricciones en cuanto a su formulación: las restricciones de igualdad, que se expresan en forma de ecuaciones; y las restricciones de desigualdad, expresadas en la forma de inecuaciones (Biegler 2010). Ambos tipos de restricciones imponen condiciones sobre las propiedades de un sistema, pero difieren en el conjunto de valores que dicha propiedad puede tomar. Una restricción de igualdad expresa que una función de las variables y/o parámetros del sistema puede tomar únicamente un determinado valor. Este tipo de restricciones pueden expresar, por ejemplo, una estequiometría constante en la composición elemental de individuos de una misma especie. Dado que el sistema observado debe satisfacer, por definición, cualquier restricción a la que esté sujeto, el

valor de las restricciones de igualdad puede ser estimado directamente a partir del estado real del sistema simplemente utilizando el valor de los parámetros en la ecuación de la restricción, en contraste con lo que sucede con restricciones de desigualdad. Las restricciones de desigualdad, por su parte, plantean que una cierta propiedad del sistema puede adquirir un rango de valores pero no puede superar un cierto límite. Las restricciones de máxima biomasa o máxima productividad antes mencionadas son ejemplos de este tipo. En general, el valor límite de restricciones de desigualdad deberá ser estimado a partir de información adicional a la que permite caracterizar el estado del sistema.

Si bien las restricciones poseen una definición matemática precisa en el marco de la formulación matemática de los modelos de optimización, en el campo de la biología existe una importante variación en cuanto a qué se refieren los autores de un texto cuando hablan de *restricción*. Al realizar el trabajo de revisión bibliográfica que presentamos en el Capítulo 2 encontramos que el uso que diferentes autores hacen de este término es fuente de una cierta confusión. En efecto, pudimos identificar usos del término *restricciones* que no se corresponden con su definición matemática formal. A continuación presentamos una muy simple y breve clasificación orientativa teniendo como base la definición matemática de las restricciones de un modelo de optimización.

**DEFINICIÓN ACOTADA :** este uso ocurre cuando el uso de *restricción* se reserva para designar, de acuerdo a algún criterio, un subconjunto de las restricciones que cumplen con la definición matemática. El concepto de *trade-off* propio de la biología evolutiva, por ejemplo, se utiliza muchas veces como perteneciente a una categoría diferente a la de restricción (e.g. Charnov y Downhower 1995; Pigliucci y Kaplan 2000), a pesar de que corresponde a su definición matemática. Amundson (1994) señala que, en el contexto del paradigma adaptacionista, varios autores consideran que para una restricción sea considerada como tal su presencia debe prevenir que el sistema alcance un óptimo global, es decir, el valor de la función objetivo que se alcanzaría de no existir una restricción debería mayor en el caso de una maximización, o menor en el caso de una minimización. Este criterio puede resultar de utilidad respecto de qué procesos limitan efectivamente un proceso de optimización. Sin embargo, desde el punto de vista de su formulación matemática tal distinción es irrelevante: la definición matemática de una restricción es independiente de la función objetivo.

**DEFINICIÓN AMPLIADA :** se da cuando el término de *restricción* es utilizado para designar no solo a elementos que cumplen con la definición matemática, sino también a otros que no respetan esta definición. Este es el caso de las llamadas "restricciones selectivas" con las que a veces se hace referencia a aspectos de la geometría del espacio muestral, que de acuerdo al proceso de optimización en cuestión pueden evitar que se alcance un estado óptimo (Getty 2000). Si bien tal

propiedad del espacio muestral puede llegar ser relevante para el resultado de un proceso de optimización, esta característica no se corresponde con la definición matemática de restricción.

Más allá del uso incorrecto del término en parte de la bibliografía desde el punto de vista matemático, las restricciones son frecuentemente reconocidas como un punto central de los modelos de optimización en biología (e.g. Yoshimura y Shields 1987; Kleidon y col. 2010; Volk y Pauluis 2010). Si un sistema biológico se encuentra sujeto a una optimización, el estado observado resultará ser uno de entre los óptimos disponibles de acuerdo a las restricciones que operen sobre el sistema. Aún si se sabe con certeza que un sistema está efectivamente sujeto a una optimización, esta información por sí sola puede tener un escaso poder predictivo si se desconocen las restricciones que limitan el fenómeno de optimización. Por lo tanto, el estado de un sistema dependerá tanto del principio de optimización involucrado como de las restricciones que definen las fronteras dentro de las que la optimización puede tener lugar.

A pesar del lugar preponderante que tienen en la definición de los modelos de optimización, son pocos los autores que plantean una relación entre las restricciones y la evaluación de hipótesis de optimización en sistemas biológicos. DeLong (2008) pone a prueba hipótesis de máxima potencia en experimentos de competencia en el laboratorio mediante el análisis de estados alternativos del sistema que se genera mediante una manipulación experimental. Esto resulta posible dada la naturaleza del sistema bajo estudio en dicho trabajo: un ensayo de laboratorio con unidades compuestas por individuos de un bajo número de especies. Tal manipulación experimental, no obstante, es impracticable en la mayoría de los casos. En la mayoría de los sistemas el número de configuraciones alternativas es muy elevado o directamente infinito. Asimismo, la manipulación experimental suele ser impracticable por múltiples razones. Cai y col. (2006) realizan un estudio experimental sobre la dinámica de un microcosmos de laboratorio y, de forma coincidente, plantean que determinar si el estado observado del sistema es efectivamente un óptimo presenta dificultades debido a las innumerables configuraciones alternativas que puede asumir los componentes de un sistema complejo. Ahora bien, si es posible plantear las restricciones que operan sobre un sistema y construir de esta forma un modelo nulo, podría ser posible establecer justamente cuáles son las configuraciones alternativas en las que el sistema puede encontrarse en *ausencia* de optimización. Esto posibilitaría, a su vez, la evaluación de la hipótesis de optimización, ya que brindaría la hipótesis nula necesaria; es a esta tarea a la que nos abocamos en las siguientes secciones.

### 3.3 GENERALIZACIÓN DEL PROBLEMA DE OPTIMIZACIÓN

La aplicación de una inferencia estadística sobre un modelo de optimización requiere la obtención de predicciones en la forma de una distribución de probabilidad, dado que únicamente a partir de esta es posible obtener una función de verosimilitud (Berger y Wolpert 1988). El problema radica en que, como ha sido previamente planteado, la predicción típica de los modelos de optimización es de tipo puntual, y plantea que el sistema se encuentra en un estado en el que la función objetivo posee un valor óptimo global. Esta predicción se deriva de un supuesto de optimización estricta, y en muchos de los casos está basada además en un supuesto de monotonía estricta.

Veremos a continuación cómo es posible plantear una generalización de un problema de optimización en su formulación típica que incluye la optimización estricta como un caso extremo y opuesto al de la hipótesis nula representada por un modelo nulo, así como un abanico de situaciones intermedias que permiten una relajación de los supuestos de optimización estricta y monotonía estricta. En esta sección delinearemos el razonamiento general mediante el uso de nociones básicas de probabilidad únicamente, para más adelante dar desarrollo a estas ideas en el marco formal de la teoría axiomática de la probabilidad basada en la teoría de la medida.

#### FORMULACIÓN DE MODELO DE OPTIMIZACIÓN ESTRICTA

Los modelos de optimización suelen basarse en un supuesto de optimización estricta (Sección 2.3.1.4). En tales modelos se obtienen predicciones puntuales, pero aún en este caso es posible formular las predicciones en términos probabilísticos. Considérese un problema de optimización compuesto por un cierto espacio muestral  $\Omega$ , un espacio de soluciones candidatas  $\Omega_c \subseteq \Omega$  definido por un conjunto de restricciones, y una función objetivo  $f: \Omega \rightarrow \mathbb{R}$ . La solución del problema de optimización se define como el conjunto de estados del sistema que pertenecen al espacio de soluciones y para los que el valor de la función objetivo es máximo<sup>1</sup>. El conjunto de soluciones óptimas  $\Omega_o \subseteq \Omega_c \subseteq \Omega$  se define entonces como  $\Omega_o \equiv \{\omega : \omega \in \Omega_c, f(\omega) = \max_{\omega \in \Omega_c} f(\omega)\}$ .

Un modelo de optimización que supone una optimización estricta plantea que el sistema se encontrará en un estado  $\omega \in \Omega_o$ . Si suponemos que existe un único estado

<sup>1</sup> Se considerará aquí el máximo de la función objetivo porque es el caso más común en la biología a causa de la aptitud biológica, pero todo lo afirmado es igualmente válido si se considera una minimización.

óptimo global <sup>2</sup>, entonces la predicción de un modelo tal se puede expresar en términos de la probabilidad de observar un cierto estado  $\omega$ :

$$P(\omega) = \begin{cases} 1, & \text{si } \omega \in \Omega_c \text{ y } f(\omega) = \max_{\omega \in \Omega_c} f(\omega) \\ 0, & \text{si } \omega \notin \Omega_c \text{ y / o } f(\omega) < \max_{\omega \in \Omega_c} f(\omega). \end{cases} \quad (3.1)$$

La expresión anterior permite observar que, de forma general, una hipótesis de optimización afirma que la probabilidad de observar a un sistema en un cierto estado varía con el valor de la función objetivo. Cuando la optimización es estricta y existe un óptimo global, como en este caso, dicha relación es extrema: la probabilidad es 1 para el estado óptimo global, y nula para cualquier otro estado.

#### RELAJACIÓN DEL SUPUESTO DE OPTIMIZACIÓN ESTRICTA

Supóngase ahora que el supuesto de optimización estricta es relajado, lo que implica que un estado puede ser una solución del problema de optimización aunque el valor de la función objetivo evaluada en dicho estado sea menor al máximo que toma en el espacio de soluciones candidatas. En este caso es posible generalizar la definición del conjunto de soluciones óptimas:

$$\Omega_o \equiv \{\omega : \omega \in \Omega_c, f(\omega) \geq q \cdot \max_{\omega \in \Omega_c} f(\omega)\}, \quad (3.2)$$

con  $q \in (0, 1]$ . El parámetro  $q$  corresponde al parámetro de posición de una función escalonada, y representa la proporción de la función objetivo que un estado del espacio de soluciones candidatas debe poseer para pertenecer al conjunto de soluciones óptimas.

Obsérvese que de esta manera se generaliza una hipótesis de optimización, ya que  $q = 1$  corresponde a una optimización estricta, mientras  $0 < q < 1$  representa una optimización donde son posibles soluciones subóptimas.

##### 3.3.0.1 Relajación del supuesto de monotonía estricta

La formulación antes expuesta permite relajar también, en caso de existir, el supuesto de monotonía estricta. Este supuesto es común en el planteo de modelos de optimización en biología, y se utiliza cuando la función objetivo que se hipotetiza optimizada,  $f_v$ , no puede calcularse en base a los datos disponibles para un cierto sistema experimental. En estos casos, se asume que existe una relación de monotonía estricta entre  $f_v$  y una función  $f$  que se utiliza como función objetivo:  $f(\omega_1) > f(\omega_2) \Leftrightarrow f_v(\omega_1) >$

<sup>2</sup> Por simplicidad se considera la existencia de un estado óptimo global, pero este no siempre ocurre ya que pueden existir múltiples o infinitas soluciones óptimas alternativas. El caso general será analizado más adelante.

$f_V(\omega_2)$ . El supuesto de monotonía estricta puede ser relajado por uno de monotonía no estricta si se considera que la relación entre  $f_V$  y  $f$  es de monotonía estricta hasta un cierto valor  $f$ , por encima del cual un aumento de  $f$  no implica un aumento de  $f_V$ . El punto en el que cambia la relación de monotonía se puede expresar en función del valor máximo de la función objetivo como  $\rho \max_{\omega \in \Omega_c} f(\omega)$ ,  $\rho \in (0, 1]$ . Una hipótesis de optimización en la que se relaja el supuesto de monotonía estricta pero se mantiene el de optimización estricta se puede representar entonces de acuerdo a la definición 3.2, donde  $q = \rho$ . En este caso,  $\rho = 1$  representa el caso particular de monotonía estricta, y  $0 < \rho < 1$  representa una relajación de dicho supuesto.

### 3.3.0.2 Formulación general

Considérese ahora el caso general de la definición 3.2 donde  $0 \leq q \leq \rho \leq 1$ . El caso particular  $q = \rho = 1$  representa la hipótesis típica de optimización bajo los supuestos de optimización y monotonía estrictas. El caso de  $q = \rho < 1$  representa una relajación del supuesto de optimización estricta, pero no del supuesto de monotonía estricta, mientras  $q < \rho < 1$  implica una relajación de ambos supuestos. El caso de  $q = \rho = 0$  corresponde, en contraste, a una ausencia de optimización, donde todo estado que cumpla con las restricciones pertenece al conjunto de soluciones óptimas:  $\Omega_o = \Omega_c$ . Por lo tanto, es posible abarcar con una misma expresión general el espectro de casos particulares, y, lo que resulta más relevante, formular una hipótesis nula asociada.

En base a lo expuesto, queda de manifiesto la relación existente entre el modelo nulo y diferentes variantes de una hipótesis de optimización y el valor de  $q$ . Por lo tanto, realizar una inferencia sobre una hipótesis de optimización equivale a realizar una sobre  $q$ . Si se puede obtener la distribución de probabilidad de  $q$  en función de los datos disponibles sobre un sistema bajo estudio será posible entonces establecer una función de verosimilitud y realizar así una inferencia estadística de alguna clase; sobre este aspecto tratan las dos siguientes secciones.

## 3.4 INTRODUCCIÓN A LA TEORÍA DE LA MEDIDA

La obtención de un modelo nulo y de un modelo general probabilístico de optimización que se presenta en las siguientes secciones depende de la capacidad de derivar un constructo matemático denominado espacio de probabilidad a partir de un problema arbitrario de optimización <sup>3</sup>, lo que permite aplicar métodos de inferencia estadística para evaluar el sustento empírico de hipótesis de optimización. El consiguiente desa-

<sup>3</sup> Estrictamente, se verá que se establecen un mínimo de requerimientos sobre el espacio muestral, las restricciones y la función objetivo, pero la clase de problemas que se abarcan es lo suficiente amplia como para incluir, sino la totalidad, la inmensa mayoría de los modelos de optimización que se encuentran en la bibliografía sobre optimización en biología.

El desarrollo matemático se basa en la teoría axiomática de la probabilidad, formalizada por primera vez por Kolmogorov (1956) y que tiene como sustento a la teoría de la medida. Si bien esta formalización de la teoría de la probabilidad es la de mayor popularidad, no es la única formalización existente. Se ha elegido utilizarla aquí dada la similitud que existe entre su formulación y la de un problema de optimización. Es importante aclarar, sin embargo, que Jaynes (2003) demuestra que la formalización de Kolmogorov es consistente con la teoría de la probabilidad también axiomática pero derivada a partir de postulados como una extensión de la lógica (Cox 1961) y que sirve de base para la estadística Bayesiana como una teoría de la inferencia. Por lo tanto, los resultados que se obtendrán en este trabajo a partir de la teoría de la probabilidad de Kolmogorov serán equivalentes a los que se arribaría partiendo de una teoría de la inferencia Bayesiana.

La principal ventaja de formalizar la construcción de un modelo nulo mediante la teoría axiomática de la probabilidad consiste en el tratamiento unificado de cualquier tipo de problema, más allá de si se trata de un dominio continuo o discreto, finito o infinito. Esta generalidad posibilita, como se verá más adelante, la obtención de modelos a través de una misma formulación para problemas de optimización que usualmente son abordados de forma separada. Como contrapartida, resulta necesario trabajar con conceptos de la teoría de la medida, una rama del análisis matemático en la que se basa la teoría axiomática de la probabilidad. Es por este motivo que se presenta a continuación una breve introducción a conceptos básicos de la teoría de la medida dirigida a un público no especialista. Para la elaboración de este apartado hemos tomado de referencia los textos de Athreya y Lahiri (2006) y Pollard (2002), y hemos incluido lo que consideramos son los conceptos esenciales necesarios para dar sustento a lo desarrollado en las siguientes secciones de este capítulo; lógicamente, este apartado será omitido por quienes manejen conceptos elementales de la teoría de la medida.

### *Espacio de medida*

Un espacio de probabilidad consiste en un caso particular de un constructo matemático denominado espacio de medida y que está formado por tres elementos que definiremos a continuación: un espacio muestral  $\Omega$ , una  $\sigma$ -álgebra sobre dicho espacio, y una medida  $\mu$  definida en base a los dos elementos anteriores.

**ESPACIO MUESTRAL** Un espacio muestral  $\Omega$  es un conjunto de elementos  $\omega \in \Omega$ . Un espacio muestral puede ser definido de manera arbitraria, pero de acuerdo a su cardinalidad podrá ser:

- Un espacio contable finito, si posee una cardinalidad menor a la del conjunto de los números naturales, es decir,  $\Omega$  posee  $n$  elementos  $\omega$ , con  $n \in \mathbf{N}$  y  $n < \infty$ .
- Un espacio contable infinito, si posee una cardinalidad igual a la de los números naturales, es decir, una cantidad discreta pero infinita de elementos.
- Un espacio no contable, si posee una cardinalidad superior a la de los números naturales, por ejemplo cualquier intervalo de los reales  $(a, b)$  con  $a, b \in \mathbf{R}$ ,  $a < b$ .

**$\sigma$ -ÁLGEBRA** Un concepto fundamental en la definición axiomática de probabilidad es el de  $\sigma$ -álgebra, también denominado *campo* de conjuntos por algunos autores (e.g. Kolmogorov 1956). Un conjunto  $\Omega$  y una  $\sigma$ -álgebra  $F$  definida sobre  $\Omega$  constituyen un espacio medible  $(\Omega, F)$ ; los elementos de  $F$  son llamados conjuntos medibles. La relevancia de esta dupla para la teoría de la probabilidad radica en que solo es posible definir una medida de probabilidad sobre conjuntos medibles.

Una  $\sigma$ -álgebra  $F$  sobre un conjunto  $\Omega$  es una colección de subconjuntos de  $A \subseteq \Omega$  que cumple con las siguientes condiciones:

1. Es cerrada para operaciones de complementación, lo que significa que para todo elemento  $A$  de  $F$  su complemento  $\Omega - A$  pertenece también a  $F$ :

$$A \in F \iff \Omega - A \in F .$$

2. Incluye a  $\Omega$ , lo que dada la propiedad antes enunciada equivale a que el conjunto vacío  $\emptyset$  también pertenece a  $F$ .
3. Es cerrada frente a uniones contables, lo que significa que dado un conjunto contable de elementos su unión también pertenece a  $F$ :

$$A_n \in F \text{ para } n \geq 1 \implies \bigcup_{n \geq 1} A_n \in F .$$

Estas condiciones implican que una  $\sigma$ -álgebra es cerrada también para intersecciones contables:

$$A_n \in F \text{ para } n \geq 1 \implies \bigcap_{n \geq 1} A_n \in F .$$

La  $\sigma$ -álgebra generada por una cierta familia  $B$  de subconjuntos de  $\Omega$  se designa como  $\sigma(B)$ , y se define como la  $\sigma$ -álgebra más pequeña que contiene a  $B$ , y es aquella que resulta de la intersección de todos las  $\sigma$ -álgebras sobre  $\Omega$  que contienen a  $B$ .

**MEDIDA** La medida de un conjunto resulta una generalización de la noción intuitiva del largo de un segmento, el área de una figura y el volumen de un cuerpo, y

usualmente se dice que corresponde a la masa de un conjunto  $A \subseteq \Omega$  dada una cierta distribución de masa sobre  $\Omega$ .

Una medida  $\mu$  es una función de conjuntos definida sobre un espacio medible  $(\Omega, \mathcal{F})$ ,

$$\mu : \mathcal{F} \rightarrow \mathbb{R}^+,$$

que satisface las siguientes propiedades:

1. La medida del conjunto vacío es nula,  $\mu(\emptyset) = 0$ .
2. La medida posee aditividad contable, esto es, la medida de una unión contable de elementos de  $\mathcal{F}$  es igual a la suma de la medida de los elementos,

$$A_{n \geq 1} \in \mathcal{F}, A_i \cap A_j = \emptyset \text{ si } i \neq j \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Una medida es finita cuando  $\mu(\Omega) < \infty$ , e infinita cuando  $\mu(\Omega) = \infty$ . Una medida sobre una  $\sigma$ -álgebra  $\mathcal{F}$  se denomina  $\sigma$ -finita si existe una colección contable de subconjuntos  $A_1, A_2, \dots \in \mathcal{F}$  que satisface: i)  $\bigcup_{n \geq 1} A_n = X$  y ii)  $\mu(A_n) < \infty$ .

*Espacio de probabilidad*

Un concepto central en la teoría axiomática de la probabilidad es el constructo matemático denominado espacio de probabilidad. Un espacio de probabilidad es simplemente un espacio de medida  $(\Omega, \mathcal{F}, \mu)$  donde  $\mu$  es una medida de probabilidad, es decir, cumple con la condición  $\mu(\Omega) = 1$ ; por convención, una medida de probabilidad se suele indicar con el símbolo  $P$ . En terminología tradicional, la probabilidad de que ocurra un cierto evento, un subconjunto  $A$  de  $\Omega$ , varía en el rango  $[0, 1]$ , con la condición de que debe suceder algún evento de  $\Omega$ :  $P(\emptyset) = 0$  y  $P(\Omega) = 1$ . Asimismo, la probabilidad de la unión de eventos disjuntos es igual a la sumatoria de la probabilidad de los eventos:  $P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$ .

En el ámbito de la teoría de la probabilidad, el espacio muestral  $\Omega$  representa el conjunto de todos los resultados posibles de un experimento aleatorio, también llamados sucesos elementales. Por su parte, una  $\sigma$ -álgebra  $\mathcal{F}$  de  $\Omega$  es una colección de subconjuntos de  $\Omega$  a los que puede asignarse una probabilidad de ocurrencia, y en este marco es común referirse a los conjuntos medibles  $A \in \mathcal{F}$  como eventos o sucesos.

*Generación de una función de medida*

La definición de un espacio de probabilidad, así como de un espacio de medida en el caso más general, depende de la capacidad de formular para un problema de interés dado una medida  $\mu$  sobre una  $\sigma$ -álgebra  $F$  de un espacio muestral  $\Omega$ . Una posibilidad es realizar una representación explícita de una  $\sigma$ -álgebra sobre  $\Omega$ , y mostrar luego que una función  $\mu$  cumple con las propiedades de una medida. Esta estrategia, sin embargo, no es en general la recomendada en la bibliografía (Pollard 2002; Athreya y Lahiri 2006). En su lugar, se suele partir de una clase de subconjuntos de  $\Omega$  que posean ciertas condiciones, y en donde sea sencillo definir una función  $\mu$  con propiedades adecuadas, y que puede ser luego extendida a una  $\sigma$ -álgebra que contenga a dichos subconjuntos de  $\Omega$ .

**MEDIDA SOBRE UNA SEMIÁLGEBRA** Para poder definir una medida, y así luego un espacio de probabilidad, se debe partir de una colección de subconjuntos  $S$  de  $\Omega$  que sea una semiálgebra de  $\Omega$ , esto es, que cumpla las siguientes condiciones:

1.  $S$  debe ser cerrada para intersecciones finitas:  $A, B \in S \Rightarrow A \cap B \in S$ .
2. Para todo elemento de  $S$ , su complemento puede ser formado por la unión finita de elementos disjuntos de  $S$ :  $A \in S \Rightarrow \exists B_{i=1}^n \subseteq S, 1 < n < \infty, i \neq j \Rightarrow B_i \cap B_j = \emptyset, A^c = \cup_{i=1}^n B_i$ .

Una cierta función  $\mu$  sobre una semiálgebra  $S$ ,  $\mu : S \rightarrow \mathbb{R}^+$ , es una medida si: i) es nula para el conjunto vacío; y ii) posee aditividad contable, es decir, la medida de la unión de elementos disjuntos de  $S$  es igual a la suma de la medida de dichos elementos.

**INDUCCIÓN DE UNA MEDIDA** Dada una medida  $\mu$  sobre una semiálgebra  $S$  sobre  $\Omega$ , se dice que  $\mu$  induce una medida  $\mu^*$  sobre cualquier subconjunto de  $\Omega$ . Sean  $S_l$  y  $S_u$  el conjunto de las uniones e intersecciones contables, respectivamente, de los elementos  $S$ , es posible aproximar el valor de  $\mu^*$  para cualquier elemento  $A$  de  $\sigma(S)$ . Se dice que la medida  $\mu^*$  es aproximada por el interior si

$$\mu^*(A) \equiv \sup\{\mu(F) : A \subseteq F \in S_l\},$$

mientras que es aproximada por el exterior si

$$\mu^*(A) \equiv \inf\{\mu(G) : A \supseteq G \in S_u\}.$$

En otros términos,  $\mu^*(A)$  puede ser aproximada como el supremo de las intersecciones contables de elementos de  $S$ , y también por el ínfimo de las uniones contables de

elementos de  $S$ . En ambos casos corresponderá a la colección contable de subconjuntos disjuntos de  $A$  donde ambas fórmulas son coincidentes. Un subconjunto  $A$  de  $\Omega$  se define como  $\mu^*$ -medible si para todo  $E \subset \Omega$ :  $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$ . La colección de los subconjuntos  $\mu^*$ -medibles constituyen una  $\sigma$ -álgebra, y  $\mu^*$  evaluada en dicho conjunto cumple por definición con los requisitos de una medida.

El teorema de extensión de Caratheodory asegura que dada una medida  $\mu$  sobre una semiálgebra  $S$  puede ser extendida a una medida  $\mu^*$  sobre una  $\sigma$ -álgebra formada por los subconjuntos  $\mu^*$ -medibles, posibilitando de este manera la definición de un espacio de medida.

*Funciones medibles*

Dados dos espacios medibles  $(X, \mathbf{A})$  e  $(Y, \mathbf{B})$ , se define una función  $(\mathbf{F}, \mathbf{G})$ -medible como una  $f : X \rightarrow Y$  que cumple con la condición

$$B \in \mathbf{B} \implies f^{-1}(B) \in \mathbf{A}$$

es decir, en la que para todo elemento  $B \in \mathbf{B}$  su imagen inversa  $f^{-1}(B)$  es un conjunto medible de  $X$ . En el caso de las funciones reales sobre  $\Omega$ ,  $f : \Omega \rightarrow \mathbb{R}$ , un tipo de  $\sigma$ -álgebra muy importante resulta la de los conjuntos de Borel, formada por la clase de todos los subconjuntos abiertos y que se indica como  $\mathbf{B}(\mathbb{R})$ . Si  $f : \Omega \rightarrow (a, b)$  con  $a, b \in \mathbb{R}$  y  $a < b$ , la  $\sigma$ -álgebra de Borel correspondiente se indica como  $\mathbf{B}(a, b)$ .

Dadas funciones  $f_1, \dots, f_k (k \in \mathbb{N})$ , con  $f_i : \Omega \rightarrow \mathbb{R}$  y en todos los casos  $(\mathbf{F}, \mathbf{B}(\mathbb{R}))$ -medibles, se cumplen las siguientes condiciones:

- Su adición es una función  $(\mathbf{F}, \mathbf{B}(\mathbb{R}))$ -medible.
- Su producto es una función  $(\mathbf{F}, \mathbf{B}(\mathbb{R}))$ -medible.
- La función  $f \equiv (f_1, \dots, f_k) : \Omega \rightarrow \mathbb{R}^k$  es  $(\mathbf{F}, \mathbf{B}(\mathbb{R}^k))$ -medible.
- La composición de  $f \equiv (f_1, \dots, f_k) : \Omega \rightarrow \mathbb{R}^k$  con una función continua  $g : \mathbb{R}^k \rightarrow \mathbb{R}^p$  con  $p \in \mathbb{N}$ , y por lo tanto medible, es una función  $(\mathbf{F}, \mathbf{B}(\mathbb{R}^p))$ -medible.

Por lo tanto, el conjunto de funciones  $(\mathbf{F}, \mathbf{B}(\mathbb{R}))$ -medibles de  $\Omega$  a  $\mathbb{R}$  es cerrado bajo operaciones finitas de adición y multiplicación, así como de producto por un escalar.

Si se considera el conjunto de los reales extendidos  $\hat{\mathbb{R}} = \mathbb{R} \cup -\infty, +\infty$ , una  $\sigma$ -álgebra  $\mathbf{B}(\hat{\mathbb{R}}) = \sigma(\mathbf{B}(\mathbb{R}) \cup -\infty, +\infty)$ , y una secuencia de funciones  $f_n : \Omega \rightarrow \hat{\mathbb{R}}$  medibles, el supremo, ínfimo, y sus límites resultan funciones medibles.

En la práctica, la verificación de que una función es medible puede ser difícil acuerdo a la definición dada, ya que se debería abarcar de forma explícita a todos los elementos de la  $\sigma$ -álgebra. Es posible utilizar la siguiente condición suficiente para la mensurabilidad de una función  $f$ : si la preimagen de todos los elementos de un conjunto  $A$  del

codominio de  $f$  pertenecen al dominio, y  $A$  es utilizado para generar una  $\sigma$ -álgebra  $\sigma(A)$  que forma un espacio medible junto al codominio, entonces  $f$  es  $\sigma(A)$ -medible.

### *Integral de Lebesgue–Stieltjes*

La integral de Lebesgue es una generalización que extiende la noción de la integral de Riemann, aquella enseñada tradicionalmente en cursos de cálculo integral y diferencial, a funciones para las que no está definida una integral de Riemann, y depende de la definición de la medida de Lebesgue sobre los reales. La integral de Lebesgue–Stieltjes es a su vez una generalización de la integral de Lebesgue para un conjunto de medidas; a veces se denomina también integral de Lebesgue–Radon.

**MEDIDAS DE LEBESGUE-STIELTJES** Considérese una función no decreciente  $F : \mathbb{R} \rightarrow \mathbb{R}$ , una semiálgebra

$$S \equiv (a, b] : -\infty \leq a \leq b \leq \infty \cup (a, \infty) : -\infty \leq a \leq \infty,$$

y una medida sobre  $S$  definida como  $\mu_F((a, b]) = F(b+) - F(a+)$ ,  $\mu_F((a, \infty)) = F(\infty) - F(a+)$ , con  $F(\infty) = \lim_{x \rightarrow \infty} F(x)$ . Entonces, el espacio generado por la extensión de Carathéodory de  $\mu_F$ ,  $(\mathbb{R}, \mathbf{M}_{\mu_F^*}, \mu_F^*)$  se denomina un espacio de medida de Lebesgue–Stieltjes, y se dice que  $\mu_F^*$  es la medida de Lebesgue–Stieltjes generada por  $F$ .

En particular, si  $F(x) \equiv x$  con  $x \in \mathbb{R}$ ,  $\mu_F^*$  se denomina medida de Lebesgue en  $\mathbb{R}$ , y  $\mathbf{M}_{\mu_F^*}$  es denominada la clase de los conjuntos Lebesgue-medibles. La medida de Lebesgue de un conjunto Lebesgue-medible  $R$  coincide con la noción intuitiva de longitud de un intervalo, pero su definición es más general y hace posible su extensión a otros conjuntos a los que no es posible asignarles una longitud — pero sí una medida de Lebesgue.

El argumento se puede generalizar para una función  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ ; en este caso la medida de Lebesgue–Stieltjes coincide con la medida de Lebesgue cuando  $F(x_1, \dots, x_n) = \prod_{i=1}^n x_i$ .

**INTEGRAL DE FUNCIONES SIMPLES NO NEGATIVAS** Una función  $F$ -medible  $f : \Omega \rightarrow \hat{\mathbb{R}}$  se denomina función simple si toma un conjunto finito de valores posibles; formalmente, se define de la siguiente manera. Sea  $\{\alpha_1, \dots, \alpha_k\} \in \hat{\mathbb{R}}$  el conjunto finito de  $k \in \mathbb{N}$  elementos de los reales extendidos que la función  $f$  puede tomar, y  $A_1, \dots, A_k \in \mathcal{F}$  definidos como  $A_i = \{\omega : f(\omega) = \alpha_i\}$ . Una función simple  $f$  puede escribirse como

$$f(\omega) = \sum_{i=1}^k \alpha_i I_{A_i}$$

donde  $I_{A_i}$  es la función índice de  $A_i$ . La función índice de un conjunto  $A \subseteq \Omega$ , también llamada función indicadora o característica, se define como

$$I_A(\omega) \equiv \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A. \end{cases}$$

La integral una función  $f$  simple no negativa en el espacio de medida  $(\Omega, \mathcal{F}, \mu)$ , o más sintéticamente la integral de  $f$  respecto de  $\mu$ , se define como

$$\int_{\Omega} f(\omega) d\mu(\omega) \equiv \sum_{i=1}^k \alpha_i \mu(A_i).$$

La notación  $\mu(f)$  también es utilizada para representar la integral de  $f$  en el espacio de medida  $(\Omega, \mathcal{F}, \mu)$ . La medida  $\mu$  de un conjunto  $A \in \mathcal{F}$  se suele indicar con la notación  $\mu(A)$ , y se define como la integral respecto de su función característica con respecto a  $\mu$  se define como

$$\mu(A) \equiv \int_{\Omega} I_A(\omega) d\mu(\omega).$$

Se observa que la integral de una función simple es la suma ponderada de los valores que toma la función por la medida del subconjunto del dominio en el que la función toma dicho valor. La integral se denomina de Lebesgue-Stieltjes si la medida  $\mu$  es una medida de Lebesgue-Stieltjes. En el caso particular de que la medida  $\mu$  se trate de la medida de Lebesgue se denomina integral de Lebesgue.

**INTEGRAL DE FUNCIONES MEDIBLES NO NEGATIVAS** Dada una función medible no negativa  $f : \Omega \rightarrow \hat{\mathbb{R}}^+$  sobre el espacio de medida  $(\Omega, \mathcal{F}, \mu)$ , y  $\{f_n\}_{n \geq 1}$  una secuencia no decreciente de funciones simples no negativas, tal que aproximen por defecto  $f_n(\omega) \uparrow f(\omega)$  para todo  $\omega \in \Omega$ , la integral de  $f$  respecto de  $\mu$  se define como:

$$\int_{\Omega} f(\omega) d\mu(\omega) \equiv \lim_{n \rightarrow \infty} \int_{\Omega} f_n(\omega) d\mu(\omega).$$

La integral de  $f$  sobre  $A \in \mathcal{F}$  con respecto a  $\mu$  se define como

$$\int_A f(\omega) d\mu(\omega) \equiv \int_{\Omega} f(\omega) I_A(\omega) d\mu(\omega).$$

Una función medible no negativa  $f$  es integrable respecto de  $\mu$  si resulta

$$\int_{\Omega} f(\omega) d\mu(\omega) < \infty.$$

Dada una función medible no negativa  $f$ , siempre es posible construir una secuencia no decreciente de funciones simple que aproxime  $f$  por defecto. El llamado teorema de convergencia monótona asegura que la integral de Lebesgue–Stieltjes es válida en el caso más general de una secuencia  $f_n$  de funciones medibles no negativas, no necesariamente simples, que converga a  $f$  en todo  $\Omega$  excepto en el subconjunto de medida nula  $A : \{\omega : \omega \in A, \mu(A) = 0\}$ , propiedad a la que se refiere con el léxico de convergencia en casi todo punto.

### Diferenciación

DERIVADA DE RADON-NIKODYM Toda función medible  $f$  en el espacio de medida  $(\Omega, \mathcal{F}, \mu)$  define una nueva medida  $\nu : \mathcal{F} \rightarrow \mathbb{R}$ ,

$$\nu(A) \equiv \int_{\Omega} I_A(\omega) \cdot f(\omega) d\mu(\omega),$$

con  $A \in \mathcal{F}$ , y  $\nu$  resulta una medida en el espacio medible  $(\Omega, \mathcal{F})$ . Se dice que  $\nu$  es absolutamente continua respecto de  $\mu$ , o de forma equivalente que es dominada por  $\mu$ , o que  $\nu \ll \mu$ .

El teorema de Radon-Nikodym establece las condiciones que aseguran la existencia de tal función  $f$ . Dada las medidas  $\sigma$ -finitas  $\nu$  y  $\mu$  sobre un espacio medible  $(\Omega, \mathcal{F})$ , si  $\nu$  está dominada por  $\mu$ , es decir,  $\mu(A) = 0 \Rightarrow \nu(A) = 0$ , con  $A \in \mathcal{F}$ , entonces existe una función  $f$  medible en el espacio  $(\Omega, \mathcal{F}, \mu)$  para la que se verifica la ecuación precedente. Dicha función es llamada derivada de Radon-Nikodym de  $\nu$  con respecto a  $\mu$ , y se denota como

$$f(\omega) = \frac{d\nu(\omega)}{d\mu(\omega)}.$$

Se dice que la medida  $\nu$  posee la densidad  $f(\omega)$  respecto de  $\mu$ .

DECOMPOSICIÓN DE LEBESGUE Dadas dos medidas  $\mu$  y  $\nu$  sobre un espacio medible  $(\Omega, \mathcal{F})$ , si existe un conjunto  $A \in \mathcal{F}$  donde se cumple:

$$\mu(A) = 0, \nu(\Omega - A) = 0,$$

se dice que  $\mu$  es singular respecto de  $\nu$  y viceversa, y que  $\mu$  y  $\nu$  son mutuamente singulares, lo que se indica como  $\mu \perp \nu$ .

El teorema de decomposición de Lebesgue asegura que dadas dos  $\mu$  y  $\nu$  sobre un espacio medible  $(\Omega, \mathcal{F})$ , la medida  $\mu$  puede descomponerse como la suma de dos medidas  $\sigma$ -finitas, una singular respecto de  $\nu$  y otra dominada por  $\nu$ :

$$\mu = \mu_a + \mu_s,$$

de manera que  $\mu_a \ll \nu$ , y  $\mu_s \perp \nu$ .

**TEOREMA FUNDAMENTAL DE LEBESGUE DEL CÁLCULO INTEGRAL** Una función  $F: [a, b] \rightarrow \mathbb{R}$  es absolutamente continua si y solo si existe una función  $f: [a, b] \rightarrow \mathbb{R}$  Lebesgue-medible e integrable respecto de la medida de Lebesgue  $m$  tal que:

$$F(x) = F(a) + \int_{[a,x]} f(x) dm(x), \quad a \leq x \leq b.$$

Una función  $F$  absolutamente continua es diferenciable respecto de la medida de Lebesgue, y la función  $f$  es la derivada de  $F$  respecto de la medida de Lebesgue.

#### *Variables aleatorias*

Dado un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$ , una variable aleatoria es una función real  $X: \Omega \rightarrow \mathbb{R}$  que resulta  $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -medible. De forma análoga, un vector aleatorio  $k$ -dimensional es una función  $X: \Omega \rightarrow \mathbb{R}^k$  que resulta  $(\mathcal{F}, \mathcal{B}(\mathbb{R}^k))$ -medible. De manera más general, un elemento aleatorio es una función  $X: \Omega \rightarrow E$  que resulta  $(\mathcal{F}, \mathcal{E})$ -medible.

**DISTRIBUCIÓN DE PROBABILIDAD** Dados dos espacios medibles  $(\Omega_1, \mathcal{F}_1)$  y  $(\Omega_2, \mathcal{F}_2)$ , y una función  $f: \Omega_1 \rightarrow \Omega_2$ , cualquier medida  $\mu$  sobre el espacio  $(\Omega_1, \mathcal{F}_1)$  induce una medida sobre  $\mathcal{F}_2$  definida como:

$$\mu f^{-1}(A) \equiv \mu(f^{-1}(A)), \quad A \in \mathcal{F}_2.$$

$\mu f^{-1}$  suele denominarse medida imagen de  $\mu$  bajo  $f$ , y se puede designar también como  $f(\mu)$ .

Dado un elemento aleatorio  $X: \Omega \rightarrow E$ , la distribución de probabilidad de  $X$  se suele indicar como  $P_X$ , y es la medida inducida en  $E$  por  $X$  bajo  $P$ :

$$P_X(A) \equiv P(X^{-1}(A)), \quad A \in \mathcal{E}.$$

**TIPOS DE VARIABLES ALEATORIAS** Una variable aleatoria se denomina continua si  $P_X(x) = 0$  para todo  $x \in \mathbb{R}$ , mientras que se denomina discreta si existe un conjunto

contable  $A \subset \mathbb{R}$  tal que  $P_X(A) = 1$ . Asimismo, una variable aleatoria puede ser mixta, y en ese caso su distribución puede ser descompuesta de acuerdo al teorema de descomposición de Lebesgue como una suma ponderada de una distribución continua y singular.

**FUNCIÓN DE DISTRIBUCIÓN** La función de distribución acumulada, o simplemente función de distribución, de una variable aleatoria  $X$  se define como

$$F_X(x) \equiv P_X((-\infty, x]), \quad x \in \mathbb{R},$$

y posee las siguientes propiedades:

1.  $F$  es no decreciente en  $\mathbb{R}$ .
2.  $F$  es continua por derecha en  $\mathbb{R}$ .
3.  $\lim_{x \rightarrow -\infty} F(x) = 0$ , y  $\lim_{x \rightarrow \infty} F(x) = 1$ .

De forma más general, dado un vector aleatorio  $X = (X_1, \dots, X_k)$ , la función de distribución acumulada resulta  $F_X = P(X \leq x)$ ,  $x \in \mathbb{R}^k$ , considerando que  $x, y \in \mathbb{R}^k$ ,  $x \leq y \Leftrightarrow x_i \leq y_i$  con  $i = 1, \dots, k$ .

Dado que hay una correspondencia entre una variable aleatoria  $X$  y la medida de probabilidad  $P_X$  inducida por  $X$ , se dice que  $P_X$  tiene una función de distribución  $F_X(x)$ . De forma más general, se habla de que una medida sobre  $\mathcal{B}(\mathbb{R})$  posee una función distribución  $F$ . Toda función de distribución acumulada es  $\mathcal{B}(\mathbb{R})$ -medible.

**FUNCIÓN CUANTIL** La inversa de una función de distribución acumulada  $F_X$  de una variable aleatoria  $X : \Omega \rightarrow \mathbb{R}$  se define como la función cuantil:

$$F_X^{-1}(y) \equiv \inf\{X(\omega) : F_X(X(\omega)) \geq y\}, \quad 0 < y < 1. \quad (3.3)$$

### *Espacios métricos*

Los espacios métricos son un tipo particular de espacio topológico y son una clase particularmente importante de espacio muestral ya que incluyen al espacio euclídeo  $\mathbb{R}^n$ .

**ESPACIO TOPOLÓGICO** Un espacio topológico es la clase más general de espacio matemático en el que pueden definirse conceptos como continuidad, conectividad y convergencia. Un espacio topológico está formado por un par  $(\Omega, T)$ , donde  $\Omega$  es un espacio muestral y  $T$  es una colección de subconjuntos de  $\Omega$  tal que:

1.  $\Omega \in T$ ;

2. una unión de elementos de  $T$ , finita o infinita, pertenecen a  $T$ ;
3. y una intersección finita de elementos de  $T$  pertenecen a  $T$ .

Se dice que  $T$  es una topología sobre  $\Omega$ , y los elementos de  $T$  se denominan conjuntos abiertos.

**ESPACIO DE BOREL** La  $\sigma$ -álgebra de Borel de un espacio topológico se define como el  $\sigma$ -álgebra generado por el conjunto de conjuntos abiertos de un espacio topológico. De esta manera, un espacio de Borel designa a un espacio medible  $(\Omega, \mathcal{B})$  conformado por un espacio topológico  $\Omega$  y una  $\sigma$ -álgebra de Borel  $\mathcal{B}$ .

**ESPACIO MÉTRICO** Un espacio métrico es un espacio topológico en el que es posible asignar una distancia entre dos elementos cualesquiera. El concepto de métrica es una generalización de la distancia euclídea, aquella que se corresponde con el concepto intuitivo de distancia en línea recta. En términos formales, un espacio métrico consiste en un par  $(M, d)$ , donde  $M$  es un conjunto y  $d$  una métrica que se define como una función  $d : M \times M \rightarrow \mathbb{R}$  tal que para todo  $x, y, z \in M$  se cumple: i)  $d(x, y) = d(y, x)$ , ii)  $d(x, y) = 0 \iff x = y$ , y iii)  $d(x, z) \leq d(x, y) + d(y, z)$ ; de estas propiedades se deduce  $d(x, y) \geq 0$ . Todo espacio métrico define un espacio topológico, y se dice que una métrica  $d$  induce una topología sobre  $M$ , debido a que quedan definidos los conjuntos abiertos.

### 3.5 FORMULACIÓN DE UN MODELO NULO PROBABILÍSTICO A PARTIR DE UN PROBLEMA DE OPTIMIZACIÓN

En esta sección se presenta cómo y en qué condiciones es posible derivar de forma general un modelo nulo a partir de un problema de optimización simple con sustento en la teoría axiomática de la probabilidad. Abordaremos el caso de un problema de optimización en su formulación clásica, conformada por un espacio muestral, restricciones y una única función objetivo, que constituye el planteo más común y que ha sido aplicado en mayor medida a la modelización de sistemas biológicos.

Los modelos nulos tienen una larga tradición en el contraste de hipótesis en biología (Gotelli y Ulrich 2012; Veech 2012). Gotelli y Graves (1996) definen un modelo nulo como un modelo generador de patrones que se basa en la aleatorización, donde ciertos elementos del modelo son mantenidos constantes mientras que se permite que otros varíen estocásticamente para producir un patrón que se esperarí en ausencia de un fenómeno en particular. De esta manera, un modelo se considera como un modelo nulo cuando omite en su formulación únicamente un proceso o mecanismo de interés, y es típicamente utilizado para derivar una distribución de probabilidad que puede compararse con datos observados. Para la construcción de un modelo nulo que permita la

evaluación de hipótesis de optimización resulta necesario establecer qué se esperaría en *ausencia* del proceso de optimización. En consecuencia, en esta sección no se realizará mención alguna a la función objetivo, sino que los únicos elementos que se tendrán en cuenta serán el espacio muestral y las restricciones de un modelo de optimización.

En base a un desarrollo en el marco de la teoría axiomática de la probabilidad se verá cómo formular, a partir de modelo de optimización arbitrario, un modelo nulo en la forma de un espacio de medida, y bajo qué condiciones es posible definir un espacio de probabilidad. Las ideas centrales que se desarrollan en esta sección se enumeran a continuación:

1. Es posible definir un espacio de medida sobre un espacio muestral arbitrario utilizando una medida que generaliza la noción de tamaño de un conjunto.
2. Todo conjunto de restricciones medibles define un nuevo espacio de medida sobre el espacio muestral.
3. Para derivar un espacio de probabilidad es necesario y suficiente que la medida del espacio muestral definida por el conjunto de las restricciones sea finita.

### 3.5.1 Definición de un espacio de medida estándar sobre el espacio muestral

La generalidad de la teoría axiomática de la probabilidad deriva de la teoría de conjuntos, por lo que aquí se tratará al espacio muestral como un conjunto abstracto. Lógicamente, es deseable mantener el mayor grado de generalidad, pero se verá que resulta conveniente restringir el desarrollo que se presenta a continuación a una clase de espacios medibles denominada Borel estándar, o simplemente estándar. Un espacio medible compuesto por un espacio muestral  $\Omega$  y una  $\sigma$ -álgebra  $F$  resulta Borel estándar si es un isomorfismo de alguno de los siguientes casos: un espacio contable finito  $\{1, \dots, n\}$  o contable infinito  $(\mathbb{N})$  con sus respectivas  $\sigma$ -álgebras discretas; el espacio medible  $([0, 1], \mathcal{B}([0, 1]))$ , o productos de estos espacios (Çınlar 2011, p. 11); este es un resultado que depende del llamado teorema del isomorfismo (Kechris 2012, p. 90). Los espacios medibles estándar incluyen una categoría extremadamente amplia de espacios muestrales, que incluyen a espacios contables y a espacios métricos completos y separables, comúnmente denominados espacios polacos, que incluye el espacio euclídeo  $R^n$ , con  $1 \leq n \leq \infty$ , así como espacios producto de dichos espacios. Por lo tanto, acotar el tratamiento a espacios estándar no representa una pérdida de generalidad en la inmensa mayoría de los problemas de optimización, a la vez que se gana en conveniencia por las propiedades que estos espacios poseen.

En el marco de la teoría de la medida es posible definir medidas que generalizan el concepto de tamaño de un conjunto, y que varían de acuerdo a la naturaleza del espacio muestral sobre el que se definen. Cuando el espacio muestral de un problema

de optimización se trata de un espacio contable  $\Omega = \{\omega : \omega_i, i \in \mathbf{N}\}$ , es adecuada la medida de conteo, que se como el número de elementos de un conjunto dado. Cuando el espacio muestral es un espacio euclídeo  $\Omega \subseteq \mathbf{R}^n$  se utiliza la medida de Lebesgue  $m$  (Elemento 3.4), cuyo valor coincide con la noción intuitiva de tamaño; la medida de Lebesgue de un segmento, por ejemplo, coincide con la noción intuitiva de largo ( $m([a, b]) = b - a$  si  $a \leq b$ ). El espacio euclídeo resulta un caso particular de un espacio métrico, un constructo matemático donde se asocia una distancia a cada par de elementos de un conjunto. La medida de Hausdorff resulta una generalización de la medida de Lebesgue definida para el caso general de los espacios métricos (Lawrence y Gariepy 2015, p. 84 y 91-92). La medida de Hausdorff  $s$ -dimensional,  $H^s$ , coincide con la medida de conteo cuando  $s = 0$ , mientras que equivale a la medida de Lebesgue para un espacio euclídeo  $\mathbf{R}^n$  cuando  $s = n, n \in \mathbf{N}$ .

De forma aún más general, mediante la teoría de la medida de Loeb (1975) es posible utilizar un análisis no estándar para construir la medida de Lebesgue (Anderson 1982; Cutland 2000) y la medida de Hausdorff (Potgieter 2009) como una extensión de la medida de conteo. Considérese un espacio muestral  $\Omega$  correspondiente a un cierto problema de optimización. En base a lo antes expuesto, resulta posible en general definir una  $\sigma$ -álgebra  $F$  sobre  $\Omega$ , así como una cierta medida  $\tau : F \rightarrow [0, \infty]$  que generalice el concepto intuitivo de tamaño a la clase de subconjuntos de  $F$ , que resultan por definición  $\tau$ -medibles. En adelante, utilizaremos  $\tau$  para referirnos a una medida de tal naturaleza para un cierto  $\Omega$ . En síntesis, para cualquier problema de optimización definido sobre un espacio muestral que sea Borel estándar se podrá definir un espacio de medida  $\sigma$ -finito  $(\Omega, F, \tau)$ .

### 3.5.2 Modelo nulo como una medida sobre el espacio muestral

Un modelo nulo respecto de una hipótesis de optimización debe ser definido únicamente a partir del espacio muestral  $\Omega$  y las restricciones planteadas sobre este, sin referencia alguna a la función objetivo, ya que por definición ésta debe ser excluida del modelo nulo. Una restricción representa, como ya se ha descrito, una condición necesaria que debe cumplir un estado  $\omega$  del espacio muestral  $\Omega$  para ser una solución de un problema de optimización. El caso más sencillo y frecuente en problemas de optimización con restricciones se caracteriza por el planteo de una o más restricciones de tipo determinista. Las restricciones deterministas representan claramente una simplificación de la realidad para cualquier sistema biológico, ya que siempre se dispone de un conocimiento incompleto sobre éste; tal incertidumbre puede ser descrita mediante el planteo de las llamadas restricciones probabilísticas. Se abordará en primer lugar el planteo de restricciones deterministas, para luego tratar el caso más amplio de las restricciones probabilísticas, y definir entonces una clase de funciones medibles

que permiten expresar cualquier conjunto de restricciones que puedan plantearse sobre el espacio muestral de un problema de optimización. Finalmente mostraremos que todo conjunto de restricciones medibles permite definir un espacio de medida que generaliza el espacio de soluciones candidatas de un problema de optimización.

### 3.5.2.1 Restricciones deterministas

La definición de un problema de optimización con restricciones incluye, por definición, el planteo de una o más restricciones al proceso de optimización que conforman un conjunto  $C = \{C_1, \dots, C_k\}$   $k \in \mathbb{N}$ . Una restricción determinista se caracteriza por definir un subconjunto del espacio muestral  $C_i \subseteq \Omega$  sin incertidumbre alguna. Un conjunto de restricciones deterministas define, en consecuencia, un subconjunto del espacio muestral  $\Omega_c \subseteq \Omega$  que se denomina espacio de soluciones candidatas:

$$\Omega_c = \bigcap_{i=1}^k C_i. \quad (3.4)$$

La solución a un problema de optimización con restricciones deterministas consiste en aquellos estados  $\omega \in \Omega_c$  para los que la función objetivo toma un valor óptimo, es decir, un valor mínimo o máximo. La pertenencia de un estado del espacio muestral al subconjunto definido por una restricción,  $\omega \in C_i$ , representa una condición necesaria que debe cumplir la solución al problema de optimización, lo que significa que  $\omega \notin C_i \implies \omega \notin \Omega_c$ , así como que  $\omega \in \Omega_c \implies \omega \in C_i$ .

Restricciones de esta naturaleza pueden representarse mediante una función índice asociada al subconjunto  $C_i$ , que se define como

$$I_{C_i}(\omega) := \begin{cases} 1 & \text{si } \omega \in C_i, \\ 0 & \text{si } \omega \notin C_i. \end{cases} \quad (3.5)$$

La función índice asocia a cada elemento  $\omega$  del espacio muestral un valor de 1 si cumple con la restricción, y un valor de 0 en caso contrario, de manera que resulta  $I_{C_i}(\omega) = 0 \implies \omega \notin \Omega_c$ , y de forma equivalente  $\omega \in \Omega_c \implies C_i(\omega) = 1$ . En conjunto, por lo tanto, las funciones índice asociadas a las restricciones permiten obtener una función índice asociada al espacio de soluciones candidatas:

$$c(\omega) = \prod_{i=1}^k I_{C_i}(\omega), \quad (3.6)$$

de manera que  $\Omega_c = \{\omega : c(\omega) = 1\}$ . De aquí en adelante designaremos como  $c$  a una función definida en base al conjunto de restricciones de un problema de optimización, y que de este modo define al espacio de soluciones candidatas.

3.5.2.2 Restricciones probabilísticas

El planteo de restricciones deterministas es típico en el caso de problemas abstractos de optimización. Esta formulación presenta limitaciones cuando se trabaja con sistemas experimentales, donde habitualmente no es posible asignar una pertenencia binaria al espacio de soluciones candidatas a causa de una disponibilidad limitada de información. En estos casos, y de forma más general, las restricciones de un problema de optimización se pueden expresar como una función no solo del estado  $\omega$ , sino también de parámetros, que en términos formales se representan como un elemento perteneciente a otro conjunto  $\theta \in \Theta$ . Este tipo de restricciones se suele designar como restricciones probabilísticas debido a que los parámetros poseen algún tipo de distribución de probabilidad. En términos de la teoría de la probabilidad, el espacio de parámetros  $\Theta$  tiene asociado un cierto espacio de probabilidad  $(\Theta, F_\Theta, P_\Theta)$ , donde  $F_\Theta$  es una  $\sigma$ -álgebra de  $\Theta$  y  $P_\Theta$  una medida de probabilidad.

Dado un espacio muestral  $\Omega$  y un espacio de parámetros  $\Theta$ , su producto cartesiano se define como el conjunto de los pares ordenados de los elementos de ambos conjuntos:  $\Omega \times \Theta \equiv \{(\omega, \theta) : \omega \in \Omega, \theta \in \Theta\}$ . Una restricción probabilística se define como una función  $I_{C_i} : \Omega \times \Theta \rightarrow \{0, 1\}$ . En esta formulación  $\theta$  puede representar tanto un parámetro escalar, como un vector o una matriz. Una restricción de este tipo está asociada a un subconjunto  $C_i \subseteq \Omega \times \Theta$  y la función índice asociada resulta:

$$I_{C_i}(\omega, \theta) := \begin{cases} 1 & \text{si } (\omega, \theta) \in C_i, \\ 0 & \text{si } (\omega, \theta) \notin C_i. \end{cases} \tag{3.7}$$

Nótese que una restricción determinista representa un caso particular de la formulación más general de una restricción probabilística, ya que dado un espacio de parámetros arbitrario  $\Theta$ , una restricción determinista  $I_{C_i}(\omega)$  equivale a la restricción probabilística  $I_{C_i}(\omega, \theta) := I_{C_i}(\omega)$ .

De forma similar a lo visto en el apartado anterior, un conjunto de restricciones que responden a la formulación general presentada en la Ecuación 3.7 determinan la función índice del espacio de soluciones candidatas:

$$c(\omega, \theta) = \prod_{i=1}^k I_{C_i}(\omega, \theta). \tag{3.8}$$

Esta ecuación es válida para una combinación de cualquier tipo de restricciones dado que, como ya se dijo, las restricciones deterministas representan un caso particular de la formulación general. La función  $c(\omega, \theta)$  permite definir, a su vez, un subespacio de soluciones candidatas para un cierto  $\theta$  que llamaremos  $S_\theta \equiv \{\omega : c(\omega, \theta) = 1\}$ .

Trabajar con restricciones probabilísticas resulta necesario cuando un modelo de optimización no se plantea como un problema matemático abstracto sino que se busca

representar el conocimiento que se tiene sobre un sistema concreto bajo estudio a partir de los datos disponibles. En vista de la incertidumbre inherente a cualquier proceso de medición, no será posible en general conocer con exactitud el valor de  $\theta$ . En consecuencia, no se podrá asociar una pertenencia unívoca al espacio de soluciones candidatas para todos los estados  $\omega$ , ya que esta dependerá del valor de  $\theta$ , y no será posible establecer taxativamente si cumple o no con una cierta restricción. En otras palabras, la incertidumbre en el valor de  $\theta$  se traslada a la definición misma del espacio de soluciones candidatas.

### 3.5.2.3 Definición de una medida $\sigma$ -finita sobre el espacio de soluciones candidatas

Como paso previo a la definición de una medida de probabilidad sobre el espacio muestral  $\Omega$  es necesario poder definir una medida  $\sigma$ -finita sobre  $\Omega$  que sea adecuada para establecer la medida del espacio de soluciones candidatas definido por las restricciones; denominaremos  $\tau$  a tal medida.

A fin de simplificar el tratamiento matemático que sigue a continuación, asumiremos de aquí en adelante que si la función  $c(\omega, \theta)$  permite definir una medida de probabilidad  $P$  sobre el espacio medible producto  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$ , la medida de probabilidad marginal sobre  $\Omega$  definida por  $c(\omega, \theta)$  es absolutamente continua respecto a una cierta medida  $\sigma$ -finita  $\tau$  sobre  $\Omega$ . Para que se cumpla este supuesto es condición suficiente que:

1. la medida  $P_\Theta$  sea absolutamente continua respecto de una cierta medida base sobre  $\Theta$ ,
2. y que la medida de probabilidad  $P$  sobre  $\Omega \times \Theta$  sea absolutamente continua respecto de  $\tau \times P_\Theta$ .

Este supuesto no representa una pérdida de generalidad: lo tratado puede extenderse al caso de un problema de optimización que *no* cumpla este supuesto, ya que será posible descomponer tal problema en un conjunto contable de problemas que *sí* cumplen con la condición señalada. De acuerdo al teorema de descomposición de Lebesgue, toda medida puede descomponerse respecto de otra medida  $\sigma$ -finita en la suma de dos medidas, una absolutamente continua y otra mutuamente singular (Lawrence y Garipey 2015, p. 52). Por lo tanto, si  $c(\omega, \theta)$  es tal que *no* cumple con el supuesto anterior será posible descomponer la medida  $P$  en una sumatoria contable de  $n \in \mathbb{N}$  medidas  $P_i$  en la que cada una de estas es absolutamente continua respecto de una medida  $\sigma$ -finita  $\tau_i$  sobre  $\Omega$ .

Dada una función  $c(\omega, \theta)$  definida por un conjunto de restricciones probabilísticas asociadas a un espacio de parámetros  $(\Theta, \mathcal{F}_\Theta, P_\Theta)$ , la medida  $\tau$  debe cumplir con las siguientes condiciones:

1. La medida debe ser  $\sigma$ -finita sobre  $\Omega$  sobre una  $\sigma$ -álgebra  $\mathcal{F}$ .

2.  $c(\omega, \theta)$  debe ser una función medible en el espacio medible  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$ .
3. La medida de  $c(\omega, \theta)$  debe ser no nula en dicho espacio <sup>4</sup>.

Estos requerimientos son, como se verá más adelante, condiciones necesarias y suficientes para poder expresar el modelo nulo como un espacio de medida mediante la aplicación del teorema de Radon-Nikodym. De acuerdo a lo expuesto en la Sección 3.5.1, la definición de una cierta medida  $\sigma$ -finita sobre un espacio muestral dado no representa desafío alguno. En virtud del teorema de Tonelli (Pollard 2002, p. 88), la condición de que  $c(\omega, \theta)$  sea  $(\mathcal{F} \times \mathcal{F}_\Theta)$ -medible en el espacio producto es equivalente a que

$$c(\omega) \equiv \int_{\Theta} c(\omega, \theta) dP_{\Theta}(\theta) \tag{3.9}$$

sea una función  $\tau$ -medible. Por último, la condición de que la medida de  $c(\omega, \theta)$  sea no nula significa que es la medida  $\tau$  definida puede efectivamente cuantificar el tamaño del conjunto de soluciones candidatas. El cumplimiento de esta condición requiere de un abordaje más extenso ya que un conjunto no vacío puede tener una medida nula. En otros términos, la integral de una función respecto de una cierta medida puede ser nula aunque la función no lo sea, y en tal caso la medida no permitirá la definición de un espacio de probabilidad.

Consideraremos la definición de una medida que cumpla con las condiciones señaladas para los casos donde el espacio muestral  $\Omega$  sea:

- un espacio contable,
- un espacio métrico *no* contable,
- o el producto cartesiano de un número finito de espacios de los tipos mencionados.

Estos casos son de una generalidad tal que incluyen los espacios muestrales tratados comúnmente en los modelos de optimización.

**ESPACIO MUESTRAL CONTABLE** En caso de que  $\Omega$  se trate de un espacio contable, la definición de  $\tau$  como la medida de conteo satisface las tres condiciones antes señaladas. En primer lugar, resulta  $\sigma$ -finita sobre cualquier espacio contable. Por definición, cualquier subconjunto de un  $\Omega$  contable es medible respecto de la medida de conteo, por lo que  $c(\omega)$  resultará  $\tau$ -medible. Por último, la medida de conteo es nula solo para un conjunto vacío, de manera que cumple la tercera condición. En este caso, cualquier subconjunto de  $\Omega$  será un conjunto medible, y la  $\sigma$ -álgebra será entonces el conjunto potencia  $\mathcal{F} = \{A : A \subseteq \Omega\}$ , es decir, el conjunto formado por todos los subconjuntos posibles de  $\Omega$ .

<sup>4</sup> Con excepción del caso trivial  $c(\omega, \theta) = 0$  para todo  $\omega$  y  $\theta$ .

ESPACIO MUESTRAL MÉTRICO En el caso de que  $\Omega$  resulte un espacio métrico que *no* sea contable, se vio en la Sección 3.5.1 que una medida  $\sigma$ -finita sobre un  $\Omega$  de dimensión  $n$  es la medida  $n$ -dimensional de Hausdorff,  $H^n$ .

Ahora bien, en este caso existirán conjuntos no vacíos  $H^n$ -medibles que, sin embargo, tendrán una medida nula. Para comprender este hecho es necesario recurrir al concepto de la dimensión de Hausdorff de un conjunto, que se define como (Mattila 1999, p. 58):

$$\begin{aligned} \dim_H A &= \sup\{s : H^s(A) > 0\} = \sup\{s : H^s(A) = \infty\} \\ &= \inf\{s : H^s(A) = 0\} = \inf\{s : H^s(A) < \infty\}. \end{aligned} \quad (3.10)$$

Para aquellos  $A \subset \Omega$  que sean  $H^n$ -medibles pero para los que se verifique  $\dim_H A < n$  resultará  $H^n(A) = 0$ ; tales conjuntos se denominan  $H^n$ -nulos.

Es posible que las restricciones definan conjuntos de una geometría compleja cuya dimensión de Hausdorff puede ser no entera, como resulta el caso de conjuntos fractales (Falconer 2004). En términos formales, una función  $c(\omega)$  puede ser tal que defina un conjunto de soluciones candidatas  $\Omega_c \equiv \{\omega : c(\omega, \theta) = 1\}$  que sea  $H^n$ -nulo. En tal caso,  $\Omega_c$  tendrá una cierta dimensión de Hausdorff  $s = \dim_H \Omega_c < n$ , y resultará entonces un conjunto  $H^s$ -medible. Sin embargo,  $H^s$  resultará por definición  $\sigma$ -finita respecto de  $\Omega_c$ , pero *no* respecto de  $\Omega$  <sup>5</sup>. Ahora bien, es posible definir en base a  $H^s$  una medida  $\sigma$ -finita sobre  $\Omega$  mediante su restricción a  $\Omega_c$  <sup>6</sup>. Por último, cabe señalar que la condición de que  $\Omega_c$  sea un conjunto  $H^s$ -medible equivale a que  $c(\omega)$  sea una función  $H^s$ -medible <sup>7</sup>.

Para que quede definido un espacio de medida  $\sigma$ -finito resta definir una  $\sigma$ -álgebra. Si se define  $F$  como la  $\sigma$ -álgebra *completa* de los conjuntos  $H^n$ -medibles, por definición ésta incluirá a todos los conjuntos  $H^n$ -nulos (Katzourakis y Varvaruca 2018, p. 195). En consecuencia, los conjuntos  $H^s$ -medibles estarán incluidos en  $F$ , ya que para todo  $0 \leq s < n$  dichos conjuntos son  $H^n$ -nulos. Como la medida de Hausdorff coincide con la de Lebesgue para dimensiones enteras, entonces  $F = L(\mathbb{R}^n)$ , la  $\sigma$ -álgebra completa de los conjuntos Lebesgue-medibles de  $\mathbb{R}^n$ .

<sup>5</sup> Un  $\Omega$   $n$ -dimensional no podrá ser expresado como la unión contable de conjuntos  $H^s$ -finitos con  $0 \leq s < n$ , ya que tales conjuntos de dimensión  $n$  tendrán medida infinita respecto de  $H^s$  (Lawrence y Garipey 2015, p. 86).

<sup>6</sup> La restricción de una medida  $\mu$  por un  $A \subset X$  se define como  $\mu \lfloor A(B) \equiv \mu(A \cap B) \ B \subseteq X$  (Mattila 1999, p. 10).

<sup>7</sup> Sea  $M_{H^s}$  la  $\sigma$ -álgebra de los conjuntos  $H^s$ -medibles, y siendo el codominio de  $c(\omega)$  el intervalo  $[0, 1]$ , si  $c(\omega)$  es una función  $(M_{H^s}, B([0, 1]))$ -medible, la preimagen de  $(0, 1]$  será  $\Omega_c$ , que por definición estará incluido en  $M_{H^s}$ .

En consecuencia, si  $c(\omega)$  es  $H^s$  medible y  $\Omega_c = c(\omega)^{-1}((0, 1])$  es  $H^s$ -medible, entonces la siguiente medida  $\tau$  cumple las condiciones establecidas:

$$\tau(A) = H^s(A) \llcorner_{\Omega_c, s} = \dim_H(\Omega_c) \quad A \in F, \tag{3.11}$$

donde  $F$  es una  $\sigma$ -álgebra que incluye la  $\sigma$ -álgebra de los conjuntos  $H^s$ -medibles de  $\Omega$ . Por lo tanto, para una  $\tau$  definida de esta forma la medida de la función  $c(\omega)$  será una generalización válida de la noción intuitiva de tamaño del conjunto de soluciones candidatas.

**CASO GENERAL** Consideremos ahora el caso donde el espacio muestral  $\Omega$  es el producto cartesiano de un espacio contable  $\Omega_1$  y un espacio métrico no contable  $\Omega_2$ :  $\Omega = \Omega_1 \times \Omega_2$ . En este caso, un conjunto de restricciones probabilísticas definirá una función  $c(\omega_1, \omega_2, \theta)$  sobre  $\Omega_1 \times \Omega_2 \times \Theta$ . Por el teorema de Tonelli resultará que si

$$c(\omega_1, \omega_2) \equiv \int_{\Theta} c(\omega_1, \omega_2, \theta) dP_{\Theta}(\theta) \tag{3.12}$$

es medible sobre una  $\sigma$ -álgebra de  $\Omega_1 \times \Omega_2$ , entonces  $c(\omega_1, \omega_2, \theta)$  resultará medible en el correspondiente espacio producto.

Se podrá definir una medida que cumpla las condiciones establecidas como la medida producto  $\tau = \tau_1 \times \tau_2$ , donde:

- $\tau_1$  será la medida de conteo sobre el espacio contable  $\Omega_1$ .
- La medida  $\tau_2$  sobre el espacio métrico no contable se definirá como la restricción de la medida de Hausdorff  $H^s$  al conjunto  $\{\omega_2 : c(\omega_1, \omega_2) > 0\} \subseteq \Omega_2$ , con la dimensión  $s$  igual a la dimensión Hausdorff de dicho conjunto.

A partir de las respectivas  $\sigma$ -álgebras  $F_1$  y  $F_2$  queda definida la  $\sigma$ -álgebra producto  $F = F_1 \times F_2$ .

La medida producto  $\tau$  sobre el espacio medible producto  $(\Omega, F)$  cumplirá las condiciones establecidas, y de acuerdo al teorema de Fubini (Athreya y Lahiri 2006, p. 153) resulta

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} c(\omega_1, \omega_2) d(\tau_1 \times \tau_2) &= \int_{\Omega_1} \int_{\Omega_2} c(\omega_1, \omega_2) d\tau_2 d\tau_1 \\ &= \int_{\Omega_2} \int_{\Omega_1} c(\omega_1, \omega_2) d\tau_1 d\tau_2. \end{aligned} \tag{3.13}$$

Por último, dada las propiedades de un espacio producto, lo anterior se puede generalizar para el caso de un espacio muestral formado por el producto de un número contable de espacios contables y/o métricos no contables.

### 3.5.2.4 Definición general de restricciones medibles

En la sección anterior se trataron los requerimientos que debe cumplir una medida  $\tau$  para que sea posible, aunque no esté asegurada, la definición un espacio de probabilidad dado un espacio muestral y un conjunto arbitrario de restricciones. Abordaremos aquí la definición general de la clase de restricciones medibles para un problema de optimización sobre un espacio muestral  $\Omega$  dada una cierta  $\sigma$ -álgebra  $F$  para la que se ha definido una medida  $\tau$  que resulta  $\sigma$ -finita sobre  $\Omega$  respecto de  $F$ . Es importante destacar que la mensurabilidad es una condición lo suficientemente laxa como para que la inmensa mayoría de las restricciones consideradas en la literatura sobre optimización puedan expresarse mediante funciones medibles dado el espacio sobre el que están definidas. Todas las funciones continuas definidas sobre un espacio euclídeo  $\mathbb{R}^n$  resultan, por ejemplo, Lebesgue-medibles.

Una restricción probabilística, definida sobre un espacio medible producto  $(\Omega \times \Theta, F \times F_\Theta)$ , la condición de mensurabilidad equivale a que sea posible asignar una medida al subconjunto de  $\Omega \times \Theta$ , es decir, que sea  $(F \times F_\Theta, \sigma(\{0, 1\}))$ -medible.<sup>8</sup> Considérese ahora un espacio de probabilidad definido sobre el espacio de parámetros asociado a las restricciones  $(\Theta, F_\Theta, P_\Theta)$ . Como ya se mencionó, el teorema de Tonelli permite obtener una expresión de una restricción medible como función únicamente del espacio muestral mediante la integración de la distribución de probabilidad sobre el espacio de parámetros:

$$c(\omega) = \int_{\Theta} c(\omega, \theta) dP_{\Theta}(\theta). \quad (3.14)$$

Dado que por definición  $c(\omega, \theta)$  toma valores de 0 o 1, esto implica que  $0 \leq c(\omega) \leq 1$ , es decir, y el codominio de  $c(\omega)$  resulta  $[0, 1]$ . La Ecuación 3.16 también es válida para una función índice que define el espacio de soluciones candidatas de un conjunto arbitrario de restricciones (Ecuación 3.8), dado que el producto de funciones medibles resulta también una función medible.

Dado un cierto problema de optimización, designaremos como  $C$  a la clase de funciones  $c(\omega)$  sobre el espacio muestral que cumplen con la definición

$$c : (\Omega, F) \rightarrow ([0, 1], B([0, 1])), \quad (3.15)$$

donde  $B([0, 1])$  resulta la  $\sigma$ -álgebra de Borel sobre el intervalo  $[0, 1]$ . En adelante utilizaremos el término restricciones medibles, o simplemente restricciones, para referirnos a  $c \in C$ . Asimismo, cuando se distinga entre restricciones de tipo deterministas o probabilísticas se asumirá en ambos casos que resultan funciones medibles, de manera que éstas definen a su vez una cierta  $c \in C$ .

<sup>8</sup>  $\sigma(\{0, 1\})$  representa la  $\sigma$ -álgebra sobre  $\{0, 1\}$ :  $\{\emptyset, 0, 1, \{0, 1\}\}$ .

3.5.2.5 *Medida generada por las restricciones*

El teorema de Radon-Nikodym afirma que cualquier función medible no negativa sobre un espacio de medida  $\sigma$ -finito define una nueva medida y, en consecuencia, un nuevo espacio de medida. Por lo tanto, a partir de una medida  $\tau$  tal como ha sido definida (Sección 3.5.2.3), una restricción  $c \in C$  define una nueva medida para todo  $A \in F$  que designaremos como  $\tau_c$ :

$$\tau_c(A) = \int_{\Omega} I_A(\omega) \cdot c(\omega) \, d\tau(\omega), \tag{3.16}$$

donde  $I_A$  es la función índice de  $A$ . De esta manera queda definido un nuevo espacio de medida  $(\Omega, F, \tau_c)$ . La función  $c(\omega)$  resulta la densidad o derivada de Radon-Nikodym de  $\tau_c$  respecto de  $\tau$ :  $c(\omega) = \frac{d\tau_c(\omega)}{d\tau(\omega)}$ .

EL espacio de medida  $(\Omega, F, \tau_c)$  resulta una representación general del espacio de soluciones candidatas del problema de optimización. En el caso particular de un conjunto de restricciones deterministas medibles, la función  $c(\omega)$  se corresponde a una función índice de un subconjunto; en este caso dicho espacio de medida es equivalente a la definición del espacio de soluciones como un subconjunto del espacio muestral  $\Omega_c \subseteq \Omega$ . En el caso más general de un conjunto de restricciones probabilísticas,  $c(\omega)$  asigna valores en el rango  $[0, 1]$  a cada  $\omega$ , y el valor de la función puede interpretarse como la probabilidad de que un  $\omega$  dado pertenezca al espacio de soluciones candidatas. Por lo tanto, de forma general  $c$  no puede ser equiparada con una función índice, sino que define un espacio de soluciones candidatas de una forma más general.

3.5.3 *Definición de una medida de probabilidad*

En esta sección veremos que, para una clase muy amplia de espacios muestrales, es condición necesaria y suficiente que el espacio de soluciones candidatas posea una medida finita para definir un modelo nulo como un espacio de probabilidad. Mostraremos en primer lugar que, en caso de existir, una medida de probabilidad definida a partir de restricciones debe tener una distribución uniforme respecto de éstas. Analizaremos luego la definición de una medida de probabilidad sobre el espacio muestral  $\Omega$  a partir del espacio de medida  $\sigma$ -finito  $(\Omega, F, \tau_c)$  definido en la sección anterior. A continuación mostraremos que, dado un problema de optimización arbitrario, está asegurada la existencia de restricciones medibles que, de ser incorporadas al problema de optimización aseguran la definición de un espacio de probabilidad. Finalmente, trataremos la definición del modelo nulo como un espacio de probabilidad sobre el espacio producto  $\Omega \times \Theta$ , lo que requiere un tratamiento explícito de distribuciones de probabilidad condicionales.

### 3.5.3.1 La medida de probabilidad debe tener una distribución uniforme

Como se ha visto, un conjunto arbitrario de restricciones medibles equivalen a una función  $c(\omega)$  (Ecuación 3.15), cuyo valor corresponde a la pertenencia de cada estado  $\omega$  del espacio muestral al espacio de soluciones candidatas<sup>9</sup>. De existir una medida de probabilidad definida a partir de las restricciones, dicha medida debe reflejar únicamente la información aportada por tales restricciones respecto de la pertenencia al espacio de soluciones candidatas; tal información está contenida en la medida definida por las restricciones,  $\tau_c$ . La distribución de probabilidad que posee una máxima entropía sobre un cierto dominio resulta ser la distribución uniforme (Cover y Thomas 2012, p. 29). Por lo tanto, una medida de probabilidad definida por las restricciones, de existir, debe poseer una distribución uniforme, dado que solo tal medida poseería la información contenida en restricciones y *solo* dicha información. En el marco de la teoría de la medida, una distribución uniforme se caracteriza por una medida que posee una derivada de Radon-Nikodym constante respecto a una cierta medida base. Para una medida de probabilidad definida a partir de un espacio de medida  $(\Omega, \mathcal{F}, \tau_c)$ , la probabilidad de un cierto conjunto  $A \in \mathcal{F}$  resultará proporcional a su medida:  $P(A) \propto \tau_c(A)$ .

### 3.5.3.2 Condiciones para la definición de una medida de probabilidad

Supongamos que un espacio de medida  $(\Omega, \mathcal{F}, \tau_c)$  definido en base a una restricción medible  $c$  permite definir un espacio de probabilidad  $(\Omega, \mathcal{F}, P_\Omega)$ <sup>10</sup>. De acuerdo a lo antes expuesto,  $P_\Omega$  poseerá una distribución uniforme respecto a  $\tau_c$ . Dada la relación que existe entre la medida de un conjunto y la integral de su función índice, la medida de probabilidad de un conjunto  $A \in \mathcal{F}$  resultará

$$P_\Omega(A) = \int_{\Omega} I_A(\omega) dP_\Omega(\omega). \quad (3.17)$$

De acuerdo al teorema de Radon-Nikodym, lo anterior puede expresarse como

$$\int_{\Omega} I_A(\omega) dP_\Omega(\omega) = \int_{\Omega} I_A(\omega) \cdot \frac{dP_\Omega(\omega)}{d\tau_c(\omega)} \cdot \frac{d\tau_c(\omega)}{d\tau(\omega)} d\tau(\omega), \quad (3.18)$$

Se ha visto que la derivada de Radon-Nikodym de  $\tau_c$  respecto de  $\tau$  es la función  $c(\omega)$  (Ecuación 3.16). Asimismo, se vio que si la distribución de  $P_\Omega$  es uniforme respecto de

<sup>9</sup> Esta interpretación es posible si se asume que  $\tau_c$ , la medida marginal sobre  $\Omega$  definida respecto de un espacio de medida  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta, (\tau \times P_\Theta)_c)$  es absolutamente continua respecto de una medida base, es decir, su componente singular es nulo. En este caso  $c(\omega)$  resulta la derivada o densidad de Radon-Nikodym de  $\tau_c$  respecto de  $\tau$ .

<sup>10</sup> Utilizaremos  $P_\Omega$  para referirnos a la medida de probabilidad sobre  $\Omega$  ya que, de existir, esta corresponde a la medida de probabilidad marginal de un espacio de probabilidad  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta, P)$ .

$\tau_c$  su derivada de Radon-Nikodym, que denominaremos  $\nu$ , es constante:  $\nu \equiv \frac{dP_\Omega(\omega)}{d\tau_c(\omega)}$ . Reemplazando estas expresiones en la Ecuación 3.18 se obtiene

$$\int_{\Omega} I_A(\omega) dP_\Omega(\omega) = \int_{\Omega} I_A(\omega) \cdot \nu \cdot c(\omega) d\tau(\omega). \tag{3.19}$$

Por definición resulta  $1 = \int_{\Omega} dP_\Omega(\omega)$ , de manera que

$$1 = \nu \int_{\Omega} c(\omega) d\tau(\omega). \tag{3.20}$$

La derivada de Radon-Nikodym (o densidad) de  $P_\Omega$  respecto de  $\tau_c$  resulta por lo tanto

$$\nu = \frac{1}{\int_{\Omega} c(\omega) d\tau(\omega)}. \tag{3.21}$$

Para que la constante de normalización  $\nu$  esté definida resulta entonces necesario y suficiente que  $c(\omega)$  sea una función  $\tau$ -integrable y que su integral sea finita, lo que por definición equivale a que su medida sea finita y no nula:  $0 < \tau(c) < \infty \iff 0 < \tau_c(\Omega) < \infty$ . En otros términos, siempre que el espacio de soluciones candidatas definido por las restricciones posea una medida finita y no nula está asegurada la definición de un modelo nulo en la forma de un espacio de probabilidad.

### 3.5.3.3 Existencia de restricciones medibles integrables

La condición de un espacio de soluciones candidatas de medida finita es trivial cuando el espacio muestral  $\Omega$  resulta finito. En este caso, cualquier restricción medible  $c \in \mathbb{C}$  definirá un espacio de probabilidad. En el caso más general de que  $\Omega$  sea un conjunto con una cardinalidad contable infinita (o superior), dada una cierta medida  $\tau$  que sea  $\sigma$ -finita sobre  $\Omega$ , las restricciones  $\tau$ -medibles deberán tener una medida finita para definir un espacio de probabilidad. Verificaremos a continuación, recurriendo a nociones elementales de la teoría de la medida, que la existencia de restricciones  $\tau$ -medibles con medida finita y no nulas está asegurada para un espacio muestral arbitrario cualquier problema de optimización.

La función índice del espacio muestral  $I_\Omega$  constituye por definición una restricción medible trivial, y en caso de que  $\tau(\Omega) = \infty$  resultará una función no integrable. Para toda función medible no negativa está asegurada la existencia de una secuencia de funciones simples que converge por defecto a dicha función (Pollard 2002, p. 25). Dada una restricción  $\tau$ -medible  $g \in \mathbb{C}$ , sea  $\{g_s\}_n$  una secuencia de funciones simples que converge por defecto a  $g$ :

$$g = \lim_{n \rightarrow \infty} \{g_s\}_n, \quad 0 < g^{s_1} \leq \dots \leq g^{s_n} \leq g. \tag{3.22}$$

De acuerdo al teorema de la convergencia monótona (Athreya y Lahiri 2006, p. 52), la integral de una función medible se define como el límite de la integral de una secuencia de funciones cuyo límite converge por defecto dicha función. Por lo tanto

$$\int_{\Omega} g(\omega) d\tau(\omega) = \lim_{n \rightarrow \infty} \int_{\Omega} g_{S_n}(\omega) d\tau(\omega). \quad (3.23)$$

Dado que  $\tau$  es  $\sigma$ -finita, cualquier función simple  $g_{S_n}$  perteneciente a una secuencia  $\{g_{S_n}\}_n$  puede expresarse como

$$\lim_{n \rightarrow \infty} g_{S_n}(\omega) = \lim_{n \rightarrow \infty} \sum_{j=1}^n \alpha_j \cdot I_{A_j}(\omega), \quad (3.24)$$

donde  $A_{nn} \subseteq F$  es cualquier secuencia de conjuntos medibles disjuntos tales que  $\bigcup_{i=1}^{\infty} A_i = \Omega$  y  $\tau(A_i) < \infty$ . De forma equivalente,

$$\int_{\Omega} g(\omega) d\tau(\omega) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \alpha_i \tau(A_i). \quad (3.25)$$

Sea  $\alpha_i = \frac{\beta_i}{\tau(A_i)}$ , con  $\beta_1, \dots, \beta_n, 0 \leq \beta_i \leq 1, 1 \leq n \leq \infty, n \in \mathbb{N}$ . Entonces resulta que

$$\begin{aligned} \int_{\Omega} g(\omega) d\tau(\omega) &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \alpha_i \cdot \tau(A_i) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{\beta_i}{\tau(A_i)} \cdot \tau(A_i) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \beta_i. \end{aligned} \quad (3.26)$$

Si se define el conjunto de valores de  $\beta$  con la condición  $0 < \sum_{i=1}^n \beta_i < \infty$ , de manera que resultará

$$0 < \lim_{n \rightarrow \infty} \sum_{i=1}^n \beta_i < \infty. \quad (3.27)$$

En consecuencia, está asegurada la existencia de secuencias cuyo límite es finito, y que por lo tanto convergen a restricciones medibles integrables en el espacio de medida  $(\Omega, F, \tau)$ .

### 3.5.3.4 Existencia de restricciones que de ser incorporadas aseguran la definición de un espacio de probabilidad

Dado un problema de optimización con restricciones que no posean una medida finita, cabe preguntarse si está asegurada la existencia de nuevas restricciones que, de ser incorporadas, aseguren la definición de un espacio de probabilidad. Mostraremos

a continuación que la existencia de tales restricciones está asegurada para todo espacio muestral  $\Omega$  y toda restricción medible pero no integrable que pueda plantearse.

En términos formales, sea  $F \subset \mathbb{C}$  el conjunto de las restricciones medibles pero no integrables sobre el espacio de medida  $(\Omega, \mathcal{F}, \tau)$ ,

$$F = \{f(\omega) : f : \Omega \rightarrow [0, 1], \int_{\Omega} f(\omega) d\tau(\omega) = \infty\}. \tag{3.28}$$

Se busca demostrar que para toda  $f$  existe un conjunto no nulo de restricciones

$$G = \{g(\omega) : g : \Omega \rightarrow [0, 1], 0 < \int_{\Omega} f(\omega) \cdot g(\omega) d\tau(\omega) < \infty\}. \tag{3.29}$$

Dado que cualquier  $f \in F \subset \mathbb{C}$  y  $g \in G \subset \mathbb{C}$  son por definición funciones medibles no negativas, está asegurada la existencia de secuencias de funciones simples  $\{f_s\}_n$  y  $\{g_s\}_n$  que convergen por defecto a  $f$  y  $g$  respectivamente. Por propiedad del límite resultará que

$$\begin{aligned} f(\omega) \cdot g(\omega) &= \lim_{n \rightarrow \infty} \{f_s\}_n \cdot \lim_{n \rightarrow \infty} \{g_s\}_n \\ &= \lim_{n \rightarrow \infty} \{f_s\}_n \cdot \{g_s\}_n. \end{aligned} \tag{3.30}$$

Expresemos ahora las funciones simples  $f_{s_n}$  y  $g_{s_n}$  como

$$\begin{aligned} f_{s_n} &= \sum_{j=1}^n \alpha_j \cdot I_{A_j}(\omega) \\ g_{s_n} &= \sum_{j=1}^n \beta_j \cdot I_{A_j}(\omega), \end{aligned} \tag{3.31}$$

donde  $A_m \subseteq \mathcal{F}$  es cualquier secuencia de conjuntos medibles disjuntos tales que  $\bigcup_{i=1}^{\infty} A_i = \Omega$  y  $\tau(A_i) < \infty$ . En consecuencia,

$$\begin{aligned} \int_{\Omega} f(\omega) \cdot g(\omega) d\tau(\omega) &= \lim_{n \rightarrow \infty} \int_{\Omega} f_{s_n}(\omega) \cdot g_{s_n}(\omega) d\tau(\omega) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^k \sum_{j=1}^n \alpha_i \cdot \beta_j \cdot \tau(A_i). \end{aligned} \tag{3.32}$$

Sea  $\beta_i = \frac{\alpha_i \gamma_i}{\tau(A_i)}$ , con  $\gamma_1, \dots, \gamma_n$ ,  $0 \leq \gamma_i \leq 1$ ,  $1 \leq n \leq \infty$ ,  $n \in \mathbb{N}$ . Entonces resulta que

$$\begin{aligned} \int_{\Omega} f(\omega) \cdot g(\omega) d\tau(\omega) &= \lim_{n \rightarrow \infty} \sum_{i=1}^k \sum_{j=1}^n \alpha_i \cdot \beta_j \cdot \tau(A_i) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^k \sum_{j=1}^n \alpha_i \cdot \frac{\alpha_i \gamma_i}{\tau(A_i)} \cdot \tau(A_i) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^k \sum_{j=1}^n \gamma_i. \end{aligned} \quad (3.33)$$

Basta definir el conjunto de valores de  $\gamma$  de manera que cumpla con la condición  $0 < \sum_{j=1}^n \gamma_i < \infty$  para que resulte

$$0 < \lim_{n \rightarrow \infty} \sum_{i=1}^k \sum_{j=1}^n \gamma_i < \infty. \quad (3.34)$$

### 3.5.3.5 Definición de una medida de probabilidad sobre $\Omega \times \Theta$

Mostraremos en primer lugar que si las restricciones de un problema de optimización cumplen la condición  $0 < \tau(c(\omega)) < \infty$  entonces aseguran la existencia de una medida de probabilidad  $P$  sobre  $\Omega \times \Theta$ . Así como  $c(\omega)$  define una medida  $\tau$  sobre el espacio de medida  $(\Omega, \mathcal{F}, \tau)$ , la función  $c(\omega, \theta)$  definirá una medida sobre el espacio  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_{\Theta}, \tau \times P_{\Theta})$  que designaremos como  $(\tau \times P_{\Theta})_c$ , y resultará  $\frac{d(\tau \times P_{\Theta})_c}{d(\tau \times P_{\Theta})} = c(\omega, \theta)$ . Resulta entonces

$$\begin{aligned} P(A) &= \int_{\Omega \times \Theta} I_A(\omega, \theta) dP \\ &= \int_{\Omega \times \Theta} I_A(\omega, \theta) \cdot \frac{dP}{d(\tau \times P_{\Theta})} \cdot \frac{d(\tau \times P_{\Theta})_c}{d(\tau \times P_{\Theta})} d(\tau \times P_{\Theta}) \\ &= \int_{\Omega \times \Theta} I_A(\omega, \theta) \cdot \frac{dP}{d(\tau \times P)} \cdot c(\omega, \theta) d(\tau \times P_{\Theta}). \end{aligned} \quad (3.35)$$

La medida de probabilidad  $P$  debe ser uniforme respecto de  $(\tau \times P_{\Theta})_c$ , por lo que  $\frac{d(dP)}{d(\tau \times P_{\Theta})_c}$  debe ser constante. Reemplazando  $A = \Omega \times \Theta$  se obtiene

$$\begin{aligned} \frac{dP}{d(\tau \times P_{\Theta})_c} &= \frac{1}{\int_{\Omega \times \Theta} c(\omega, \theta) d(\tau \times P_{\Theta})} \\ &= \frac{1}{\int_{\Omega} \int_{\Theta} c(\omega, \theta) dP_{\Theta}(\theta) d\tau(\omega)} \\ &= \frac{1}{\int_{\Omega} c(\omega) d\tau(\omega)}, \end{aligned} \quad (3.36)$$

y por lo tanto  $\frac{dP}{d(\tau \times P_{\Theta})_c} = v$ .

3.5.3.6 Existencia de kernel de probabilidad de  $(\Theta, F_\Theta)$  a  $(\Omega, F)$

Si bien hemos verificado que hasta este punto la existencia de una medida  $P$  sobre  $\Omega \times \Theta$  dada la condición de restricciones de medida finita, para algunas aplicaciones será necesario que puedan definirse la distribución condicional sobre  $\Omega$  dado  $\Theta = \theta$ . En términos más precisos, la distribución condicional corresponde a una variable aleatoria  $\Omega \times \Theta \rightarrow \Omega$  definida por una función  $\Omega(\omega, \theta) = \omega$  respecto de una variable aleatoria  $\Omega \times \Theta \rightarrow \Theta$  definida por una función  $\Theta(\omega, \theta) = \theta$ <sup>11</sup>. La definición clásica de probabilidad condicional es

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0, \tag{3.37}$$

que aplicada a este caso corresponde a  $A \in F$  y  $B \in F_\Theta$ . La distribución de probabilidad condicional se indica como  $P(\cdot|B)$  y resulta una medida de probabilidad, dado que  $P(\Omega|B) = 1$ . Esta definición verifica  $P(B|B) = 1$  y  $P(\Theta - B|B) = 0$ , es decir que la medida de probabilidad condicional  $P(\cdot|B)$  está concentrada en  $B \in F_\Theta$ . Tradicionalmente la distribución de probabilidad condicional de la variable- $\Omega$  respecto de la variable- $\Theta$  se expresa como  $P(\Omega = \omega|\Theta = \theta)$ .

Si el modelo nulo define una medida de probabilidad  $P$  sobre el espacio medible producto  $(\Omega \times \Theta, F \times F_\Theta)$ , quedará definida una medida de probabilidad marginal  $P_\Omega$ , que no es otra cosa que la distribución de probabilidad de la variable- $\Omega$  definida sobre dicho espacio producto:

$$P_\Omega(A) = P(A \times \Theta), \tag{3.38}$$

con  $A \in F$  (Pollard 2002, p. 84). De acuerdo a lo ya visto resultará

$$P_\Omega(A) \equiv \int_\Omega I_A(\omega) \cdot \nu \cdot c(\omega) d\tau(\omega), \tag{3.39}$$

con  $A \in F$ .

Ahora bien, dado que en general  $\theta_1 \neq \theta_2 \Rightarrow c(\omega, \theta_1) \neq c(\omega, \theta_2)$  la medida de probabilidad sobre  $\Omega \times \Theta$  no se podrá definir como la medida producto:  $P_\Omega \times P_\Theta$ , ya que la variable- $\Omega$  y la variable- $\Theta$  no son variables independientes (Pollard 2002, 84). En otros términos, la distribución de probabilidad conjunta de las variables  $\Omega$  y  $\Theta$  no resulta el producto de su distribución marginal de probabilidad.

Para especificar una medida de probabilidad sobre  $\Omega \times \Theta$  será necesario entonces definir una familia de medidas de probabilidad  $\{P_\theta : \theta \in \Theta\}$  sobre  $F$  que resulten distribuciones condicionales o kernel de probabilidad de  $(\Theta, F_\Theta)$  a  $(\Omega, F)$ ; es decir,

<sup>11</sup> A fin de evitar la introducción de nuevos símbolos, nos referiremos a las variables así definidas como variable- $\Omega$  y variable- $\Theta$ . Nótese que, tal como se definieron, estas variables resultan  $F \times F_\Theta$ ,  $F$ -medible y  $F \times F_\Theta$ ,  $F_\Theta$ -medible respectivamente.

que para todo  $\theta$  resulte  $P_\theta(\Omega) = 1$  y toda función  $\theta \rightarrow P_\theta(A)$  sea  $F_\Theta$ -medible para  $A \in \mathcal{F}$ . Designaremos la restricción de la función  $c(\omega, \theta)$  a cada  $\theta$  mediante la notación

$$c_\theta(\omega) \equiv c(\omega, \theta)|_\theta. \quad (3.40)$$

Esta notación se utilizará de aquí en adelante para indicar la restricción de una función definida sobre el producto cartesiano de un número de conjuntos a alguno de los elementos de dichos conjuntos.

Supongamos que una  $c_\theta(\omega)$  permite definir una medida de probabilidad sobre  $\Omega$  en base al espacio de medida  $(\Omega, \mathcal{F}, \tau_{c_\theta})$ , donde  $\tau_{c_\theta}$  es la medida generada por  $c_\theta(\omega)$ . La medida de probabilidad  $P_\theta$  de un  $A \in \mathcal{F}$  resultará

$$P_\theta(A) = \int_{\Omega} I_A(\omega) dP_\theta. \quad (3.41)$$

De acuerdo al teorema de Radon-Nikodym resulta

$$\int_{\Omega} I_A(\omega) dP_\theta = \int_{\Omega} I_A(\omega) \cdot \frac{dP_\theta(\omega)}{d\tau_{c_\theta}(\omega)} \cdot \frac{d\tau_{c_\theta}(\omega)}{d\tau(\omega)} d\tau(\omega). \quad (3.42)$$

De igual forma que lo visto anteriormente para  $c(\omega)$ , resulta por definición  $c_\theta(\omega) = \frac{\tau_{c_\theta}}{\tau}$ . Asimismo, la densidad de Radon-Nikodym de la medida de probabilidad  $P_\theta$  respecto de  $\tau_{c_\theta}$  debe ser constante, y la designaremos como  $v_\theta \equiv \frac{dP_\theta}{d\tau_{c_\theta}}$ . Resulta entonces

$$\int_{\Omega} I_A(\omega) dP_\theta(\omega) = \int_{\Omega} I_A(\omega) \cdot v_\theta \cdot c_\theta(\omega) d\tau(\omega). \quad (3.43)$$

Dado que  $P_\theta(\Omega) = 1$  por definición, resulta

$$v_\theta = \frac{1}{\int_{\Omega} c_\theta(\omega) d\tau(\omega)}, \quad (3.44)$$

lo que da lugar a la definición de la función  $v(\theta) \equiv v_\theta$ .

En esta punto cabe señalar que la existencia de una distribución de probabilidad condicional dado un espacio de probabilidad arbitrario no es en absoluto trivial (Pollard 2002, pp. 112 y 339-343). Esto significa que, aún cuando esté probado que exista una medida  $P$  sobre  $\Omega \times \Theta$ , puede no ser posible definir un kernel de probabilidad  $\{P_\theta : \theta \in \Theta\}$ . Sin embargo, en el caso de que  $(\Omega, \mathcal{F})$  sea un espacio medible estándar la existencia de una distribución de probabilidad condicional está asegurada (Çınlar 2011, p. 152-156). Como ya hemos señalado, la clase de espacios medibles estándar es tan amplia que virtualmente la totalidad de modelos de interés biológicos han si-

<sup>12</sup>  $c_\theta(\omega)$  es medible respecto de  $\tau$  dado que  $c(\omega, \theta)$ , ya que la restricción de una función medible resulta medible.

do planteados en espacios de esta naturaleza, por lo que el presente análisis es de aplicación general.

Analizaremos de forma explícita, y a modo ilustrativo, el caso en que  $\Theta$  sea un espacio contable. Si  $\Theta$  se trata de un espacio veremos que  $\tau(c(\omega)) < \infty \Rightarrow v(\theta) < \infty$ , de manera que para todo  $\theta$  para el que resulte  $v(\theta) > 0$  está asegurada la existencia de la medida de probabilidad condicional  $P_\theta$ . De acuerdo a lo expuesto anteriormente, y en virtud del Teorema de Tonelli resulta

$$\begin{aligned} \int_{\Theta} \int_{\Omega} c(\omega, \theta) d\tau(\omega) dP_{\Theta}(\theta) &= \sum_{\theta \in \Theta} \int_{\Omega} c(\omega, \theta) d\tau(\omega) P_{\Theta}(\theta) \\ &= \sum_{\theta \in \Theta} \int_{\Omega} c_{\theta}(\omega) d\tau(\omega) P_{\Theta}(\theta), \end{aligned} \quad (3.45)$$

y por lo tanto

$$\begin{aligned} \infty &> \sum_{\theta \in \Theta} \int_{\Omega} c_{\theta}(\omega) d\tau(\omega) P_{\Theta}(\theta) \\ \infty &> \int_{\Omega} c_{\theta}(\omega) d\tau(\omega), \quad \theta \in \Theta. \end{aligned} \quad (3.46)$$

En consecuencia, si resulta  $c_{\theta}(\omega) = 1$  al menos un  $\omega$  para todo  $\theta$ ,  $P_{\theta}$  estará definido para todo  $\theta$ .<sup>13</sup>

La existencia de la distribución de probabilidad condicional en el caso de un espacio contable se deriva de forma general de la definición 3.37, dado que para cualquier  $B \in \mathcal{F}_{\Theta}$  que no sea  $P_{\Theta}$ -nulo resulta  $P(B) > 0$ . En el caso de que  $\Theta$  no sea un espacio contable, resultará  $P_{\Theta}(\theta) = 0$  para todo  $\theta \in \Theta$ . En consecuencia la definición 3.37, de forma más general, debe interpretarse en términos de la derivada de Radon-Nikodym o densidad de probabilidad.

### 3.6 FORMULACIÓN DE UN MODELO GENERAL PROBABILÍSTICO DE OPTIMIZACIÓN

En la formulación del modelo nulo probabilístico a partir de un problema de optimización ha sido omitida la función objetivo ya que, por definición, ésta debe estar ausente de la hipótesis nula asociada a dicho modelo. A continuación, mostraremos cómo incorporar la función objetivo al modelo nulo desarrollado, y formular así un modelo de optimización generalizado en la forma de un espacio de probabilidad.

<sup>13</sup> De existir un subconjunto no nulo  $N = \{\theta : c_{\theta}(\omega) = 0, \omega \in \Omega\}$  la condición establecida se cumplirá para todo  $\theta \in (\Theta - N)$ , y la medida  $P$  se concentrará en  $(\Theta - N) \times \Omega$ .

### 3.6.1 Construcción de un modelo general de optimización

En esta sección desarrollaremos la formulación de un modelo de optimización generalizado en la forma de un espacio de medida que incluye como casos particulares al modelo nulo y a la optimización estricta. Esta formulación se logra incorporando la función objetivo a lo desarrollado previamente para el caso de un modelo nulo, e introduciendo los conceptos de función de normalización y función de transferencia. En primer lugar abordamos el caso particular de una optimización estricta con restricciones deterministas, para luego relajar el supuesto de optimización estricta y abordar el caso más general de restricciones probabilísticas. A continuación definimos las funciones de normalización y de transferencia, y finalmente presentamos una formulación general de un problema de optimización arbitrario como un espacio de medida.

#### 3.6.1.1 Generalización del modelo de optimización mediante relajación de supuestos

De acuerdo a lo expuesto en la Sección 3.3, dado un espacio muestral  $\Omega$  y un conjunto de restricciones, la función objetivo de un problema de optimización define el conjunto de soluciones óptimas. Consideremos en principio el caso de restricciones de tipo deterministas en la forma  $c : \Omega \rightarrow \{0, 1\}$  que definen un espacio de soluciones candidatas como un conjunto  $\Omega_c \subseteq \Omega$ .

En un problema de optimización típico se asume una optimización estricta, y en ese caso se ha visto que el conjunto de soluciones óptimas  $\Omega_o \subseteq \Omega$  se define como <sup>14</sup>

$$\Omega_o = \{\omega : \omega \in \Omega_c, f(\omega) = \sup_{\omega \in \Omega_c} f(\omega)\}. \quad (3.47)$$

La función objetivo permite definir la función indicadora de  $\Omega_o$ ,  $I_{\Omega_o} : \Omega \rightarrow \{0, 1\}$ , como:

$$I_{\Omega_o}(\omega) = \begin{cases} 1 & \text{si } \omega \in \Omega_c \text{ y } f(\omega) = \sup_{\omega \in \Omega_c} f(\omega) \\ 0 & \text{si } \omega \notin \Omega_c \text{ y / o } f(\omega) < \sup_{\omega \in \Omega_c} f(\omega) . \end{cases} \quad (3.48)$$

De forma equivalente, es posible definir  $I_{\Omega_o}$  a partir de la introducción de las funciones que permitirán generalizar el problema de optimización. Introducimos con este fin una función  $n$  que denominaremos función de normalización. Más adelante presentamos una definición general de la función de normalización (Sección 3.6.1.2); en el caso

<sup>14</sup> A diferencia de lo utilizado tradicionalmente en la literatura de optimización, en lugar de elemento máximo utilizamos el elemento supremo para definir a las soluciones del problema de optimización; el motivo de esta discrepancia se explica más adelante.

particular de un problema determinista con optimización estricta  $n$  puede expresarse como:

$$n(\omega) = c(\omega) \cdot \frac{f(\omega) - \inf_{\omega \in C} f(\omega)}{\sup_{\omega \in C} f(\omega) - \inf_{\omega \in C} f(\omega)}, \tag{3.49}$$

donde  $C = \{\omega : c(\omega) = 1\}$ . Esta función tomará valores en el rango  $[0, 1]$ <sup>15</sup>, y resultará  $n(\omega) = 1$  para un estado óptimo.

Introduciremos ahora una función  $t$ , que denominaremos función de transferencia, cuya definición general también presentamos más adelante (Sección 3.6.1.3); aplicada a la función  $n(\omega)$ , la función de transferencia dará como resultado el conjunto de soluciones óptimas:

$$I_{\Omega_0}(\omega) = c(\omega) \cdot t \ n(\omega) . \tag{3.50}$$

En el caso de una optimización estricta y restricciones deterministas, la función de transferencia  $t : [0, 1] \rightarrow \{0, 1\}$  resulta una función definida de a tramos:

$$t(\varphi) = \begin{cases} 1 & \text{si } \varphi = 1 \\ 0 & \text{si } \varphi < 1, \end{cases} \tag{3.51}$$

con  $\varphi \in [0, 1]$ .

Considérese ahora la formulación del problema de optimización presentado en la Sección 3.3.0.2, que permite relajar los supuestos asociados de optimización y monotonía estricta. De acuerdo a esta formulación, dada una función objetivo y un cierto valor del parámetro  $q$  queda determinado si un estado  $\omega$  es o no una solución del problema de optimización, lo que equivale a asignar a cada par  $(\omega, q)$  un valor de 0 o 1. En este caso, la función de transferencia resulta una función escalonada con un parámetro de posición  $q$ :

$$t(\varphi, q) = \begin{cases} 1 & \text{si } \varphi \geq q, \\ 0 & \text{si } \varphi < q. \end{cases} \tag{3.52}$$

En el caso de que las restricciones sean deterministas, la definición de la función indicadora del subconjunto de soluciones óptimas resulta similar a la Ecuación 3.55, con el agregado que depende del parámetro  $q$  de la función de transferencia

$$I_{\Omega_0}(\omega, q) = c(\omega) \cdot t \ n(\omega), q , \tag{3.53}$$

---

<sup>15</sup> El término  $\frac{f(\omega) - \inf_{\omega \in C} f(\omega)}{\sup_{\omega \in C} f(\omega) - \inf_{\omega \in C} f(\omega)}$  será negativo si  $f(\omega) < \inf_{\omega \in C} f(\omega)$ , y será mayor a 1 si  $f(\omega) > \sup_{\omega \in C} f(\omega)$ , pero esto solo sucederá si  $c(\omega) = 0$ , con lo que la multiplicación por dicho término hace que  $n(\omega)$  tenga a  $[0, 1]$  como codominio.

por lo que  $I_{\Omega_0} : \Omega \times [0, 1] \rightarrow \{0, 1\}$ .

Considérese ahora el caso general de un conjunto de restricciones probabilísticas que definen una función  $c : \Omega \times \Theta \rightarrow \{0, 1\}$ . Dado que a cada  $\theta$  corresponde un problema de optimización determinista, la función objetivo dependerá en general de  $\theta$ . En este caso, expresaremos la función de normalización  $n : \Omega \times \Theta \rightarrow \mathbb{R}^+$  como:

$$n(\omega, \theta) = c_{\theta}(\omega) \cdot \frac{f_{\theta}(\omega) - \inf_{\omega \in C_{\theta}} f_{\theta}(\omega)}{\sup_{\omega \in C_{\theta}} f_{\theta}(\omega) - \inf_{\omega \in C_{\theta}} f_{\theta}(\omega)}, \quad (3.54)$$

donde  $C_{\theta} \equiv \{\omega : c_{\theta}(\omega) = 1\}$ . El subconjunto de soluciones óptimas queda definido como

$$I_{\Omega_0}(\omega, \theta, q) = c(\omega, \theta) \cdot t \ n(\omega, \theta), q, \quad (3.55)$$

por lo que  $I_{\Omega_0} : \Omega \times \Theta \times [0, 1] \rightarrow \{0, 1\}$ .

Es importante destacar que el ínfimo y el supremo de  $f(\omega, \theta)$  no están evaluados en todo el dominio  $\Omega \times \Theta$ : dado que cada  $\theta$  define un problema de optimización determinista, la notación  $\sup_{\omega \in C_{\theta}}$  indica su evaluación sobre un subconjunto formado por todos los pares  $(\omega, \theta)$  para un  $\theta$  fijo en los que  $c(\omega, \theta) = 1$ <sup>16</sup>.

### 3.6.1.2 Definición de función de normalización

Una función de normalización representa el criterio con el que se normalizan los valores de la función objetivo para distintos valores de  $\theta$ . Dado que para cada  $\theta \in \Theta$  queda definido un problema de optimización determinista y asociado a este una función objetivo  $f_{\theta}(\omega)$ , mediante la función de normalización se expresa cómo se ponderarán los diferentes valores de  $f_{\theta}(\omega)$  para determinar el grado de pertenencia de un cierto  $\omega$  al conjunto de soluciones óptimas.

De forma general, dado un espacio medible  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_{\Theta})$ , definimos la función de normalización como una función  $\mathcal{F} \times \mathcal{F}_{\Theta}, \mathcal{B}([0, 1])$ -medible  $n : \Omega \times \Theta \rightarrow [0, 1]$  que cumple con los siguientes requerimientos:

1. La restricción de la función de normalización a cada  $\theta \in \Theta$ ,  $n_{\theta} : \Omega \rightarrow [0, 1]$ , resulta una función monótonamente no decreciente respecto de  $c_{\theta}(\omega) \cdot f_{\theta}(\omega)$ :

$$c_{\theta}(\omega_1) \cdot f_{\theta}(\omega_1) > c_{\theta}(\omega_2) \cdot f_{\theta}(\omega_2) \iff n(\omega_1, \theta) \geq n(\omega_2, \theta). \quad (3.56)$$

<sup>16</sup> Dada la naturaleza del espacio de medida producto, a cada  $\theta$  se asocia una cierta restricción  $c_{\theta}(\omega) \equiv c(\omega, \theta)$  que resulta una función  $c_{\theta} : \Omega \rightarrow \{0, 1\}$ . De la misma forma, para un  $\theta$  dado queda definida la función  $f_{\theta}(\omega) \equiv f(\omega, \theta)$  que resulta una función  $f_{\theta} : \Omega \rightarrow [0, \infty)$ . Por lo tanto, fijado un valor de  $\theta$  queda definida una cierta  $c_{\theta}(\omega)$ , y con esta un subespacio de soluciones candidatas  $\Omega_{c_{\theta}} \equiv \{\omega : \omega \in \Omega, c_{\theta}(\omega) = 1\}$  y, a su vez, un subconjunto de soluciones óptimas  $\Omega_{o_{\theta}} \subseteq \Omega_{c_{\theta}}$ . Dado que  $\sup_{\omega \in \Omega} c(\omega, \theta) \cdot f(\omega, \theta) = \sup_{\omega \in \Omega_{c_{\theta}}} f(\omega, \theta)$ , consideramos que la primera expresión resulta más accesible que la segunda.

2. Para todo  $\omega$  tal que  $c_\theta(\omega) = 1$  debe verificarse que

$$\begin{aligned} f_\theta(\omega) = \inf_{\omega \in C_\theta} f_\theta(\omega) &\implies n_\theta(\omega) = 0, \\ f_\theta(\omega) = \sup_{\omega \in C_\theta} f_\theta(\omega) &\implies n_\theta(\omega) = 1. \end{aligned} \tag{3.57}$$

De acuerdo a la definición dada, cualquier función  $n$  coincidirá en el caso de la optimización estricta, ya que  $n_\theta(\omega) = 1$  para los óptimos globales para cada  $\theta$ . En contraste, diferentes funciones  $n$  plantean diferentes maneras de relajar el supuesto de optimización estricta. Trataremos a continuación dos tipos de funciones de acuerdo al criterio utilizado para realizar la normalización.

**NORMALIZACIÓN POR VALOR SUPREMO E ÍNFI MO** Esta es la formulación con la que se ha ejemplificado la función de normalización hasta ahora:

$$n(\omega, \theta) = c_\theta(\omega) \cdot \frac{f_\theta(\omega) - \inf_{\omega \in C_\theta} f_\theta(\omega)}{\sup_{\omega \in C_\theta} f_\theta(\omega) - \inf_{\omega \in C_\theta} f_\theta(\omega)}, \tag{3.58}$$

con  $C_\theta \equiv \{\omega : c_\theta(\omega) = 1\}$ <sup>17</sup>.

Las funciones  $c$  y  $f$  son, por definición, funciones medibles en  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\theta)$ . La suma, producto y cociente de funciones medibles resultan medibles, así como el ínfimo y el supremo (Folland 2013, p. 45). En consecuencia, la formulación de la función de normalización formulada presentada en la Ecuación 3.58 cumple el requisito de mensurabilidad; se verifican asimismo el resto de los requisitos antes establecidos en la definición general.

Se ha utilizado el elemento supremo en lugar del máximo que se utiliza típicamente en la formulación de modelos optimización a causa de su mayor generalidad de aplicación. En caso de existir, el máximo de una función coincide con su supremo:

$$\exists \max_{\omega \in \Omega} f(\omega) \implies \max_{\omega \in \Omega} f(\omega) = \sup_{\omega \in \Omega} f(\omega). \tag{3.59}$$

Sin embargo, puede darse que el supremo exista pero la función no tome el valor supremo, y en tal caso no existirá un elemento máximo:

$$\arg \sup_{\omega \in \Omega} f[\Omega] \notin \Omega \implies \nexists \max_{\omega \in \Omega} f(\omega). \tag{3.60}$$

<sup>17</sup> Hemos supuesto que la optimización corresponde a una maximización de la función objetivo, pero de tratarse una minimización basta con definir una nueva función objetivo invirtiendo el signo de la función original.

Esto ocurre cuando  $\Omega$  se trata de un espacio métrico y el supremo de  $f(\omega)$  coincide con la frontera de un conjunto abierto definido por una restricción de desigualdad. Por ejemplo, con  $\Omega_c = [0, 1)$  y  $f(\omega) = \omega$  resulta

$$\arg \sup_{\omega \in \Omega} f(\omega) = 1 / [0, 1), \quad (3.61)$$

y, por lo tanto, el máximo de la función objetivo no existe. Si se asume una optimización estricta no existirá una diferencia entre utilizar el elemento supremo o el máximo en la formulación de un modelo de optimización, ya que un  $\omega$  dado será una solución óptima solo si se verifica  $f(\omega) = \sup_{\omega \in \Omega} f(\omega)$  y esto solo sucede si el elemento máximo existe. En cambio, si se relaja el supuesto de optimización estricta, el elemento supremo permite definir un modelo de optimización aún si  $f(\omega) \neq \sup_{\omega \in \Omega} f(\omega)$ , con la salvedad que dicho modelo no podrá incluir la optimización estricta como caso particular; será posible, sin embargo, aproximarse a la hipótesis de optimización estricta tanto como se desee.

**NORMALIZACIÓN CUANTIL EN BASE AL MODELO NULO** Como alternativa, es posible realizar una normalización de acuerdo a la distribución de probabilidad de la función objetivo para cada  $\theta$  que se espera en base al modelo nulo. Dicha distribución es la que se predice de acuerdo a la hipótesis nula y, por lo tanto, suele designarse como distribución nula. Se verá más adelante que esta normalización presenta una serie de ventajas en relación al planteo e interpretación de hipótesis, así como también en cuanto a la dificultad de obtener una aproximación numérica, lo que será típicamente necesario para realizar una inferencia estadística.

Una vez que a partir de un problema de optimización se ha definido el modelo nulo como un espacio de probabilidad  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\theta, P)$ , tal como se vio en la Sección 3.5.3.5, está asegurada la existencia de una familia de medidas de distribución de probabilidad condicional  $P_\theta(\omega)$  para  $\theta \in \Theta$ . La función objetivo  $f(\omega, \theta)$  definirá una familia de funciones  $\mathcal{F}$ -medibles  $f_\theta(\omega)$  que resultarán variables aleatorias sobre las respectivas distribuciones condicionales  $P_\theta(\omega)$ , e inducirán por lo tanto medidas de probabilidad  $P_{f_\theta}$  sobre  $\mathbf{B}(\mathbf{R})$ <sup>18</sup>:

$$P_{f_\theta}(A) \equiv P_\theta \circ f_\theta^{-1}(A), \quad A \in \mathbf{B}(\mathbf{R}). \quad (3.62)$$

Dado que  $f_\theta(\omega)$  es una función sobre los reales, la función de distribución acumulada resulta

$$F_{f_\theta}(x) \equiv P_{f_\theta}(-\infty, x], \quad x \in \mathbf{R}. \quad (3.63)$$

<sup>18</sup> A pesar de que es convención utilizar una letra mayúscula para indicar una variable aleatoria, nos referiremos a la variable aleatoria definida por una función  $f_\theta(\omega)$  como  $f_\theta$  ya que el símbolo  $F$  se utiliza para la función de distribución acumulada.

Definimos la normalización cuantil como

$$n(\omega, \theta) = c_\theta(\omega) \cdot F_{f_\theta, q_N} f_\theta(\omega) , \tag{3.64}$$

donde  $F_{f_\theta, q_N}$  designa la distribución acumulada de la función objetivo para  $\Theta = \theta$  de acuerdo a la distribución correspondiente al modelo nulo,  $P_{f_\theta, q_N}$  ; el significado de  $q_N$  se verá en la siguiente sección, donde se da una definición general de la función de transferencia. El valor de esta función de normalización para un par  $(\omega, \theta)$  corresponde al cuantil de la función objetivo  $f(\omega, \theta)$  de acuerdo al modelo nulo y condicional a  $\Theta = \theta$ .

La función definida cumple los requisitos establecidos para una función de normalización. Para todo  $\theta \in \Theta$  resulta por definición que  $F_{f_\theta, q_N}$  es creciente respecto de  $c_\theta(\omega) \cdot f_\theta(\omega)$ , y resulta

$$\begin{aligned} F_{f_\theta, q_N} \inf_{\omega \in C_\theta} f_\theta(\omega) &= 0 \\ F_{f_\theta, q_N} \sup_{\omega \in C_\theta} f_\theta(\omega) &= 1 . \end{aligned} \tag{3.65}$$

La función de normalización cuantil cumple asimismo el requisito de mensurabilidad, ya que la función de distribución acumulada es  $B(\mathbb{R})$ -medible.

### 3.6.1.3 Definición de función de transferencia

Se ha utilizado hasta aquí una función escalonada como la función de transferencia  $t$  que, a partir de una función de normalización  $n(\omega, \theta)$ , define el conjunto de soluciones óptimas. La función escalonada resulta un ejemplo conveniente de función de transferencia porque posee un único un parámetro de posición,  $q \in [0, 1]$ , pero se verá a continuación que representa solo un caso particular de una clase de funciones que definimos a continuación.

De forma general, dado un espacio medible  $(Q, F_Q)$  definimos una función de transferencia como una función  $B([0, 1]) \times F_Q, B([0, 1])$  -medible  $t: [0, 1] \times Q \rightarrow [0, 1]$  que cumple con los siguientes requerimientos:

**EL MODELO NULO RESULTA UN CASO PARTICULAR** Existe un valor  $q_N \in Q$  tal que

$$t n(\omega, \theta), q_N = 1 \tag{3.66}$$

para todo  $\omega \in \Omega$  y  $\theta \in \Theta$ , de manera que se verifique

$$I_{\Omega_0}(\omega, \theta, q_N) = c(\omega, \theta) , \tag{3.67}$$

y el modelo nulo resulte entonces un caso particular del modelo de optimización.

ES no DECRECIENTE PARA RANGO INFERIOR DEL DOMINIO Para todo  $q \in Q$  existe un rango inferior del codominio  $[0, u]$  de la función de normalización  $n$ , con  $0 < u \leq 1$ , donde se verifica una relación monótonamente no decreciente entre  $t$  y  $n$ :

$$n_1 < n_2 \leq u \implies t(n_1, q) \leq t(n_2, q). \quad (3.68)$$

Dependiendo de los supuestos asociados a las hipótesis de optimización, puede existir un tercer requerimiento. Sea  $f_v(\omega, \theta)$  la función objetivo que se hipotetiza verdaderamente sujeta a optimización, y  $f(\omega, \theta)$  la efectivamente utilizada en el modelo de optimización:

VALOR MÁXIMO PARA RANGO SUPERIOR DOMINIO Si resulta  $f \equiv f_v$  para todo  $q \in Q$ , o si  $f \neq f_v$  pero se supone una monotonía no estricta, debe resultar  $t(\varphi, q) = 1$  para  $\varphi \in [s, 1]$ . Esta condición, sumada al requerimiento de que  $T$  sea no decreciente para rango inferior del dominio, implica además que existirá una relación monótonamente no decreciente entre  $t(n(\omega, \theta), q)$  y  $n(\omega, \theta)$  para todo  $q$  y todo el rango de  $n$ .

Estos requisitos representan un conjunto de características que consideramos esenciales para que la formulación general pueda representar efectivamente un modelo de optimización. En el caso de que resulte  $f \neq f_v$  y no se asuma una monotonía entre ambas, para  $n > u$  no existirá otro requerimiento en la relación entre  $t$  y  $n$ . En tal caso, una función  $t$  que cumpla con los tres requerimientos será adecuada pero, sin embargo, no abarcará todos los modelos posibles ya que será incapaz de representar el caso hipotético de una relación decreciente entre  $t$  y  $n$  para  $n(\omega, \theta) > u$ .

En el caso particular de la función escalonada como función de transferencia resulta  $Q = [0, 1]$ , y se verifica que cumple los requisitos establecidos: el modelo nulo se obtiene como caso particular cuando  $q_N = 0$ ;  $t$  resulta una función no decreciente para todo  $q \in Q$ ; y  $t(\varphi, q) = 1$  para  $\varphi \geq q$ .

#### 3.6.1.4 Formulación general de un modelo de optimización

Sean  $(\Omega, F)$ ,  $(\Theta, F_\Theta)$  y  $(Q, F_Q)$  espacios medibles, dado un modelo de optimización definido por una restricción  $c(\omega, \theta)$  y una función objetivo  $f(\omega, \theta)$ , y definidas una función de normalización  $n(\omega, \theta)$  y una función de transferencia  $t: [0, 1] \times Q \rightarrow [0, 1]$ , queda definido un modelo general de optimización en la forma de una función  $o: \Omega \times \Theta \times Q \rightarrow [0, 1]$ , que denominaremos función de optimización, de acuerdo a

$$o(\omega, \theta, q) \equiv c(\omega, \theta) \cdot t(n(\omega, \theta), q). \quad (3.69)$$

La función de optimización  $o$  representa una generalización de un modelo de optimización: en lugar de definir un conjunto de soluciones óptimas define un grado de pertenencia a dicho conjunto para cada  $\omega$  en función de los parámetros  $\theta$  y  $q$ .

Se mostrará a continuación que, tal como está definida, la función  $o$  resulta  $F \times F_{\Theta} \times F_Q, \mathcal{B}([0, 1])$ -medible, entonces  $o$  define una medida sobre el espacio de medida  $(\Omega \times \Theta \times Q, F \times F_{\Theta} \times F_Q, \tau \times P_{\Theta} \times P_Q)$ . Las funciones  $c$ ,  $t$  y  $n$  son medibles en sus respectivos espacios medibles. Dado que el producto, composición y cociente entre funciones medibles para una cierta  $\sigma$ -álgebra resulta medible, es suficiente con mostrar que dichas funciones son  $F \times F_{\Theta} \times F_Q, \mathcal{B}([0, 1])$ -medibles para demostrar que  $o$  lo es. Dado que la extensión de una función medible a un dominio más amplio resulta siempre una función medible<sup>19</sup>, la extensión de  $c(\omega, \theta)$  al dominio  $\Omega \times \Theta \times Q$  resulta  $F \times F_{\Theta} \times F_Q, \mathcal{B}([0, 1])$ -medible. La función de transferencia tiene como entrada a  $n(\omega, \theta)$  y  $q$ . Una función  $g(q) \equiv q$  puede extenderse para definir una función  $g(\omega, \theta, q)$  que resultará  $F \times F_{\Theta} \times F_Q, F_Q$ -medible. La función de normalización  $n(\omega, \theta)$  resulta por definición una función  $\tau, ([0, 1])$ -medible. En consecuencia,  $t n(\omega, \theta), q$  resulta  $F \times F_{\Theta} \times F_Q, \mathcal{B}([0, 1])$ -medible, con lo que se verifica que la función  $o$  posee la propiedad de mensurabilidad requerida.

### 3.6.2 Definición de una medida sobre el espacio muestral

Trataremos en esta sección la definición de una medida sobre el espacio muestral a partir de la formulación general de un modelo de optimización en la forma de una función  $o(\omega, \theta, q)$ .

#### 3.6.2.1 Espacio de medida sobre $\Omega \times \Theta$

En primer lugar, abordaremos como se puede definir una medida sobre  $(\Omega \times \Theta, F \times F_{\Theta})$  a partir de la restricción de la función  $o(\omega, \theta, q)$  al dominio  $\Omega \times \Theta \times q$ ; designaremos a la restricción de la función general como  $o_q(\omega, \theta)$ , o simplemente  $o_q$ , y como  $\tau^q$  a la medida que definirá. Los requerimientos que debe cumplir una medida  $\tau^q$  respecto de  $o_q$  son los mismos que se establecieron en la Sección 3.5.1 para el caso particular del modelo nulo:

1.  $\tau^q$  debe ser una medida  $\sigma$ -finita sobre  $\Omega$  para una  $\sigma$ -álgebra  $F$ .

<sup>19</sup> Dada una función  $f: X \rightarrow Y$  que resulta  $(A, B)$ -medible, su extensión a un espacio medible  $(X \times \Omega, A \times F)$  resulta una función  $g: X \times \Omega \rightarrow Y$  definida como  $g(x, \omega) = f(x)$  resulta  $(A \times F, B)$ -medible. Que  $g(x, \omega)$  sea medible significa que si  $B \in \mathcal{B}$ , entonces su imagen inversa  $C = g^{-1}[B] \equiv \{(x, \omega) : f(x, \omega) \in B\}$  debe estar incluida en  $A \times F$ . Dado que  $f$  es medible por definición,  $f^{-1}[B] \equiv \{x : f(x) \in B\} = A \in \mathcal{A}$ , y por lo tanto  $x \in A \Rightarrow x \in C$ . Dado que  $g$  es una extensión de  $f$ , el valor de  $g$  no varía con  $\omega$ , por lo que resulta que  $\omega \in \Omega \Rightarrow \omega \in C$ . En consecuencia,  $C = g^{-1}[B]$  estará formado por todos los pares  $\{(x, \omega) : x \in A, \omega \in \Omega\}$ , y esto resulta el producto cartesiano  $A \times \Omega$ , y por definición la  $\sigma$ -álgebra producto resulta entonces que  $A \times \Omega \in A \times F$ , con lo que se verifica que  $g$  es medible.

2.  $o_q(\omega, \theta)$  debe resultar una función medible en el espacio medible  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$ .
3. La medida  $\tau^q$  de  $o_q(\omega, \theta)$  debe ser no nula a excepción del caso trivial en el que  $o_q(\omega, \theta) = 0$  para todo  $\omega$  y  $\theta$ .

Es importante señalar que, a fin de simplificar el tratamiento matemático y sin pérdida de generalidad, en la Sección 3.5.2.3 se estableció el siguiente supuesto. Si la función  $c(\omega, \theta)$  define una medida de probabilidad sobre  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$ , la medida de probabilidad marginal sobre  $\Omega$  definida por  $c(\omega, \theta)$  es absolutamente continua respecto a una cierta medida  $\sigma$ -finita  $\tau$  sobre  $\Omega$ . Extenderemos este supuesto a la función  $o_q(\omega, \theta)$ , de manera que si  $o_q$  define efectivamente una medida de probabilidad  $P_q$  sobre  $\Omega \times \Theta$ , la medida de probabilidad marginal sobre  $\Omega$  resulta absolutamente continua respecto de una medida  $\sigma$ -finita  $\tau^q$ . Al igual que lo indicado para el modelo nulo, si un problema de optimización no respeta este supuesto será posible expresar o aproximar dicho problema en un conjunto contable de modelos en el que cada  $o_q$  defina una medida marginal sobre  $\Omega$  que sea absolutamente continua respecto de una cierta  $\tau^q$ . En consecuencia, si se puede definir tal medida  $\tau^q$ , la función  $o_q$  definirá una nueva medida sobre  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$  absolutamente continua respecto de la medida producto  $\tau^q \times P_\Theta$ .

En el caso de que  $\Omega$  se trate de un espacio contable, la definición de  $\tau$  es trivial, ya que será la medida de conteo para cualquier  $q$ .

Siguiendo un razonamiento similar al utilizado para el modelo nulo, en caso de que  $\Omega$  sea un espacio métrico de dimensión  $n$ , la función  $o_q$  define un conjunto  $\Omega_q = \{\omega : \int_{\Theta} o_q(\omega, \theta) dP_\Theta(\theta) > 0\}$ . La geometría de  $\Omega_q$  dependerá en última instancia de las funciones de restricción, objetivo, de normalización y de transferencia, pero dado un cierto modelo de optimización variará únicamente con el parámetro  $q$ . Dada una  $o_q$ , queda definida una medida  $\tau^q$  como

$$\tau^q(A) = H^s(A \cap \Omega_q), \quad A \in \mathcal{F}, \quad s = \dim_H(\Omega_q), \quad (3.70)$$

donde  $\mathcal{F}$  es la  $\sigma$ -álgebra completa de los conjuntos  $H^n$ -medibles de  $\Omega$ . En la Sección 3.5.2.3 se estableció para el caso particular del modelo nulo ( $q = q_N$ ) que la medida  $\tau^q$  definida es  $\sigma$ -finita; esto es asimismo válido para cualquier  $q \in \mathcal{Q}$ .

Dado que la función de optimización  $o(\omega, \theta, q)$  resulta una función medible, por el teorema de Tonelli su restricción  $o_q(\omega, \theta)$  es una función medible sobre el espacio  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$ . De acuerdo al teorema de la convergencia monótona, una función medible no negativa define una nueva medida sobre un espacio de medida  $\sigma$ -finito.

Por lo tanto, a partir de una medida producto  $\tau^q \times P_\Theta$  sobre dicho espacio medible,  $o_q$  define una medida

$$\tau^q \times P_\Theta \circ_{o_q}(A) \equiv \int_{\Omega \times \Theta} I_A(\omega, \theta) \cdot o_q(\omega, \theta) d \tau^q \times P_\Theta, \tag{3.71}$$

y la función  $o_q$  resulta la derivada de Radon-Nikodym (o densidad) de la nueva medida respecto de la medida base:  $o_q(\omega, \theta) = \frac{d(\tau^q \times P_\Theta) \circ_{o_q}}{d(\tau^q \times P_\Theta)}$ . De esta manera, para un cierto  $q$  el modelo de optimización queda expresado en la forma de un espacio de medida  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta, (\tau^q \times P_\Theta) \circ_{o_q})$ .

Por último, este resultado se extiende al caso general en que el espacio muestral resulte un espacio producto de un número contable de espacios contables y/o espacios métricos no contables, tal como se vio en la Sección 3.5.1; en otros términos, es válido para el caso general de espacios medibles estándar (Sección 3.5.1).

### 3.6.2.2 Espacio de medida sobre $\Omega \times \Theta \times Q$

Consideremos ahora la expresión del modelo general como una medida definida por la función  $o(\omega, \theta, q)$  sobre el espacio medible producto  $(\Omega \times \Theta \times Q, \mathcal{F} \times \mathcal{F}_\Theta \times \mathcal{F}_Q)$ . Dado que  $o(\omega, \theta, q)$  es medible en dicho espacio por definición, de acuerdo al teorema de Radon-Nikodym la función definirá una nueva medida sobre cualquier medida  $\sigma$ -finita en dicho espacio medible.

Hasta ahora hemos asumido que se parte de un espacio de medida con una cierta medida base  $\tau$  sobre  $\Omega$ ; de esta manera, hemos evitamos lidiar con medidas con componentes singulares no nulos gracias a un supuesto que, como se ha visto, no implica una pérdida de generalidad. Sin embargo, hemos visto en la sección anterior que la definición de  $\tau^q$  podría variar al en función de  $q$  en caso de ser  $\Omega$  un espacio métrico no contable (Sección 3.6.2.2). Por lo tanto, para dar un tratamiento unificado al modelo general de optimización debemos considerar de forma explícita la posibilidad de que al variar  $q$  existan medidas  $\tau^q$  con un componente singular no nulo respecto a la medida definida por el modelo nulo  $\tau^{qN}$ .

Tal como fue definida (Ecuación 3.70),  $\tau^q$  es  $\sigma$ -finita sobre  $(\Omega, \mathcal{F})$  para todo  $q \in Q$ . De acuerdo al teorema de descomposición de Lebesgue (Lawrence y Gariepy 2015, p. 52),  $\tau^q$  puede descomponerse respecto de la medida correspondiente al modelo nulo  $\tau^{qN}$  como

$$\begin{aligned} \tau^q(A) &= \tau_s^q(A) + \tau_a^q(A) \\ &= \int_{\Omega} I_A(\omega) \cdot \frac{d\tau^q}{d\tau^{qN}} d\tau^{qN}(\omega) + \tau_s^q(A) \end{aligned} \tag{3.72}$$

para todo  $A \in \mathcal{F}$ .  $\tau_s^q(A)$  es el componente de  $\tau^q$  que resulta absolutamente continuo respecto de  $\tau^{qN}$  (es decir,  $\tau_a^q(A) = 0$ ), y  $\frac{d\tau^q}{d\tau^{qN}}$  es la derivada de Radon-Nikodym

de  $\tau^q$  respecto de  $\tau^{qN}$ .  $\tau^q$ , por su parte, es el componente de  $\tau^q$  mutuamente singular respecto de  $\tau^{qN}$ , lo que significa que existe un conjunto  $E \subseteq \Omega$  tal que  $\tau^q(\Omega - E) = \tau^{qN}(E) = 0$ .

La descomposición de la medida  $\tau^q$  para cada  $q$  respecto de  $\tau^{qN}$  dependerá que cómo varíe la geometría del conjunto  $\Omega_q$  en función  $q$ . En primer lugar, nótese que por como fueron definidas las funciones de transferencia y normalización es trivial verificar que para cualquier  $q \in Q$  resulta

$$o_q(\omega, \theta) \leq o_{qN}(\omega, \theta) = c(\omega, \theta), \quad (3.73)$$

y, por lo tanto,  $\Omega_q \subseteq \Omega_{qN} = \Omega_c$ . La dimensión de Hausdorff del espacio de soluciones candidatas  $\Omega_c$  resultará entonces igual o mayor a la dimensión del conjunto de soluciones  $\Omega_q$  (Mattila 1999, p. 59):

$$\dim_H \Omega_q \leq \dim_H \Omega_c \leq \dim_H \Omega. \quad (3.74)$$

Analizaremos en primer lugar el caso de que  $\dim_H \Omega_q = \dim_H \Omega_c$ ; veremos que si esto sucede  $\tau^q$  resultará absolutamente continua respecto de  $\tau^{qN}$ . Efectivamente, en este caso se verificará

$$\begin{aligned} A \cap \Omega_c = \emptyset &\implies A \cap \Omega_q = \emptyset \\ A \cap \Omega_c \neq \emptyset &\implies A \cap \Omega_q \neq \emptyset, \end{aligned} \quad (3.75)$$

y en consecuencia

$$\mathbf{H}^{SN}(A \llcorner \Omega_c) \geq \mathbf{H}^{SN}(A \llcorner \Omega_q), \quad SN = \dim_H \Omega_c. \quad (3.76)$$

Por lo tanto,  $\tau^{qN}(A) \geq \tau^q(A)$ , y

$$\tau^{qN}(A) = 0 \implies \tau^q(A) = 0, \quad (3.77)$$

lo que confirma  $\tau^q \ll \tau^{qN}$ .

Por el contrario, en el caso de que  $\dim_H \Omega_q < \dim_H \Omega_c$  resultará que  $\tau^{qN}$  y  $\tau^q$  serán mutuamente singulares. La medida de Hausdorff de una cierta dimensión es nula para cualquier conjunto de dimensión menor, de manera que  $\tau^{qN}(\Omega_q) = 0$ , y dado que  $(\Omega - \Omega_q) \cap \Omega_q = \emptyset$  resulta  $\tau^q(\Omega - \Omega_q) = 0$ , lo que por definición significa  $\tau^{qN} \perp \tau^q$ .

Ahora bien, incluso en el caso de que la definición de  $\tau^q$  varíe con  $q$ , la familia de medidas  $\{(\tau^q \times P_\Theta)_{o_q}, q \in Q\}$  definida en la Sección 3.6.2.2 constituirá un kernel de  $(Q, F_Q)$  a  $(\Omega \times \Theta, F \times F_\Theta)$  (Pollard 2002, p. 84). Por lo tanto, a fin de mantener una formulación general que incluya como instancias particulares todos los casos que han sido considerados, es conveniente definir la medida que genera  $o(\omega, \theta, q)$  sobre  $\Omega \times \Theta \times Q$  mediante el producto de  $P_Q$  con dicho kernel.

### 3.6.3 Definición de un espacio de probabilidad

En esta sección analizamos bajo qué condiciones un modelo general de optimización en la forma de la función  $o(\omega, \theta, q)$  permite definir un espacio de probabilidad sobre el espacio producto  $\Omega \times \Theta \times Q$ . Analizaremos en primer lugar qué condiciones deben darse para que la restricción de  $o(\omega, \theta, q)$  a un cierto valor de  $q$ ,  $o_q(\omega, \theta)$ , defina un espacio de probabilidad correspondiente a una hipótesis particular donde  $q$  tiene un cierto valor, es decir, una distribución de probabilidad condicional de  $(Q, F_Q)$  a  $(\Omega \times \Theta, F \times F_\Theta)$ . A continuación veremos como la familia de distribuciones de probabilidad condicional definen un espacio de probabilidad que contiene el conjunto de hipótesis de optimización para cada  $q \in Q$ . Finalmente, presentaremos una expresión de la medida de probabilidad del modelo de optimización como una familia de distribuciones de probabilidad condicional de  $(\Theta \times Q, F_\Theta \times F_Q)$  a  $(\Omega, F)$ .

#### 3.6.3.1 Definición de una medida de probabilidad en base a $o_q$

Se vio en la formulación del modelo nulo, que corresponde al caso particular  $o_{q_N}$ , que la derivada de Radon-Nikodym de la medida de probabilidad debe ser constante respecto a una cierta medida base (Sección 3.5.3.1). El mismo planteo es válido para cualquier  $o_q$ ; el análisis que sigue a continuación representa una generalización de tratado anteriormente para dicho caso particular.

Supongamos que, una vez planteado un modelo general de optimización, para un cierto valor  $q \in Q$  la función  $o_q(\omega, \theta)$  permite definir una medida de probabilidad en base al espacio de medida  $(\Omega \times \Theta, F \times F_\Theta, (\tau^q \times P_\Theta)_{o_q})$ . Para un modelo dado, si tal medida de probabilidad existe dependerá únicamente del valor de  $q$ , por lo que designaremos tal medida como  $P_q$ . La medida  $P_q$  de un  $A \in F \times F_\Theta$  resultará

$$P_q(A) = \int_{\Omega \times \Theta} I_A(\omega, \theta) dP_q. \quad (3.78)$$

Por el teorema de Radon-Nikodym resulta

$$\int_{\Omega \times \Theta} I_A(\omega, \theta) dP_q = \int_{\Omega \times \Theta} I_A(\omega, \theta) \cdot \frac{dP_q}{d(\tau^q \times P_\Theta)_{o_q}} \cdot \frac{d(\tau^q \times P_\Theta)_{o_q}}{d(\tau^q \times P_\Theta)} d(\tau^q \times P_\Theta). \quad (3.79)$$

Una distribución uniforme de probabilidad sobre  $\Omega \times \Theta$  implica que la densidad de Radon-Nikodym de  $P_q$  respecto de la medida  $(\tau^q \times P_\Theta)_{o_q}$  debe ser constante; designaremos dicha densidad como  $v_q \equiv \frac{dP_q}{d(\tau^q \times P_\Theta)_{o_q}}$ . Por su parte, de acuerdo a lo expuesto en la Sección 3.6.2.2, resulta  $\frac{d(\tau^q \times P_\Theta)_{o_q}}{d(\tau^q \times P_\Theta)} = o_q(\omega, \theta)$ . Por lo tanto,

$$\int_{\Omega \times \Theta} I_A(\omega, \theta) dP_q = \int_{\Omega \times \Theta} I_A(\omega, \theta) \cdot v_q \cdot o_q(\omega, \theta) d(\tau^q \times P_\Theta). \quad (3.80)$$

Dado que  $1 = \int_{\Omega \times \Theta} dP_q$ , se define entonces la siguiente función:

$$v(q) \equiv \frac{1}{\int_{\Omega \times \Theta} o_q(\omega, \theta) d(\tau^q \times P_\Theta)} \quad (3.81)$$

De forma similar a lo visto en el caso del modelo nulo (Sección 3.5.3.2), la constante de normalización  $v(q)$ <sup>20</sup> estará definida si la función  $o_q(\omega, \theta)$  resulta una función  $(\tau^q \times P_\Theta)$ -integrable y su integral resulta no nula, lo que equivale por definición a que su medida sea positiva y finita:

$$0 < (\tau^q \times P_\Theta)(o_q) < \infty \iff 0 < (\tau^q \times P_\Theta)_{o_q}(\Omega \times \Theta) < \infty. \quad (3.82)$$

Para el caso particular  $q = q_N$  resulta  $o_{q_N}(\omega, \theta) = c(\omega, \theta)$ , y en consecuencia

$$\begin{aligned} v(q_N) &= \frac{1}{\int_{\Omega \times \Theta} c(\omega, \theta) d(\tau^{q_N} \times P_\Theta)} \\ &= \frac{1}{\int_{\Omega} c(\omega) d\tau^{q_N}}, \end{aligned} \quad (3.83)$$

con lo que se verifica la equivalencia con la condición obtenida para la definición de un espacio de probabilidad a partir del modelo nulo (Ecuación 3.21).

Veremos ahora que si que la constante es finita y no nula para el modelo nulo ( $0 < v(q_N) < \infty$ ), entonces lo mismo sucede para todo  $q$  para el que la medida definida por  $o_q$  coincida con esta, es decir, donde se verifique  $\tau^q(A) = \tau^{q_N}(A)$  para todo  $A \in \mathcal{F}$ . De acuerdo a la definición de la función  $o$ , para cualquier  $q \in Q$  resulta  $o_q(\omega, \theta) \leq c(\omega, \theta)$ . Dada la monotonía de la integración (Athreya y Lahiri 2006, p. 49), resulta

$$\int_{\Omega \times \Theta} o_q(\omega, \theta) d(\tau^q \times P_\Theta) \leq \int_{\Omega \times \Theta} c(\omega, \theta) d(\tau^{q_N} \times P_\Theta), \quad (3.84)$$

o de forma equivalente  $(\tau^q \times P_\Theta)(o_q) \leq (\tau^{q_N} \times P_\Theta)(o_{q_N})$ . En consecuencia, si la restricción  $c(\omega, \theta)$  es integrable, cualquier  $o_q(\omega, \theta)$  también lo será:  $(\tau^q \times P_\Theta)(o_{q_N}) < \infty \implies (\tau^q \times P_\Theta)(o_q) < \infty$ .

Ahora bien, hemos vistos que si  $\Omega$  es un espacio métrico no contable puede suceder que  $\tau^q$  difiera de la medida  $\tau^{q_N}$  definida por el modelo nulo. En este será necesario pero *no* suficiente que  $0 < \tau^{q_N}(\Omega_c) < \infty$ , ya que puede ser que para  $q \in Q$  ocurra que  $\tau^q(\Omega_q) = \infty$  aún cuando  $\Omega_q \subset \Omega_c$ <sup>21</sup>. Por lo tanto, se deberá verificar *adicionalmente*

20 Para un cierto valor de  $q$ ,  $v(q)$  se trata de una constante de normalización ya que permite que  $P_q(\Omega \times \Theta) = 1$ . Si bien el valor de  $v(q)$  varía en función de  $q$ , optamos por referirnos a ésta con el término de *constante* en este contexto para no confundir con la función de normalización  $n$ .

21 En general, un subconjunto de un conjunto de medida finita puede tener una medida infinita si posee una dimensión de Hausdorff menor. Un ejemplo clásico de tal caso es la curva denominada copo de nieve de Koch, que posee una dimensión de Hausdorff de  $\frac{\ln 4}{\ln 3} \approx 1,2619$  y un perímetro resulta infinito, si bien está incluido en un conjunto finito de  $\mathbb{R}^2$ .

que se pueda definir una medida de probabilidad para aquellos  $q \in Q$  para lo que  $\tau^q(A) \neq \tau^{qN}(A)$ , con  $A \in F$ .

3.6.3.2 Definición de un espacio de probabilidad sobre  $\Omega \times \Theta \times Q$

El paso restante es la definición de un espacio de probabilidad en base a la función  $o(\omega, \theta, q)$  que incluya todos los casos particulares de modelos de optimización sobre el espacio de parámetros  $Q$  de la función de transferencia. Es posible definir tal espacio de probabilidad como el producto de la medida de probabilidad  $P_Q$  sobre  $Q$  con la familia de distribuciones de probabilidad condicional  $\{P_q : q \in Q\}$  definida en el apartado anterior.

Definimos la medida de probabilidad  $P$  sobre el espacio producto  $\Omega \times \Theta \times Q$  para un  $A \in F \times F_\Theta \times F_Q$  como:

$$P(A) = \int_Q \int_{\Omega \times \Theta} I_A(\omega, \theta, q) \cdot v_q \cdot o_q(\omega, \theta) d(\tau^q \times P_\Theta) dP_Q(q). \tag{3.85}$$

Verificaremos que  $P(\Omega \times \Theta \times Q) = 1$ . Por definición, para todo  $q \in Q$  resulta

$$\int_{\Omega \times \Theta} v_q \cdot o_q(\omega, \theta, q) d(\tau^q \times P_\Theta) = 1, \tag{3.86}$$

de manera que

$$\begin{aligned} P(\Omega \times \Theta \times Q) &= \int_Q \int_{\Omega \times \Theta} v_q \cdot o_q(\omega, \theta) d(\tau^q \times P_\Theta) dP_Q(q) \\ &= \int_Q dP_Q(q). \end{aligned} \tag{3.87}$$

Dado que la definición de la función de normalización  $v(q)$  depende de que se verifique  $0 < \tau^{qN}(\omega) < \infty$ , esta es condición necesaria y suficiente para la definición del espacio de probabilidad sobre  $\Omega \times \Theta \times Q$ .

3.6.3.3 Kernel de probabilidad de  $(\Theta \times Q, F_\Theta \times F_Q)$  a  $(\Omega, F)$

Para evaluar el soporte empírico de hipótesis en el espacio  $\Theta \times Q$  es necesario poder formular una familia de distribuciones de probabilidad condicional  $\{P_{\theta,q} : (\theta, q) \in \Theta \times Q\}$ , lo que constituirá un kernel de probabilidad de  $(\Theta \times Q, F_\Theta \times F_Q)$  a  $(\Omega, F)$ .

De igual manera que lo tratado en la Sección 3.5.3.5 para el modelo nulo, para un  $A \in F$  resultará

$$P_{\theta,q}(A) = \int_\Omega I_A(\omega) \cdot v_{\theta,q} \cdot o_{\theta,q}(\omega) d\tau, \tag{3.88}$$

con  $v(\theta, q) \equiv v_{\theta, q}$  dado por

$$v(\theta, q) = \frac{1}{\int_{\Omega} o_{\theta, q}(\omega) d\tau^{\Omega}(\omega)}. \quad (3.89)$$

La existencia del kernel de probabilidad depende por lo tanto de que se cumpla la condición  $0 < v(\theta, q) < \infty$  para todo  $(\theta, q) \in \Theta \times Q$ . Como se discutió en la Sección 3.5.3.6, si la medida  $P$  sobre  $\Omega \times \Theta \times Q$  está definida, la existencia de un kernel de probabilidad sobre  $\Omega$  está asegurada si  $(\Omega, F)$  es un espacio medible estándar (Çınlar 2011, p. 152-156).

### 3.6.3.4 Kernel de probabilidad de $(\Theta, F_{\Theta})$ a $(\Omega \times Q, F \times F_Q)$

Finalmente, resta formular una familia de distribuciones de probabilidad condicional  $\{P_{\theta} : \theta \in \Theta\}$  que constituyan un kernel de probabilidad de  $(\Theta, F_{\Theta})$  a  $(\Omega \times Q, F \times F_Q)$ . Para un  $A \in F \times F_Q$  resultará

$$P_{\theta}(A) = \int_{\Omega \times Q} I_A(\omega, q) \cdot v_{\theta} \cdot o_{\theta}(\omega, q) d(\tau^{\Omega} \times P_Q), \quad (3.90)$$

con  $v(\theta) \equiv v_{\theta}$  dado por

$$v(\theta) = \frac{1}{\int_{\Omega \times Q} o_{\theta}(\omega, q) d(\tau^{\Omega} \times P_Q)}. \quad (3.91)$$

Como ya ha sido mencionado, si la medida  $P$  sobre  $\Omega \times \Theta \times Q$  está definida, la existencia de un kernel de probabilidad sobre  $\Omega \times Q$  está asegurada si  $(\Omega \times Q, F \times F_Q)$  es un espacio medible estándar, lo que se cumple si tanto  $(\Omega, F)$  como  $(Q, F_Q)$  lo son (Çınlar 2011, p. 152-156).

## 3.7 INFERENCIA ESTADÍSTICA

Una vez expresado un modelo general de optimización en la forma de un espacio de probabilidad, se hace posible evaluar de forma objetiva el soporte empírico de una hipótesis mediante una inferencia estadística. Jaynes (2003) define la inferencia estadística como una forma de razonamiento inductivo que, a diferencia del razonamiento deductivo, resulta óptima cuando se cuenta con información incompleta. Tradicionalmente, la disciplina de la inferencia estadística ha sido considerada como independiente de la teoría de la probabilidad, si bien se nutre de ésta. Esta visión se opone a la promovida por Jaynes (2003) y, en general, por quienes abogan por una interpretación bayesiana objetiva de la probabilidad como una extensión de la lógica: de acuerdo a esta escuela de pensamiento, la distinción entre estas disciplinas es artificial.

En esta sección se abordará cómo realizar una inferencia estadística a partir de un modelo general de optimización. En primer lugar, trataremos la formulación de hipótesis y la consiguiente derivación de predicciones que puedan ser comparadas con las observaciones disponibles sobre el sistema bajo estudio. A continuación, introduciremos una transformación cuantil en la función de transferencia que facilita el planteo e interpretación de hipótesis. Luego abordaremos la definición y cómputo de la función de verosimilitud, y cómo el uso de la transformación cuantil permite sortear limitaciones inherentes al de la verosimilitud como base de una inferencia estadística. Por último, discutiremos la aplicación de los métodos más comunes de inferencia basadas en una visión frecuentista o bayesiana de la probabilidad, que varían tanto en sus fundamentos como en su grado de complejidad.

### 3.7.1 Hipótesis de optimización y predicciones

#### 3.7.1.1 Formulación de hipótesis

Hemos visto cómo, mediante la variación del parámetro  $q$ , la función de optimización  $o(\omega, \theta, q)$  permite generalizar un problema de optimización y representar un abanico de hipótesis entre el modelo nulo y la optimización estricta. Tal como fue definido, el espacio de parámetro  $Q$  asociado a la función de transferencia representa el espacio de las hipótesis de optimización consideradas.

Hasta ahora hemos definido de forma arbitraria la medida de probabilidad sobre  $Q$ , pero es justamente al especificar una cierta distribución de probabilidad  $P_Q$  que se realiza el planteo de una hipótesis particular sobre el modelo de optimización. En el caso de una función de transferencia escalonada, por ejemplo,  $q = 1$  corresponde a la hipótesis de optimización estricta,  $q = 0$  representa el modelo nulo, y el intervalo  $(0, 1)$  representa el conjunto de hipótesis de optimización en las que se ha relajado el supuesto de optimización estricta el supuesto y/ el de monotonía estricta. Tales hipótesis de tipo puntual se pueden plantear definiendo una medida  $P_Q$  concentrada en  $q = 0$  y  $q = 1$  respectivamente.

De forma general, la hipótesis nula de un modelo de optimización se asociará a una medida de probabilidad sobre  $Q$  concentrada en  $q_N$ , que denominaremos  $P_{Q_N}$  a fin de distinguirla de la distribución de probabilidad condicional  $P_{q_N}$ . En términos formales,  $P_{Q_N}$  resulta una medida de Dirac concentrada en  $q_N$ : para todo  $A \in \mathcal{F}_Q$ :

$$P_{Q_N}(A) \equiv \begin{cases} 1 & \text{si } q_N \in A, \\ 0 & \text{si } q_N \notin A. \end{cases} \quad (3.92)$$

Una hipótesis de optimización, por su parte, se representará mediante una cierta medida de probabilidad sobre  $Q$ , que denominaremos  $P_{Q_0}$ . Al contrario que la medida

correspondiente al modelo nulo,  $P_{Q_0}$  estará concentrada en  $Q - q_N$ , lo que significa que deberá ser mutuamente singular con respecto a la medida de probabilidad definida por el modelo nulo:  $P_{Q_0}(q_N) = P_{Q_N}(Q - q_N) = 0$ .

### 3.7.1.2 Interpretación de hipótesis y función de normalización

El planteo de una hipótesis de optimización particular dependerá, lógicamente, de la relación que guarde el valor de  $q$  con la función de transferencia utilizada en el modelo de optimización. Dado un conjunto de restricciones y una función objetivo, las predicciones de modelos con una distinta función de normalización y/o de transferencia coincidirán en los casos extremos del modelo nulo y de la optimización estricta, dada la forma en que fueron definidas (Sección 3.6.1.2 y Sección 3.6.1.3). Esto *no* sucederá para los casos intermedios que corresponden a una relajación de los supuestos de optimización estricta y/o monotonía estricta. Veremos que en tales casos la interpretación dependerá fundamentalmente del tipo de función de normalización que se utilice.

Supondremos a partir de este punto que, más allá de la forma funcional específica que se utilice, la función de transferencia posee un parámetro de posición en el rango  $[0, 1]$ , tal como es el caso de la función escalonada que se ha utilizado hasta ahora como ejemplo. Cuando se utiliza una función de normalización por valores ínfimo y supremo, la interpretación de una hipótesis parece sencilla a primera vista. Una hipótesis que plantea  $q = 0,95$ , por ejemplo, afirma que un estado  $\omega \in \Omega_c$  pertenece al conjunto de soluciones óptimas si la función objetivo  $f_\theta(\omega)$  tiene un valor igual o mayor al 95 % de su valor supremo evaluado en  $\Omega_c$ . En consecuencia, si se observa que efectivamente tal valor de  $q$  posee sustento empírico puede parecer, en principio, que esto sería una evidencia unívoca en favor de la existencia de optimización. Ahora bien, ¿qué pasaría si la distribución de la función objetivo de acuerdo al modelo nulo se concentrase en torno de valores cercanos a un 95 % del valor supremo? No podría diferenciarse entonces entre tal hipótesis de optimización y la hipótesis nula, ya que a partir de ambas se derivarían predicciones similares. En un caso como este, por lo tanto, la interpretación de una hipótesis de optimización variará si se conoce la distribución de probabilidad de  $f_\theta(\omega)$  que se espera de acuerdo al modelo nulo.

Las dificultades descriptas al interpretar hipótesis y predicciones pueden evitarse si se utiliza una función de normalización cuantil basada en el modelo nulo. En este caso no se necesita de información extra para interpretar una hipótesis: un parámetro de posición  $q = 0,95$  significa que un estado  $\omega \in \Omega_c$  es un óptimo si la función objetivo  $f_\theta(\omega)$  pertenece al percentil 95 de la distribución esperada de acuerdo al modelo nulo.

### 3.7.1.3 Predicciones y observaciones

Así como las hipótesis de optimización se plantean respecto de una probabilidad sobre el espacio de parámetros  $Q$ , las predicciones que se deriven de éstas depende-

rán de cuales sean los datos disponibles realizar su evaluación empírica. Los datos que pueden ser obtenidos de un sistema biológico que se hipotetiza bajo optimización representan información sobre el estado de una cierta propiedad del sistema. Eventualmente, la información disponible puede ser suficiente para especificar unívocamente que el sistema se encuentra en un cierto estado  $\omega \in \Omega$ , pero típicamente esto no será posible.

En términos de la teoría de la probabilidad, los datos representan un conjunto de observaciones sobre el estado de una variable aleatoria definida sobre un espacio muestral. Dado el espacio medible  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$  y un espacio medible  $(E, \mathcal{E})$ , una variable aleatoria  $X$  se define como cualquier función  $(\mathcal{F} \times \mathcal{F}_\Theta, \mathcal{E})$ -medible  $X : \Omega \times \Theta \rightarrow E$ . El conjunto  $E$  se denomina espacio de observación, y los datos de los que se dispone permiten establecer el estado de una variable aleatoria  $X(\omega) = x$ , con  $x \in E$ . Habitualmente las variables aleatorias sean reales, con  $E = \mathbb{R}$ , pero de forma general  $E$  será un conjunto cuya naturaleza dependerá de los datos disponibles.

Dado que una variable aleatoria  $X$  es una función medible sobre  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$ , su extensión a un dominio  $(\Omega \times \Theta \times Q, \mathcal{F} \times \mathcal{F}_\Theta \times \mathcal{F}_Q)$  definida como  $X(\omega, \theta, q) \equiv X(\omega, \theta)$  es una función medible sobre  $(\Omega \times \Theta \times Q, \mathcal{F} \times \mathcal{F}_\Theta \times \mathcal{F}_Q)$ . Por lo tanto, definido un modelo de optimización general en la forma de una medida de probabilidad  $P$ , se induce una medida de probabilidad  $P_X$  sobre  $X$ :

$$P_X(A) \equiv P(X^{-1}(A)), A \in \mathcal{E}. \tag{3.93}$$

Dado que la medida  $P$  depende de  $P_Q$ , cada hipótesis determinará una cierta distribución de probabilidad  $P_X$  de cualquier variable aleatoria  $X$  que se defina:

$$P_X(A) = \int_Q \int_{\Omega \times \Theta} I_{X^{-1}(A)}(\omega, \theta, q) \cdot v_q \cdot o_q(\omega, \theta, q) d(\mathcal{T}^\Omega \times P_\Theta) dP_Q(q). \tag{3.94}$$

En el caso de una hipótesis que plantee un valor puntual de  $q \in Q$  para una hipótesis de optimización, la medida  $P_X$  se obtendrá a partir de la distribución de probabilidad condicional para  $Q = q$  de acuerdo a la Ecuación 3.80:

$$P_X(A) = \int_{\Omega \times \Theta} I_{X^{-1}(A)}(\omega, \theta) \cdot v_q \cdot o_q(\omega, \theta, q) d(\mathcal{T}^\Omega \times P_\Theta). \tag{3.95}$$

Denominaremos como  $x \in E$  a una única observación del estado de la variable  $X$ . Dado que el espacio de observación  $E$  es un conjunto abstracto, una variable  $X$  sobre  $(E, \mathcal{E})$  podrá utilizarse para designar a cualquier elemento aleatorio. Una observación  $x \in E$  podría resultar un único valor real, pero también un vector o una matriz, o cualquier otro elemento.

Ya que  $x$  designa una observación de  $X$ , designaremos  $\mathbf{x}$  a un conjunto de  $n$  observaciones de dicha variable:  $\mathbf{x} = \{x_1, \dots, x_n\}$ ,  $x_i \in E$ , o de forma equivalente  $\mathbf{x} \in \times_{i=1}^n E$ .

Dado que la distribución de una única observación sobre  $X$  resulta  $P_X$ , la distribución de  $n$  observaciones *independientes* de una misma variable  $X$  resultará un vector aleatorio que designaremos como  $\mathbf{X}$ . El vector aleatorio así definido no será otra cosa que una variable aleatoria  $\mathbf{X} : \times_{i=1}^n \Omega \rightarrow \times_{i=1}^n E$ , y su distribución de probabilidad estará dada por la medida  $P_{\mathbf{X}} \times_{i=1}^n P_X$ .

Puede darse el caso de que se disponga de observaciones de variable aleatorias que *no* se distribuyan de manera independiente. Sean  $X_1$  y  $X_2$  variables aleatorias sobre  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$  a espacios  $(E_1, \mathcal{E}_1)$  y  $(E_2, \mathcal{E}_2)$ . Las distribuciones  $P_{X_1}$  y  $P_{X_2}$  *no* serán en general independientes, por lo que salvo en casos particulares resultará  $P_{X_1, X_2}(A) \neq P_{X_1}(A_1) \cdot P_{X_2}(A_2)$ , con  $A = A_1 \times A_2$ ,  $A_1 \in \mathcal{E}_1$  y  $A_2 \in \mathcal{E}_2$ . La distribución de observaciones  $(x_1, x_2) \in X_1 \times X_2$  tendrá una distribución conjunta definida por la variable aleatoria  $X(\omega, \theta) \equiv (X_1(\omega, \theta), X_2(\omega, \theta))$ , una función de  $(\Omega \times \Theta, \mathcal{F} \times \mathcal{F}_\Theta)$  a  $(E_1 \times E_2, \mathcal{E}_1 \times \mathcal{E}_2)$ . Esta expresión se puede generalizar a un número contable de variables aleatorias y de observaciones sobre estas.

Finalizamos esta sección con la definición de densidad de probabilidad, un concepto fundamental en la inferencia estadística debido a que es la base de la definición de la función de verosimilitud. Dada una observación  $x$  de una variable aleatoria  $X$  definida sobre  $\Omega \times \Theta \times \mathcal{Q}$ , con una distribución  $P_X$ , se define la densidad de probabilidad  $p(x)$  como la derivada de Radon-Nikodym de  $P_X$  respecto de una medida base  $\mu$  sobre  $(E, \mathcal{E})$ :

$$p(x) \equiv \frac{dP_X(x)}{d\mu(x)}. \quad (3.96)$$

La densidad  $p(x)$  depende de la medida de probabilidad  $P$  sobre  $\Omega \times \Theta \times \mathcal{Q}$ , que será la que determine la distribución  $P_X$  de la variable aleatoria. A pesar de que en la literatura se suele mencionar a  $p(x)$  sin hacer referencia explícita a la medida base  $\mu$  que domina a  $P_X$  (e.g. Casella y Berger 2002; Cox 2006), la definición de  $p(x)$  descansa en la existencia de  $\mu$ , por lo que creemos que sería más acertado hablar de densidad de probabilidad *respecto de*  $\mu$ . Veremos en la siguiente sección que este hecho puede ser de suma importancia a la hora de comparar el soporte empírico de hipótesis de optimización.

### 3.7.2 Función de verosimilitud

En el marco del modelo general de optimización planteado, realizar una inferencia sobre hipótesis de optimización equivale a inferir la distribución del parámetro  $q$  a partir de los datos observacionales de los que se dispone sobre el sistema bajo estudio. Más allá del tipo de análisis estadístico que se utilice, la capacidad de realizar una inferencia dependerá en general de una estimación de la distribución de variables alea-

torias sobre el espacio de probabilidad a fin de computar la función de verosimilitud <sup>22</sup>.

### 3.7.2.1 Definición de verosimilitud

Dado el modelo general de optimización planteado, la familia de medidas de probabilidad condicional  $\{P_q : q \in Q\}$  sobre  $\Omega \times \Theta$  definen una familia de medidas de probabilidad condicional sobre  $\Omega \times \Theta$ . Dado que con cada valor  $q$  variará la medida  $P_q$  sobre  $\Omega \times \Theta$ , como consecuencia variará a su vez la distribución de  $X$ , lo que define una distribución condicional de  $X$  respecto de  $q$ :  $\{P_{X_q} : q \in Q\}$ . Vimos en la sección anterior que dada una observación  $x$  sobre una variable aleatoria definida por una función  $F \times F_\Theta$ , E-medible  $X : \Omega \times \Theta \rightarrow E$  queda definida la densidad de probabilidad  $p(x)$  (Ecuación 3.110). En consecuencia, queda también así definida la densidad de probabilidad *condicional* de  $x$  dado  $Q = q$  como:

$$p_q(x) \equiv \frac{dP_{X_q}(x)}{d\mu(x)}. \quad (3.97)$$

Se suelen utilizar varias notaciones diferentes para designar a la función de densidad condicional:  $p(x, q)$ ,  $p(x | q)$ , y  $p(x | Q = q)$ .

La función de verosimilitud de un parámetro  $q$  dada una observación  $x$  se define como <sup>23</sup>:

$$L(x | q) \equiv p_q(x). \quad (3.98)$$

La distinción entre  $L(x | q)$  y  $p(x, q)$  se debe a que la función de verosimilitud varía en función de  $q$  y la observación  $x$  es una constante, mientras que en la función de densidad  $p_q(x)$  la variable es  $x$  mientras que  $q$  es fijo.

La relevancia de la función de verosimilitud deriva del llamado Principio de la Verosimilitud, que sostiene que *toda* la información que puede obtenerse sobre un parámetro  $q$  a partir de un experimento que resulte en una observación  $x$  está contenida en la función de verosimilitud  $L(x | q)$  (Berger y Wolpert 1988). Esto significa que, al realizar una inferencia sobre  $q$ , conocer el valor de la observación  $X = x$  no aporta información alguna si ya se conoce el valor de  $L(x | q)$  <sup>24</sup>. Una inferencia basada en la

<sup>22</sup> Salvo para el caso de una prueba de significación, que como se verá depende únicamente del modelo nulo y no requiere necesariamente el cómputo de la función de verosimilitud.

<sup>23</sup> Salvo que  $\Theta$  forme parte del espacio de observación  $E$  y  $\theta$  sea por lo tanto observado,  $q$  no es el único parámetro del modelo de optimización y, en consecuencia, la función  $L(x | q)$  definida es estrictamente una función de verosimilitud marginal, ya que se obtiene mediante la marginalización de la distribución de probabilidad  $P_\Theta$  sobre  $\Theta$ .

<sup>24</sup> En su formulación tradicional, el Principio de la Verosimilitud se aplica estrictamente si el parámetro  $q$  es el único que afecta la distribución de la variable aleatoria  $X$ . Sin embargo, si la distribución de probabilidad de otros parámetros es conocida, como en el caso de  $\theta$  ya que  $P_\Theta$  se asume conocida, entonces dicho principio se aplica efectivamente a la función de verosimilitud marginal. Berger y Wolpert (1988, sección

función de verosimilitud consiste en la comparación del valor de dicha función para diferentes valores del parámetro  $q$  en cuestión: mientras mayor sea la verosimilitud de un modelo para un cierto valor  $q$  se considera que hay una mayor evidencia en favor de dicho modelo, ya que la probabilidad de observar datos *similares* al observado resulta mayor.

### 3.7.2.2 Limitantes de la inferencia mediante verosimilitud y normalización cuantil

Como consecuencia de su definición, el uso de la función verosimilitud impone un requisito para realizar una inferencia respecto de  $q$ : las medidas de probabilidad condicional  $P_{x_q}$  sobre  $(E, \mathcal{E})$  deben ser absolutamente continuas respecto de una cierta medida base  $\mu$  que resulte  $\sigma$ -finita sobre  $(E, \mathcal{E})$ ; en otras palabras, su componente singular respecto de  $\mu$  debe ser nulo. Para que esto suceda, es condición necesaria y suficiente que las medidas de probabilidad  $P_q$  sobre  $\Omega \times \Theta$  sean absolutamente continuas con respecto a una medida base sobre  $\Omega \times \Theta$  (Albeverio y Torbin 2004, Teorema 3).

Este hecho cobra relevancia en los casos en que  $\Omega$  se trata de un espacio métrico no contable. Como se ha visto en la Sección 3.6.2.2, puede suceder entonces que la medida  $\pi$  sobre  $\Omega$  no posea una misma medida base para todo  $q \in Q$  como consecuencia de una reducción en la dimensión de Hausdorff del espacio de soluciones óptimas definido ciertos valores de  $q$ . Si esto sucede solo sería posible realizar una inferencia sobre un subconjunto de valores de  $q$  que tengan una misma medida base, es decir, una medida de Hausdorff con una misma dimensión. Dado que resulta  $\dim_H \Omega_q \leq \Omega_c$ , puede existir un valor crítico  $q_D \in (0, 1]$  tal que si  $q > q_D$  el espacio de soluciones candidatas definido tenga una dimensión de Hausdorff menor que la del modelo nulo:  $q_D \equiv \sup\{q : \dim_H \Omega_q = \Omega_c\}$  <sup>25</sup>.

Supongamos que un modelo de optimización describe adecuadamente un sistema bajo estudio con una normalización diferente de la cuantil; utilizaremos  $\hat{q}$  para referirnos al parámetro de posición de este modelo, mientras que usaremos  $q$  para un modelo idéntico salvo por la aplicación de una normalización cuantil. Si existe efectivamente un valor crítico  $\hat{q}_D$  tal que  $\hat{q} > \hat{q}_D \Leftrightarrow \dim_H \Omega_{\hat{q}} < \Omega_c$ , puede suceder que la verdadera distribución de probabilidad verificase  $P_Q(\hat{q} > \hat{q}_D) > 0$  <sup>26</sup>, es decir, que exista una probabilidad no nula de que  $\hat{q} > \hat{q}_D$ .

3.5) discuten en profundidad la relación entre definiciones alternativas de la función de verosimilitud y la aplicación del Principio de Verosimilitud.

25 Es importante recordar que, a fin de simplificar el razonamiento, que la función de transferencia posee un parámetro de posición en el rango  $[0, 1]$ . De cualquier forma, el desarrollo es válido para cualquier función de transferencia que cumpla con los requerimientos establecidos en la Sección 3.6.1.3.

26 La interpretación de una distribución de probabilidad  $P_Q$  sobre  $Q$  es directa en términos de estadística bayesiana, mientras que en términos frecuentistas correspondería al caso particular en donde  $P_Q$  es una medida de Dirac concentrada en un cierto valor  $\hat{q} > \hat{q}_D$ .

Veremos a continuación que la aplicación de una normalización cuantil permite realizar una correcta inferencia sobre el parámetro  $\hat{q}$  a través de la función de verosimilitud si el modelo de optimización cumple con una condición que analizaremos luego, con la única limitante de que se podrá estimar la evidencia en favor del conjunto de valores  $\hat{q} > \hat{q}_D$  como un solo bloque. Llamaremos a  $\Omega_{\hat{q}_D}$  al conjunto singular de soluciones óptimas definido por  $q_D$ . Sea  $P_{f_{\theta}, q_N}$  la distribución de probabilidad de la función objetivo para un  $\theta$  de acuerdo al modelo nulo, resultará

$$P_{f_{\theta}, q_N} f_{\theta}(\omega) > \hat{q}_D \times \sup_{\omega \in \Omega_c} f_{\theta}(\omega) = 0, \tag{3.99}$$

y por lo tanto

$$P_{f_{\theta}, q_N} \hat{q}_D \times \sup_{\omega \in \Omega_c} f_{\theta}(\omega) = 1. \tag{3.100}$$

Esto significa que la medida de  $P_Q$  sobre  $(\hat{q}_D, 1]$  se concentrará en  $q = 1$  si se aplica en cambio una función de normalización cuantil. Aún si el sistema está sujeto en verdad a una optimización con un valor  $q_D$  o mayor y se encuentra en un estado  $\omega_{obs} \in \Omega_{q_D}$ , con una normalización cuantil resultará  $\omega_{obs} \in \Omega_{q=1}$  y la función de verosimilitud será nula, ya que  $P(\Omega_{q^*}) = 0$ .

Ahora bien, si para un  $\omega_{obs} \in \Omega_{\hat{q}_D}$  se cumple que

$$\lim_{q \rightarrow 1} \mathbb{L}(q \mid \omega_{obs}) = \infty, \tag{3.101}$$

si bien en  $q = 1$  existirá una singularidad, se podrá realizar una aproximación para valores de  $q$  arbitrariamente cercanos a 1 y con una densidad absolutamente continua respecto de una misma medida base. Analizaremos cuándo se cumplirá esta condición. Dada una colección de conjuntos  $A \in \mathbf{A}$  con  $A \subset \Omega$  que constituya una cobertura de Vitali de  $\Omega_{q_D}$ , es decir que para cada  $\omega \in \Omega_{q_D}$  y  $\epsilon > 0$  existe un  $A$  tal que  $\omega \in A$  y  $A < \epsilon$ <sup>27</sup> (Edgar 2013, p. 10). La secuencia  $\{\Omega_{q_i}\}$  con  $q_1 < q_2 < \dots < q_n < 1$ ,  $n \in \mathbf{N}$  constituya una cobertura de  $\Omega_{q_D}$ , dado que por definición  $\Omega_{q_D} \subset \Omega_{q_i}$ . La condición que se debe cumplir es que para todo  $\Omega_{q_i}$  y todo  $\omega \in \Omega_{q_D}$  y todo  $A \in \mathbf{A}$  con  $\omega \in A$  resulta

$$P_{\Omega}^{q_N}(\Omega_{q_i} \cap A) > 0, \tag{3.102}$$

donde  $P_{\Omega}^{q_N}$  es la medida de probabilidad marginal sobre  $\Omega$  de acuerdo al modelo nulo, absolutamente continua respecto de una medida base  $\tau^N$  sobre  $\Omega$ . El espacio singular de soluciones óptimas  $\Omega_{q_D}$  constituye por definición un conjunto  $\tau^{q_N}$ -nulo, pero si se cumple la condición anterior significa que para cualquier  $\omega \in \Omega_{q_D}$  será parte

<sup>27</sup> El diámetro de un conjunto se define como  $A \equiv \sup\{\rho(x, y) : x, y \in A\}$ .

de un subconjunto de  $\Omega_{q_i}$  que tendrá una medida  $\tau^N$  no nula. En consecuencia, la densidad de probabilidad sobre  $\omega \in \Omega_{q_D}$  será no nula; resta probar que la densidad de probabilidad se incrementa cuando  $q_i \rightarrow 1$ .

De acuerdo a lo visto en la Sección 3.6.3.1, ante un mismo valor de la función  $o_{\frac{1}{q}}(\omega) = \int_{\Theta} o(\omega, \theta) dP(\theta)$  la densidad de probabilidad  $v(q)$  sobre  $\omega$  será inversamente proporcional a  $\int_{\Omega} o_q(\omega) d\tau(\omega)$ . Mientras mayor sea el valor de  $q$  mayor será el valor de  $v(q)$  y por lo tanto mayor será la densidad sobre un  $\omega \in \Omega_{q_D}$ <sup>28</sup>. Por lo tanto, la verosimilitud de  $q_i$  para un  $\omega_{obs} \in \Omega_{q_D}$  aumentará cuando  $q_i \rightarrow 1$ . El que un modelo de optimización cumpla con la condición descrita depende de las propiedades de las curvas de nivel de la función  $o(\omega, \theta, q)$ , lo que a su vez dependerá de las funciones de restricción y objetivo.

### 3.7.3 Cuantil de la función objetivo como variable aleatoria

En función de la relevancia de la normalización cuantil, una variable aleatoria que resultará de suma importancia será el cuantil de la función objetivo al que pertenece un sistema bajo estudio de acuerdo a la distribución nula.

En primer lugar definiremos una función medible  $X$  que, al aplicarse al espacio de probabilidad correspondiente a un modelo general de optimización, una observación  $x$  sobre  $X$  corresponda al cuantil de la distribución nula que se esperaríamente tenga la función objetivo para un cierto  $q \in Q$ . Tal como se vio al presentar la normalización cuantil en la Sección 3.6.1.2, dado un modelo nulo queda definida la distribución condicional  $P_{f_{\theta, q|N}}$  de la función objetivo dado  $\Theta = \theta$ , lo que define a su vez una distribución acumulada condicional  $F_{f_{\theta, q|N}}$ . Definimos una función medible  $h : \mathbb{R} \times \Theta \rightarrow [0, 1]$  como

$$h(y, \theta) = F_{f_{\theta, q|N}}(y), y \in \mathbb{R}, \theta \in \Theta. \quad (3.104)$$

La composición de la función  $h(y, \theta)$  sobre  $f(\omega, \theta)$  define una función medible

$$X(f(\omega, \theta), \theta) \equiv F_{f_{\theta, q|N}}(f(\omega, \theta)). \quad (3.105)$$

Dado un cierto sistema bajo estudio, la función  $X$  así definida será observable si y solo si se dan las siguientes condiciones:

<sup>28</sup> Dada la definición de la función  $o$  y de la función de transferencia  $t$ , para un parámetro de posición  $q_1 < q_2$  resulta  $o_{q_1}(\omega) \geq o_{q_2}(\omega)$ , y dada la monotonía de la integración resultará

$$\int_{\Omega} o_{q_1}(\omega) d\tau(\omega) > \int_{\Omega} o_{q_2}(\omega) d\tau(\omega), \quad (3.103)$$

y en consecuencia  $v(q_1) < v(q_2)$ . La densidad de probabilidad  $p_q(\omega)$  definida por una  $o_q$  es igual a  $o_q(\omega) \cdot v(q)$ . Por lo tanto, si para un valor observado  $\omega_{obs}$  dado resulta  $o_{q_1}(\omega_{obs}) = o_{q_2}(\omega_{obs})$ , entonces resultará  $L(\omega_{obs} | q_2) > L(\omega_{obs} | q_1)$ .

1. El valor del parámetro  $\theta$  al que corresponde una cierta observación debe ser conocido, es decir,  $\Theta$  debe ser parte del espacio de observación experimental.
2. El valor de la función objetivo debe ser observable, ya sea mediante una cuantificación directa de la función objetivo, en cuyo caso  $X$  definirá una función  $\mathbf{R} \times \Theta \rightarrow [0, 1]$ ; o mediante una determinación indirecta a partir de la observación directa de  $\omega$ , en cuyo  $X$  definirá una función  $\Omega \times \Theta \rightarrow [0, 1]$ .

Cabe destacar que la primera condición está asegurada en el caso de un modelo de optimización determinista, donde existirá un único  $\theta$ .

Dado que la variable  $X$  está definida en el espacio medible Borel estándar  $([0, 1], \mathbf{B}([0, 1]))$ , está asegurada la existencia de una distribución de probabilidad condicional respecto a cualquier elemento de  $\mathbf{F} \times \mathbf{F}_\Theta \times \mathbf{F}_Q$ . Una función  $h_\theta : \mathbf{R} \rightarrow [0, 1]$  aplicada a una variable aleatoria  $f_\theta : \Omega \rightarrow \mathbf{R}$  define una variable aleatoria  $X_\theta : \Omega \rightarrow [0, 1]$ . La distribución de probabilidad condicional  $P_{\theta,q}$  sobre  $\Omega$  (Sección 3.6.3.3) define la distribución condicional de la función objetivo:

$$P_{f_{\theta,q}}(A) \equiv P_{\theta,q} f_\theta^{-1}(A) , A \in \mathbf{B}(\mathbf{R}) . \tag{3.106}$$

Al aplicar sobre esta distribución la función  $h_\theta(y)$  se puede obtener  $P_{X_{\theta,q}}$  como (Çınlar 2011, p. 52)

$$\begin{aligned} P_{X_{\theta,q}}(A) &\equiv P_{\theta,q} X_{\theta,q}^{-1}(A) \\ &= P_{f_{\theta,q}} h_\theta^{-1}(A) \\ &= P_{f_{\theta,q}} F_{f_{\theta,qN}}^{-1}(A) , \end{aligned} \tag{3.107}$$

donde  $A \in \mathbf{B}([0, 1])$ .

La distribución de probabilidad de  $X$  de acuerdo al modelo nulo resulta:

$$P_{X_{\theta,qN}}(A) = P_{f_{\theta,qN}} F_{f_{\theta,qN}}^{-1}(A) . \tag{3.108}$$

Por definición de la función cuantil, para un intervalo  $(a, b]$  con  $a < b$  y  $a, b \in [0, 1]$  resulta

$$\begin{aligned} P_X(F_X^{-1}((a, b])) &= P_X F_X^{-1}(a), F_X^{-1}(b) \\ &= P_X F_X^{-1}(a) < X \leq F_X^{-1}(b) \\ &= F_X F_X^{-1}(b) - F_X F_X^{-1}(a) \\ &= b - a . \end{aligned} \tag{3.109}$$

La medida  $P_{X_{\theta,qN}}$  tiene por lo tanto una distribución uniforme en el intervalo  $[0, 1]$ , y dado que la medida es igual para todo  $\theta$ ,  $P_{X_{qN}}$  posee la misma distribución. Para una distribución uniforme en  $[a, b]$ , la derivada de Radon - Nikodym respecto de la medida

de Lebesgue resulta constante e igual a  $\frac{1}{b-a}$ . Por lo tanto, para una observación  $x$  la función de densidad de probabilidad condicional a  $Q = q_N$  resulta

$$\begin{aligned} p_{q_N}(x) &= \frac{dP_{X_{q_N}}(x)}{d\mu(x)} \\ &= 1. \end{aligned} \quad (3.110)$$

Por lo tanto, la verosimilitud de  $q_N$  en función de una observación  $x$  resulta constante

$$L(q_N | x) = 1. \quad (3.111)$$

Esto significa que de acuerdo al modelo nulo esperamos que la distribución de probabilidad del cuantil de una función objetivo observada es uniforme en  $[0, 1]$ .

Para el caso general de una hipótesis  $Q = q$  resulta

$$P_{X_{\theta,q}}(A) = P_{f_{\theta,q}} F_{f_{\theta,q_N}}^{-1}(A). \quad (3.112)$$

Supongamos que la función de transferencia se trata de una función escalonada. Reemplazando  $A = (q, 1]$  resulta

$$\begin{aligned} P_{X_{\theta,q}}((q, 1]) &= P_{f_{\theta,q}} F_{f_{\theta,q_N}}^{-1}((q, 1]) \\ &= P_{f_{\theta,q}} F_{f_{\theta,q_N}}^{-1}(q), F_{f_{\theta,q_N}}^{-1}(1) \\ &= P_{f_{\theta,q}} F_{f_{\theta,q_N}}^{-1}(q), \infty \\ &= F_{f_{\theta,q}}(\infty) - F_{f_{\theta,q}} F_{f_{\theta,q_N}}^{-1}(q) \\ &= 1 - F_{f_{\theta,q}} F_{f_{\theta,q_N}}^{-1}(q). \end{aligned} \quad (3.113)$$

Ahora bien, nótese que por definición  $F_{f_{\theta,q_N}}^{-1}(q)$  es el valor de  $f_{\theta}(\omega)$  al que le corresponde el percentil  $q$ , pero de acuerdo a la función de transferencia escalonada para todo  $0 < q_i < q$  resultará  $t(F_{f_{\theta,q_N}}^{-1}(q_i), q) = 0$ , y por lo tanto  $F_{f_{\theta,q}} F_{f_{\theta,q_N}}^{-1}(q) = 0$ . En consecuencia, para todo  $\theta \in \Theta$  resulta  $P_{X_{\theta,q}}((q, 1]) = 1$ , lo que significa que la medida de probabilidad está concentrada en  $(q, 1]$  y es uniforme en dicho intervalo, mientras que el intervalo  $(0, q]$  resulta un conjunto  $P_{X_{\theta,q}}$ -nulo. La derivada de Radon - Nikodym determina, de hecho, una densidad uniforme en el intervalo  $(q, 1]$ :

$$\begin{aligned} p_q(x) &= \begin{cases} \frac{1}{1-q} & \text{si } x \geq q, \\ 0 & \text{si } x < q. \end{cases} \end{aligned} \quad (3.114)$$

Por lo tanto, la verosimilitud de  $q$  en función de una observación  $x$  resulta

$$L(q | x) = \begin{cases} \frac{1}{1-q} & \text{si } q \leq x, \\ 0 & \text{si } q > x. \end{cases} \quad (3.115)$$

Nótese que se verifica que si  $q = q_N = 0$  resulta  $L(q | x) = 1$  para todo  $x \in E$ . Para la función de transferencia escalonada, tanto para la hipótesis nula y la familia de hipótesis alternativas la variable aleatoria  $X$  posee una distribución uniforme en un intervalo  $[q_m, 1]$  que difiere únicamente en el límite inferior  $q_m$ : para la nula resulta  $q_m = q_N = 0$ , mientras que para una alternativa  $q_m \neq q_N > 0$ .

### 3.7.4 Métodos de inferencia estadística

En esta sección trataremos brevemente los métodos estadísticos que pueden ser utilizados para realizar una inferencia sobre las hipótesis de optimización. Comenzaremos por tratar el clásico contraste de una hipótesis nula mediante una prueba de significación, análisis basado únicamente en el modelo nulo. A diferencia de este, los métodos más generales se basan en la comparación entre el modelo nulo y diferentes hipótesis alternativas a través de la función de verosimilitud mediante, e incluyen a la prueba de cociente de verosimilitud, el enfoque de selección de modelos basado en la teoría de la información, y métodos estadísticos bayesianos.

#### 3.7.4.1 Prueba de significación de la hipótesis nula.

El método de inferencia más simple que puede aplicarse a la evaluación de una hipótesis de optimización se trata de una prueba frecuentista de significación de hipótesis nula. La hipótesis nula supone la ausencia de un mecanismo de interés, que en caso de estudio es un fenómeno de optimización que se da para  $q = q_N$ . La hipótesis alternativa debe ser mutuamente excluyente con la hipótesis nula, y en el caso de la optimización  $q \neq q_N$ ; si se trabaja con la función de transferencia escalonada sería  $q > q_N$ .

Una prueba de significación se basa en la comparación del valor que toma una propiedad observada del sistema con la distribución de probabilidad que se espera de esta de acuerdo a una hipótesis nula. En primer lugar se debe especificar en base a las observaciones disponibles  $\mathbf{x}$  un vector aleatorio  $\mathbf{X}$ , que tendrá una distribución de probabilidad  $P_{\mathbf{x}}$ . A continuación se define un estadístico de prueba  $T$  como una función del vector aleatorio  $\mathbf{X}$  de forma que el valor observado del estadístico resulta  $t = T(\mathbf{x})$ . Al ser una función sobre  $\mathbf{X}$  el estadístico  $T$  resulta a su vez una variable

aleatoria sobre  $\Omega$  con una distribución  $P_T$ <sup>29</sup>. El estadístico  $T$  se debe definir de tal manera que valores extremos de  $t$  sean improbables si la hipótesis nula es verdadera.

Si el modelo nulo de optimización es verdadero, la probabilidad de observar al sistema bajo estudio en un estado  $\omega$  dependerá únicamente de la función de restricción, y *no* de la función objetivo como sucedería si la optimización se da efectivamente. Una variable aleatoria  $X$  que reúne tales características es el cuantil de la función objetivo de acuerdo a la distribución nula, definida en la Ecuación 3.105 presentada Sección 3.7.3, que como se vio tendrá una distribución uniforme en  $[0, 1]$  de acuerdo a la hipótesis nula. Definida la variable aleatoria, resta definir el estadístico de prueba, que idealmente debería tener la propiedad de suficiencia respecto al modelo nulo y el parámetro  $q$ , que significa que ningún otro estadístico que pueda ser calculado sobre la misma muestra proporciona una información adicional sobre el valor de dicho parámetro (Fisher 1922). Si se considera una función de transferencia escalonada, la distribución de  $X$  será uniforme en  $[q_m, 1]$  con  $0 < q_m \leq 1$ , por lo que diferirá de la hipótesis nula únicamente en el parámetro de posición inferior. El estadístico

$$T(\mathbf{X}) = \min \mathbf{X} \quad (3.116)$$

resulta un estadístico suficiente para el parámetro de posición inferior de una distribución uniforme dado un conjunto de observaciones independientes  $\mathbf{x}$  (Lehmann y Romano 2006, p. 20)

Para finalizar la construcción de la prueba de significación es necesario establecer la denominada región de rechazo, el intervalo o intervalos de valores extremos que se esperaría tome la variable  $T$  en caso de que la hipótesis nula sea falsa. El estadístico  $T(\mathbf{X})$  es una variable aleatoria, y por lo tanto tendrá una cierta distribución  $P_T$  de acuerdo a la hipótesis nula. Suponiendo una región de rechazo en torno a valores máximos de  $T$ , se define el valor  $p$  como<sup>30</sup>

$$p \equiv P_T T(\mathbf{X}) \geq T(\mathbf{x}) \quad (3.117)$$

La hipótesis nula es rechazada si  $p$  es menor que un valor fijado *a priori* que se denomina nivel de significación ( $\alpha$ ), es decir, si la probabilidad de observar un valor del estadístico  $T$  tan extremo como el correspondiente a los datos observados es lo suficientemente

29 El estadístico para un vector aleatoria  $n$ -dimensional resulta una función  $T : \mathbb{X}_{i=1}^n E \rightarrow \mathbb{R}$ , y su distribución resulta  $P_T(A) \equiv P_{\mathbf{X}}(T^{-1}(A))$ ,  $A \in \mathbb{B}(\mathbb{R})$ .

30 Veech (2012) establece una distinción entre una pruebas de significación directas, como la que aquí se describe, y pruebas indirectas. En una prueba directa el valor  $p$  se computa como el percentil empírico del estadístico de acuerdo al modelo nulo. En contraste, en una prueba indirecta la significación del estadístico  $T$  se evalúa mediante otro análisis, por ejemplo una prueba de chi cuadrado, una prueba de Kolmogorov-Smirnov o una regresión. Veech recomienda las pruebas directas dado que no requieren de ningún supuesto sobre la distribución nula.

baja. El valor del nivel de significación usado comúnmente en estudios biológicos es de  $\alpha = 0,05$ .

Las pruebas de significación han sido ampliamente utilizados en la biología así como en otras ciencias fácticas, al punto que se han convertido en un estándar de facto. Una prueba de significación posee la ventaja de una mayor simplicidad de implementación en comparación con otros métodos estadísticos de inferencia. La mayoría de los modelos nulos en áreas de la biología como la ecología son, de hecho, utilizados como una prueba de significación a fin de poner a prueba una hipótesis nula (Veech 2012). En el caso de la optimización, esta aproximación puede aportar información valiosa si la existencia misma de una optimización es relevante más allá de su magnitud. Este resulta el caso cuando una función objetivo carece de suficiente sustento empírico, por ejemplo en el caso de hipótesis de optimización de fundamento termodinámico a niveles superior al individual.

Como contraparte de su simplicidad, las pruebas de significación presentan severas limitaciones. En el caso de un modelo de optimización, la aplicación de este tipo de pruebas está limitada a los casos particulares donde el cuantil de la función objetivo sea una variable observable, lo que como se vio sucederá si tanto  $\theta$  como la función objetivo o el estado  $\omega$  son observables. Desde el punto de vista de su interpretación, y de forma más general, la principal limitante de las pruebas de significación reside en que esta aproximación *no* permite evaluar la probabilidad de que la hipótesis nula sea verdadera, si bien es común que el valor  $p$  sea erróneamente interpretado de esta manera. Una prueba de significación permite evaluar, únicamente, qué es lo que se esperaría si la hipótesis nula fuese verdadera: su resultado depende solamente del acuerdo que exista entre los datos disponibles y una predicción derivada de la hipótesis nula, ya que no se lo compara con predicciones derivadas de la hipótesis alternativa <sup>31</sup>. Una prueba de significación no permite en consecuencia diferenciar entre el soporte empírico de múltiples hipótesis alternativas, por ejemplo, si se consideran posibles diferentes funciones objetivo. A causa de estos motivos, la popularidad de esta metodología no la exime de cuestionamientos, y las pruebas de significación han sido, de hecho, severamente criticadas (e.g. Gill 1999; Johnson 1999; Nicholls 2001). Como conclusión, de ser posible su implementación, se recomienda optar por algunos de los otros análisis estadísticos que se describen a continuación.

#### 3.7.4.2 Prueba de cociente de verosimilitud.

La prueba de cociente de verosimilitud es un método de inferencia frecuentista que compara el ajuste de modelos anidados, un modelo que resulta un caso particular de otro modelo, que en el caso de la optimización corresponde al modelo nulo y al modelo

<sup>31</sup> Esto se aplica al caso general de una prueba; en una prueba particular puede ser que el estadístico depende tanto del modelo nulo como de la hipótesis alternativa, como en el caso de la prueba de cociente de verosimilitud, que se presenta más adelante.

general. El nombre de esta prueba se debe a que el estadístico asociado se define como el cociente entre el modelo nulo de máxima verosimilitud, donde el parámetro está restringido a un subconjunto  $q \in Q_{nulo}$ , y la máxima verosimilitud para  $q \in Q_{alt}$  correspondiente al modelo de optimización general:

$$T(\mathbf{x}) = \frac{\sup_{q \in Q_{nulo}} L(q_N | \mathbf{x})}{\sup_{q \in Q_{alt}} L(q | \mathbf{x})}. \quad (3.118)$$

En el caso del modelo nulo resultará  $Q_{nulo} = q_N$ . El conjunto de valores de  $q$  de la hipótesis alternativa,  $Q_{alt} \subseteq Q$ , puede coincidir con el espacio de parámetros  $Q$  o no, pero se debe verificar que  $Q_{nulo} \subset Q_o$ , es decir, que la hipótesis alternativa incluya al modelo nulo como caso particular.

Una prueba de cociente de verosimilitud se define estableciendo la región de rechazo como  $\{\mathbf{x} : T(\mathbf{x}) \leq c\}$ , con  $0 \leq c \leq 1$ . En función de la naturaleza de la hipótesis alternativa se distinguen dos casos en cómo se define el valor de  $c$  y se obtiene una inferencia a partir del estadístico observado  $t = T(\mathbf{x})$ . Se observa que esta prueba es un tipo de prueba de significación de hipótesis nula, pero se lo presenta en un diferente apartado porque tiene la particularidad de que su resultado depende no solo del modelo nulo, sino también de la hipótesis alternativa mediante el modelo de máxima verosimilitud, y de ahí deriva su particular relevancia.

Una hipótesis que contempla una única distribución posible para la variable aleatoria  $\mathbf{X}$  se denomina hipótesis simple. Una hipótesis de optimización simple plantea que el parámetro  $q$  toma un cierto valor  $q_{alt}$ , que es planteado antes de conocer las observaciones  $\mathbf{x}$ , y en este caso  $Q_{alt}$  resulta el conjunto  $\{q_N, q_{alt}\}$ . Neyman y Pearson (1933) demostraron que para hipótesis simples el estadístico de cociente de verosimilitud constituye la prueba con el mayor poder estadístico con un nivel de significación  $\alpha$  si el valor de  $c$  verifica  $P_T(T(\mathbf{x}) \leq c) = \alpha$ .

Una hipótesis que contempla más de una distribución posible para  $\mathbf{X}$  se denomina compuesta. El teorema de Karlin-Rubin es una generalización del lema de Neyman-Pearson, y afirma que el cociente de verosimilitud es la prueba uniformemente más poderosa para hipótesis compuestas si dicho cociente posee la propiedad de monotonía (Casella y Berger 2002, p. 391). La propiedad de monotonía se cumple para una variable aleatoria  $X$  si dadas las funciones de densidad de probabilidad para la hipótesis nula  $p_{q_N}(x)$  y la hipótesis alternativa  $p_q(x)$  se verifica que su cociente es monótonamente no decreciente

$$x_2 > x_1 \implies \frac{p_q(x_2)}{p_{q_N}(x_2)} \geq \frac{p_q(x_1)}{p_{q_N}(x_1)}, \quad (3.119)$$

o alternativamente monótonamente no creciente. Una variable aleatoria  $X$  que cumple estos requisitos es el cuantil de la función objetivo de acuerdo a la distribución nula,

definida en la Ecuación 3.105. Nuevamente analizaremos como ejemplo el caso particular de una función de transferencia escalonada, pero las conclusiones serán válidas para cualquier función de transferencia dada su definición. De acuerdo a la función de verosimilitud presentada en la Ecuación 3.115,  $L(q | x)$  es no decreciente respecto de  $x$  para todo  $q \in Q$ , mientras que es constante para  $q_N$ , por lo que se verifica la condición de monotonía del cociente de verosimilitud expresada en la Ecuación 3.119. Dado que toda función de transferencia debe ser creciente respecto de  $n(\omega, \theta)$ , la densidad de probabilidad y por lo tanto la verosimilitud de un cierto  $q$  será no decreciente respecto de  $x$ <sup>32</sup>.

Una vez definidas las hipótesis y una apropiada prueba de cociente de verosimilitud, para determinar la región de rechazo es necesario conocer la distribución del estadístico  $T(\mathbf{X})$  que se espera si la hipótesis nula es verdadera. En general dicha distribución es difícil de obtener, pero el teorema de Wilks (1938) provee una estimación asintótica de su distribución: cuando el número de observaciones  $n \rightarrow \infty$  la distribución de  $-2 \log T(\mathbf{x})$  tiende a una distribución  $\chi_k^2$ , donde  $k$  es la diferencia entre los grados de libertad de ambos modelos, es decir, la dimensionalidad entre  $Q_{nulo}$  y  $Q_{alt}$ . El valor crítico  $c$  se ajusta entonces de acuerdo a la distribución  $\chi_k^2$ . La explicación de este resultado es que si la hipótesis nula es verdadera, entonces los parámetros libres extra que tiene el modelo más general solo introducen variaciones aleatorias en el ajuste del modelo. La prueba de Wald y la prueba de los multiplicadores de Lagrange son asintóticamente equivalentes a la prueba de cociente de verosimilitud (Engle 1984), es decir, tienen el mismo resultado cuando  $n \rightarrow \infty$ .

La prueba de cociente de verosimilitud presenta ventajas respecto de una prueba general de significación de hipótesis nula, ya que tiene en cuenta el ajuste tanto del modelo nulo como del modelo general de optimización. Esta prueba es aplicable a cualquier modelo de optimización, ya que tal como ha sido definido el modelo nulo es siempre un caso particular de este. A pesar de dichas ventajas, esta prueba resulta un tipo particular de prueba de significación de hipótesis nula, con lo que es sujeto de las críticas antes expuestas a este tipo de inferencia estadística. En el caso de hipótesis compuestas es requisito que el cuantil de la función objetivo sea una variable observable, lo que restringe su aplicación. La estimación de la distribución esperada por el modelo del cociente de verosimilitud es generalmente compleja, y la aproximación mediante el teorema de Wilks puede diferir mucho dependiendo del modelo y del número de observaciones disponibles. Dado que esta prueba solo permite evaluar modelos anidados, no será posible comparar entre sí hipótesis de optimización alternativas si son de tipo compuestas. Esta es una severa limitación si, por ejemplo, se desea evaluar hipótesis de optimización basadas en diferentes funciones objetivo.

32 Salvo en el caso de que la función objetivo efectivamente computada no sea la función objetivo que verdaderamente se hipotetiza bajo optimización y se relaje el supuesto de monotonía entre ambas, con lo que la función de transferencia podría ser decreciente respecto de  $n(\omega, \theta)$  cuando  $n(\omega, \theta) \rightarrow 1$ .

### 3.7.4.3 Criterio de información de Akaike.

En el marco de la teoría de la información, el criterio de Akaike (abreviado AIC en sus siglas en inglés) representa una herramienta que permite realizar una selección de modelos estadísticos de acuerdo a un balance entre el ajuste y la complejidad de los modelos, basada en una minimización de la información de Kullback-Leibler perdida al utilizar un cierto modelo candidato en relación a otro modelo (Burnham y Anderson 2002). El AIC prioriza modelos que poseen una mayor verosimilitud en relación a la complejidad del modelo, cuantificada en la forma de los grados de libertad que posee, ya que un modelo más complejo tenderá a un mejor ajuste de los datos simplemente por efecto del azar. Consideremos que para un problema de optimización existen un conjunto de  $h$  hipótesis que se expresan en relación al espacio de parámetros  $Q$  como afirmaciones  $q \in Q_1, \dots, q \in Q_h$ , con  $Q_1, \dots, Q_h \subseteq Q$ , y donde para cada hipótesis se obtiene el modelo de máxima verosimilitud. Lógicamente, una de las hipótesis debe ser la hipótesis nula con  $Q_i = q_N$ .

El AIC del modelo candidato  $i$  se define como

$$AIC_i = 2 \cdot k - 2 \cdot \ln \sup_{q \in Q_i} L(q | \mathbf{x}), \quad (3.120)$$

donde  $k$  representa la dimensionalidad del subespacio de parámetros  $Q_i$ , lo que se denomina comúnmente como el número de parámetros libres (Akaike 1974). Es importante destacar que la expresión 3.120 es asintóticamente válida, es decir, cuando el número de observaciones  $n \rightarrow \infty$  el  $AIC_i$  es un estimador insesgado de la pérdida de información y, por lo tanto, asegura la selección de los modelos que la minimizan. Cuando  $n$  es bajo en relación al número de parámetros libres  $k$  el uso de AIC resulta fuertemente sesgado a la selección de modelos con un exceso de parámetros, y por lo tanto debe utilizarse otro estimador de la información. Sugiura (1978) demostró que el llamado Índice de Información de Aikake corregido o de segundo orden,

$$AIC_{Ci} = 2 \cdot k - 2 \cdot \ln \sup_{q \in Q_i} L(q | \mathbf{x}) + \frac{2k^2 + 2k}{n - k - 1}, \quad (3.121)$$

es insesgado en la selección de modelos univariados, lineales y con distribución normal de errores si el modelo verdadero está incluido entre los candidatos. La performance del AICc se ha estudiado para otros modelos, y su uso es recomendado en general en lugar del AIC (e.g. Cavanaugh y col. 1997; Burnham y Anderson 2002), ya que ambos son asintóticamente equivalentes, es decir,  $\lim_{n \rightarrow \infty} AIC_{Ci} = \lim_{n \rightarrow \infty} AIC_i$ .

Una vez que se ha computado para cada uno de los modelos candidatos el AIC, AICc u otra versión del índice, los modelos son comparados entre sí mediante el  $\Delta AIC_i$ , que se define como

$$\Delta AIC_i = AIC_i - \min_{i \in \{1, \dots, h\}} AIC_i. \quad (3.122)$$

Los modelos seleccionados son aquellos con el menor valor de  $\Delta AIC_i$ , dado que dichos modelos poseen una mayor verosimilitud en relación a la complejidad que poseen. Una práctica recomendada al utilizar el AIC es la revisión de resultados para comparar los modelos candidatos y excluir aquellos que incluyen un único parámetro extra respecto de otro modelo pero cuya mejora en el log L es menor a dos unidades respecto del modelo restringido. Esta práctica se recomienda para evitar la selección de modelos que incluyan los así denominados parámetros no informativos, lo que sucede si un modelo tienen un AIC cercano pero menor a otro modelo más sencillo simplemente porque la inclusión de un único parámetro extra tiene una penalización reducida aún si dicho parámetro no aporta información alguna (Arnold 2010). La evidencia en favor de cada una de las hipótesis, finalmente, se determina mediante el cómputo del peso de Akaike de cada modelo candidato, un valor en el rango (0, 1):

$$w_i = \frac{e^{-\frac{\Delta AIC(i)}{2}}}{\sum_{j \in \{1, \dots, h\}} e^{-\frac{\Delta AIC(j)}{2}}}. \quad (3.123)$$

Bajo el supuesto de que el verdadero modelo que explica las observaciones se encuentra entre los modelos candidatos, el peso de Akaike para un modelo candidato se interpreta como la probabilidad de que dicho modelo sea correcto.

La selección de modelos mediante AIC tiene múltiples ventajas respecto de las pruebas de significación estadística. Una de las principales ventajas es que se pueden comparar modelos no anidados: en un mismo análisis se puede evaluar la evidencia en favor del modelo nulo y de múltiples hipótesis de optimización alternativas. Otra ventaja fundamental respecto de las pruebas de significación es que no existen limitantes en la definición de la variable aleatoria  $X$  sobre la que se realizan las observaciones  $\mathbf{x}$ . Se ha demostrado que la selección de modelos mediante AIC es asintóticamente equivalente a otros métodos de selección de modelos como la validación cruzada dejando uno afuera (Stone 1977). También se ha mostrado que puede ser derivado en un marco estadístico bayesiano similar al utilizado para definir el Índice de Información Bayesiano (BIC) y con una performance comparable, con la única diferencia de utilizar diferentes distribuciones previas para las hipótesis (Burnham y Anderson 2002, p. 293-305).

La principal desventaja en el uso del AIC o AICc es que para un conjunto arbitrario de modelos y un número reducido de observaciones  $n$ , la única forma de asegurar

una inferencia correcta es la realización de estudios de simulación. Dichos estudios se realizan típicamente por el método de Monte Carlo, y consisten en simular observaciones a partir de diferentes modelos, para luego evaluar el resultado de una inferencia basada en las observaciones a través de AIC u otros índices similares (e.g. Burnham y Anderson 2002, p. 271).

#### 3.7.4.4 Factor de Bayes

En el marco de la estadística bayesiana, es posible realizar una inferencia respecto de modelos candidatos mediante el cómputo del factor de Bayes, que se define como el cociente entre la densidad de probabilidad marginal de los datos  $\mathbf{x}$  de dos modelos alternativos (Kass y Raftery 1995). Consideremos un conjunto de  $h$  hipótesis que se expresan en relación al espacio de parámetros  $Q$  como afirmaciones  $q \in Q_1, \dots, q \in Q_h$ , con  $Q_1, \dots, Q_h \subseteq Q$ . El factor de Bayes en favor de la hipótesis  $H_i$  respecto de la hipótesis  $H_j$  resulta

$$B_{ij} = \frac{p(\mathbf{x} | H_i)}{p(\mathbf{x} | H_j)}. \quad (3.124)$$

Los términos  $p(\mathbf{D} | H_i)$  son obtenidos por integración de la verosimilitud de los modelos con respecto a los parámetros libres que poseen

$$p(\mathbf{x} | H_i) = \int p(\mathbf{x} | q, H_i) \cdot p(q | H_i) dq, \quad (3.125)$$

y es por esto que recibe el nombre de densidad marginal de probabilidad, o la denominación equivalente de verosimilitud marginal. El factor de Bayes  $B_{ij}$  representa la evidencia en favor de la hipótesis  $i$  vs.  $j$  de acuerdo a las observaciones  $\mathbf{x}$ . De hecho, el cociente entre la probabilidad posterior de las hipótesis es el producto de  $B_{ij}$  y la probabilidad previa de las hipótesis:

$$\frac{p(H_i | \mathbf{x})}{p(H_j | \mathbf{x})} = B_{ij} \cdot \frac{p(H_i)}{p(H_j)}. \quad (3.126)$$

El factor de Bayes establece de manera natural un compromiso entre la complejidad del modelo y su ajuste (Kass y Raftery 1995). Jeffreys (1998) plantea una escala para la interpretación de  $B$  a fin de juzgar el soporte en favor de la hipótesis, y plantea que  $B > 3,2$  puede considerarse como evidencia a favor,  $B > 10$  se interpreta como evidencia fuerte a favor de la hipótesis, y con  $B > 100$  se considera soporte concluyente a favor de la hipótesis. En el caso particular de que ambas sean hipótesis simples que planteen  $Q_i = q_i, Q_j = q_j$ , el factor de Bayes coincide con el cociente de verosimilitud.

El uso del Factor de Bayes posee las ventajas asociadas a la fundamentación teórica de la estadística bayesiana, y no existen limitantes respecto del modelo en el que puede

ser aplicado. Como contrapartida, la estimación del factor de Bayes puede requerir la estimación numérica de integrales en múltiples dimensiones que representan un desafío mayor que en las otras alternativas. Por lo tanto, puede requerirse el uso de métodos computacionalmente intensivos, como algoritmos de tipo Monte Carlo por Cadenas de Markov (MCMC en sus siglas en inglés).

### 3.8 DISCUSIÓN

En este trabajo hemos presentado una generalización de un modelo de optimización con restricciones en su formulación típica que, bajo ciertas condiciones, define un modelo probabilístico y posibilita entonces una inferencia mediante métodos estadística estándar basados en la función de verosimilitud. Este tipo de modelos de optimización son los que se utilizan en la mayoría de los trabajos sobre sistemas biológicos, y representan el 80 % de los modelos utilizados en la selección de trabajos analizados en el Capítulo 2 (Tabla 2.3).

De acuerdo a lo desarrollado, las condiciones necesarias y en su conjunto suficientes para que a partir de un problema de optimización pueda definirse un modelo probabilístico son las siguientes:

1. La función de optimización  $o(\omega, \theta, q)$  deber ser una función  $F \times F_{\Theta} \times F_{\Omega}$ ,  $\mathbf{B}([0, 1])$ -medible.
2. La función que define el modelo nulo debe tener una medida finita:

$$0 < \tau^{q_N}(o_{q_N}(\omega, \theta)) < \infty .$$

3. Si el espacio muestral  $\Omega$  se trata de un espacio métrico no contable, para aquellos  $q \in \{q : \dim_H \Omega_q < \dim_H \Omega_c\}$  debe cumplirse que  $\Omega_q$  sea  $\sigma$ -finito respecto a  $\mathbf{H}^{\dim_H \Omega_q}$ .

Las condición de mensurabilidad de la función que define al modelo general de optimización es una propiedad heredada de las funciones que la conforman: las restricciones, la función objetivo, y las funciones de transferencia y normalización. Dado que el conjunto de las funciones medibles incluye no solo a las funciones elementales, sino a todas aquellas que son típicamente utilizadas en el planteo de modelos matemáticos, la condición de mensurabilidad no es en modo alguna restrictiva. En efecto, las funciones utilizadas en la formulación de las restricciones y funciones objetivo de los modelos de optimización son medibles, esta condición se encuentra asegurada si se utilizan funciones de transferencia y normalización medibles. La segunda condición establece que es posible expresar un modelo de optimización de la formulación general planteada siempre que el modelo nulo conformado por las restricciones aplicadas

sobre el espacio muestral definan un conjunto de soluciones candidatas con una medida positiva y finita. En otros términos, esta condición significa que el modelo nulo define un espacio de probabilidad y permite así el cómputo o estimación de la función de verosimilitud, y esta es a su vez una condición necesaria y suficiente para que el modelo de optimización general también lo haga. La última condición se aplica cuando el espacio muestral es métrico y no contable, y se relaciona con la geometría del conjunto de soluciones del modelo de optimización para ciertos valores de los parámetros de la función de transferencia. Un espacio muestral métrico y no contable es típico en los modelos de optimización, ya que es el caso de un espacio  $\mathbb{R}^n$  que poseen la mayoría de estos modelos.

Al comienzo de este trabajo nos propusimos desarrollar una metodología que permita una evaluación objetiva del soporte empírico de modelos de optimización a través de una inferencia estadística y que posea un sustento matemático formal. El planteo de este objetivo fue motivado por la imposibilidad de encontrar una metodología que permitiese una evaluación empírica de modelos de optimización sin poseer las deficiencias que han sido señaladas para los métodos utilizados en la bibliografía (y que han sido tratados en el Capítulo 2). El desarrollo presentado en este trabajo muestra que, cumpliendo las condiciones antes mencionadas, es ciertamente posible una evaluación de las predicciones de los modelos de optimización que posee múltiples ventajas respecto a los métodos tradicionalmente utilizados. En efecto, la derivación de predicciones en forma de distribuciones de probabilidad permite realizar una inferencia estadística en el marco de la teoría de información o de la estadística bayesiana y obtener así un resultado objetivo respecto del soporte empírico en favor de una hipótesis de optimización.

Un aspecto a destacar del método propuesto es que permite una generalización de la hipótesis de optimización a evaluar. En lugar de considerar únicamente el o los estados óptimos, la generalización de un modelo de optimización hace posible que en base a una misma formulación se cubran múltiples modelos alternativos: desde una total ausencia de optimización, representada por el modelo nulo, a una optimización en sentido estricto, pasando por un espectro de casos intermedios que incluyen estados tradicionalmente denominados sub-óptimos. De esta forma, aún si los autores plantean en un principio una única hipótesis de optimización como modelo matemático a evaluar, si se cumplen las condiciones para definir un modelo general probabilístico está asegurada la evaluación de hipótesis alternativas. Más aún, el modelo nulo corresponde a una hipótesis nula respecto de la optimización, por lo que se cumple asimismo con un requerimiento indispensable para una correcta evaluación de la hipótesis de optimización en sí.

La relajación de supuestos que suelen estar asociados a las hipótesis de optimización es otro aspecto a destacar de la formulación propuesta. De hecho, la relajación del supuesto de monotonía estricta y la consideración de los estados sub-óptimos es,

precisamente, lo que hace posible la generalización del modelo de optimización y la derivación de un espacio de probabilidad. En lugar de considerar únicamente los estados óptimos, como sucede comúnmente al evaluar modelos de optimización, la introducción de una función de transferencia permite cubrir un abanico de casos particulares que se diferencian por cuán probable consideran que el sistema se encuentre en estados sub-óptimos de acuerdo al valor de la función objetivo.

Ward (1993) plantea que los investigadores impiden el progreso de la ciencia que practican si no están dispuestos a usar explicaciones por fuera del modelo prevaeciente en su disciplina. En el marco específico de la teoría del forrajeo, Ward (1992) apela a incluir modelos basados en el concepto de suficiencia, noción que apela a la inclusión de estados subóptimos y se plantea como alternativa a la optimización. En este sentido, creemos que la generalización de los modelos de optimización representa un puente entre los paradigmas de suficiencia vs. optimización (estricta). Dado que las explicaciones que apelan a la optimización suelen tener preponderancia en la biología en su conjunto cuando están basadas en la evolución por selección natural. En un marco más amplio, por lo tanto, la metodología propuesta puede facilitar a que explicaciones alternativas a la optimización estricta sean evaluadas en conjunto y de forma rutinaria, y podría contribuir así al avance de esta rama del conocimiento.

En la mayoría de las aplicaciones de modelos de optimización no resulta interesante saber si el estado observado del sistema se encuentra en una configuración óptima, ya que con seguridad este no será el caso aún si se da una optimización como consecuencia de limitantes en la determinación de los parámetros del sistema así como de las idealizaciones inherentes a todo modelo matemático. Por ejemplo, si el estado óptimo global resulta superior en un porcentaje muy bajo frente al valor medio, el conocer cuál es la configuración óptima puede brindar poca información respecto a la configuración del sistema real aún existiendo un proceso de optimización. Dada su formulación, la metodología expuesta posee la ventaja de permitir responder a preguntas formuladas en términos probabilísticos. El uso de una función de normalización cuantil en base al modelo nulo, en particular, permite definir estos estados en un contexto estadístico más allá del valor absoluto de la función objetivo, lo que puede resultar más provechoso para el estudio de un sistema de acuerdo a lo expuesto. Todo sistema natural se halla sujeto a variaciones estocásticas de diferentes tipos, por lo que puede suceder que la configuración óptima no se imponga si el proceso de optimización existe pero tiene una intensidad baja en relación dichas variaciones.

En función de lo expuesto, entendemos que la formulación presentada en este trabajo tiene múltiples aplicaciones potenciales en el estudio de la optimización en biología, y esperamos que su aplicación contribuya a ampliar el repertorio de respuestas posibles al evaluar un modelo de optimización.



## APLICACIÓN DE LA METODOLOGÍA GENERAL A MODELOS DE FORRAJEO ÓPTIMO

---

### Resumen

La teoría de forrajeo representa uno de los campos de la biología donde se ha trabajado con modelos de optimización con mayor énfasis. Un trabajo que tuvo una particular repercusión en este ámbito fue el de Belovsky (1986b), que evalúa la capacidad de modelos de optimización para predecir la composición de la dieta de 14 especies de herbívoros en el campo y muestra un notorio ajuste del modelo de maximización de la ingesta de energía. Este estudio ha sido criticado severamente por varios autores y los modelos planteados han sido incluso reanalizados con el fin de resaltar deficiencias en la evaluación empírica.

En este trabajo evaluamos una generalización de los modelos planteados por Belovsky (1986b) en la forma de modelos probabilísticos con el objetivo de realizar una primera aplicación de la metodología presentada en el Capítulo 3 y analizar cómo la evaluación empírica a través de una inferencia estadística puede afectar las conclusiones a las que se llega a partir de una misma hipótesis y conjunto de datos. Para cada especie de herbívoros se analizaron dos versiones de cada modelo que difieren en la distribución de probabilidad de los parámetros: una versión determinista donde tienen un valor fijo e idéntico al del trabajo original, y una versión no determinista que da cuenta de la variabilidad intrínseca y/o incertidumbre en la estimación de los parámetros. Se plantearon modelos probabilísticos de optimización mediante el uso de una función de transferencia escalonada y, en el caso no determinista, una función de normalización cuantil. Se utilizó un método de Monte Carlo por cadenas de Markov para realizar un muestreo aleatorio sobre el espacio de soluciones de los modelos para diferentes valores del parámetro  $q$  de la función de transferencia, y luego se estimó la distribución de probabilidad de la composición de dieta predicha mediante un método de suavizado por kernel gaussiano. Se utilizaron como métodos de inferencia estadística la prueba de cociente de verosimilitud, el criterio de información de Akaike y el factor de Bayes.

El análisis del modelo nulo reveló que la dieta de una de las especies resulta incompatible con las restricciones planteadas, un hecho que no es mencionado en el trabajo original ni en otras publicaciones que también analizan los modelos de este estudio, por lo que fue excluida de los análisis estadísticos. Los resultados de los métodos de inferencia estadística fueron similares entre sí e igualmente concluyentes en favor de la hipótesis de máxima ingesta de energía en concordancia con las conclusiones de Belovsky (1986b), tanto en la versión determinista de los modelos planteados en el trabajo original como en la versión no determinista. Al ser el producto de métodos estadísticos estándar, los resultados obtenidos en este trabajo permiten realizar una evaluación objetiva de los modelos de optimización que carece de las deficiencias que varios autores señalan respecto del trabajo original.

#### 4.1 INTRODUCCIÓN

Los estudios sobre la ecología del forrajeo tiene un lugar destacado entre las disciplinas que han aplicado modelos de optimización en biología. Con origen en los trabajos pioneros de MacArthur y Pianka (1966) y Emlen (1966), el énfasis en la aproximación de la optimización en el estudio del comportamiento de forrajeo fue tal que a mediados de la década de 1980 se había impuesto la denominación de teoría de forrajeo *óptimo*. A partir del trabajo clásico de Gould y Lewontin (1979), tal como se ha expuesto en el Capítulo 2, se sucedieron críticas al uso de modelos de optimización en la biología, y una de las áreas que sufrió con mayor intensidad dichos cuestionamientos fue la teoría de forrajeo *óptimo*. Ante las afirmaciones acerca del éxito de la teoría realizadas por algunos de los autores más destacados (Pyke y col. 1977; Krebs y col. 1978; Stephens y Krebs 1986, e.g.) se antepusieron ácidas críticas (Heinrich 1983; Pierce y Ollason 1987, e.g.). Como resultado de tal polémica, Perry y Pianka (1997) señalan que luego de la década del 1980 el uso del término de *teoría de forrajeo óptimo* fue discontinuado en favor del más general *teoría de forrajeo*, pero lejos de abandonarse los modelos de optimización pasaron a ser parte del marco teórico de la disciplina (ver Stephens y Krebs 1986).

Si bien el tenor de la polémica en torno a la optimización alcanzó uno de los puntos más álgidos en relación a la teoría de forrajeo *óptimo*, los aspectos criticados resultan comunes a los señalados en otras áreas de aplicación de la optimización en biología tal como se discutió en el Capítulo 2. Un trabajo que tuvo una particular repercusión en la polémica en torno al forrajeo *óptimo* fue el de Belovsky (1986b). Este estudio evalúa si la composición de la dieta de campo de 14 especies de herbívoros generalistas puede ser explicada por un mismo modelo de optimización, y representa la culminación de varios trabajos del autor en esta línea de investigación (Belovsky 1978; Belovsky y Jordan 1978; Belovsky 1981a, 1981b; Belovsky 1984a; Belovsky 1984b, 1984c, 1986a). De manera notable, los resultados de Belovsky (1986b) presentan una alta correspondencia cuantitativa entre las observaciones de campo y las predicciones del modelo de optimización en especies de insectos y mamíferos cuyo tamaño corporal varía en 7 órdenes de magnitud.

Belovsky (1986b) fue objeto de severas críticas respecto de la formulación del modelo y de metodología utilizada para evaluar las hipótesis de optimización (Hobbs 1990; Owen-Smith 1993; Huggard 1994; Owen-Smith 1996), lo que dio lugar a varias respuestas por parte de Belovsky (Belovsky 1990; Belovsky y Schmitz 1993; Belovsky 1994). En el trabajo original se evaluó a las hipótesis de optimización mediante un análisis de regresión de la composición de dieta observada en función de las predicciones obtenidas mediante un modelo de optimización. En particular Belovsky (1986b) utiliza una prueba de significación sobre la relación entre la dieta predicha y observada, y

obtiene como resultado un valor altamente elevado y significativo del coeficiente de determinación ( $R^2 = 0,984$ ,  $P < 0,001$ ). El autor concluye en base a este análisis que la hipótesis de máxima adquisición de energía efectivamente explica la composición de dieta en las especies y ambiente bajo estudio. Dado que toda la información necesaria para reproducir los resultados presentados está disponible en la publicación, esto permitió el reanálisis de los modelos planteados por parte de otros autores. Hobbs (1990) utiliza un análisis de Monte Carlo para evaluar la sensibilidad del modelo de optimización lineal, y concluye que los resultados de Belovsky (1986b) no son reproducibles debido a la gran sensibilidad a variaciones en el valor de los parámetros del modelo. Belovsky (1990) señala como respuesta a estas críticas que el análisis realizado por Hobbs utiliza el *desvío* estándar de los parámetros en lugar del *error* estándar, lo que a su entender es incorrecto ya que el objetivo del trabajo original es evaluar comportamientos promedio. Huggard (1994), por su parte, realiza un análisis de Monte Carlo más exhaustivo, y a pesar de que utiliza el error estándar sus conclusiones son similares a las de Hobbs (1990). La respuesta de Belovsky (1994) en este caso es señalar que Huggard (1994) excluye 2 especies que no presentan variabilidad en la predicción del modelo, y esto produce un sesgo en los resultados. Belovsky (1994) también señala que las estimaciones de variabilidad de los parámetros presentadas en el trabajo sobrestiman su varianza real, pero esto parece ser un argumento *ad hoc* ya que no había sido señalado con antelación ante similares críticas. Las críticas a Belovsky (1986b) no se circunscriben a cuestiones de índole estadística únicamente, pero fueron este tipo de cuestionamientos los que permitieron sostener afirmaciones como la siguiente, que cierra el trabajo de Huggard (1994, pp. 474): “La notoria capacidad de los modelos de programación lineal de predecir la dieta de manera precisa debería ser vista con escepticismo, y debería ser re-examinada cuidadosamente en estudios independientes”.

Las características mencionadas acerca de Belovsky (1986b) lo hacen adecuado para aplicar la metodología presentada en el Capítulo 3. Asimismo, y como es común a la mayoría de las publicaciones, en los modelos planteados se asume una optimización estricta, de manera que las predicciones corresponden a un óptimo global para cada especie que se obtiene mediante una técnica de programación lineal. De la misma manera, se asume una relación de monotonía estricta entre la aptitud biológica y la función objetivo que efectivamente se utiliza para derivar las predicciones: una relación inversa con el tiempo de forrajeo, o una relación directa con la energía adquirida. En consecuencia, estos modelos representan una oportunidad para analizar el efecto de su generalización. Los objetivos de este capítulo fueron realizar una primera aplicación de la metodología presentada en el Capítulo 3 a los modelos de forrajeo óptimo planteados por Belovsky (1986b) y analizar cómo pueden afectar las conclusiones que obtienen al evaluar una misma hipótesis de optimización con un mismo conjunto de datos.

## 4.2 METODOLOGÍA

Los modelos de optimización de Belovsky (1986b) incluyen restricciones relacionadas con la fisiología del sistema digestivo, y considera dos hipótesis de optimización alternativas: una minimización del tiempo de forrajeo y una maximización de la energía adquirida. La descripción de las variables y parámetros de ambos modelos de optimización se presenta en la Tabla 4.1. El modelo plantea la maximización de la función objetivo

$$f = a_1 \cdot g + a_2 \cdot d, \quad (4.1)$$

sujeta a las restricciones

$$\begin{aligned} D_c \cdot D_t &\geq b_1 \cdot g + b_2 \cdot d \\ T &\geq c_1 \cdot g + c_2 \cdot d, \end{aligned} \quad (4.2)$$

que expresan la máxima capacidad digestiva y el máximo tiempo de forrajeo respectivamente.

El modelo de mínimo tiempo de forrajeo, por su parte, está definido por la minimización de la función objetivo

$$f = c_1 \cdot g + c_2 \cdot d, \quad (4.3)$$

sujeta a las restricciones

$$\begin{aligned} D_c \cdot D_t &\geq b_1 \cdot g + b_2 \cdot d \\ E &\leq a_1 \cdot g + a_2 \cdot d, \end{aligned} \quad (4.4)$$

que expresan la máxima capacidad digestiva y el mínimo requerimiento energético respectivamente. Por último, la variable aleatoria observada es el porcentaje de monocotiledóneas en la dieta

$$X = \frac{g}{g + d}. \quad (4.5)$$

El valor de la totalidad de los parámetros está disponible en la Tabla 3 de Belovsky (1986b). Los valores de la variable aleatoria observada fueron obtenidos a partir del análisis de la Figura 3 de una versión digital del trabajo original mediante el software de procesamiento de imágenes digitales ImageJ (Schneider y col. 2012). El valor observado de porcentaje de monocotiledóneas en la dieta resultó el promedio de los valores cuantificados en la Figura 3a y 3b, que difirieron en todos los casos en menos del < 1,5 %.

Belovsky (1986b) fija el valor de los parámetros del modelo al promedio observado. Al no existir incertidumbre alguna respecto del valor de los parámetros, en consecuencia los modelos de optimización que plantea son de tipo determinista. En este trabajo consideramos además para cada modelo de optimización una versión no determinista, donde se asocia al espacio de parámetros ( $\Theta$ ) una distribución de probabilidad ( $P_\Theta$ )<sup>1</sup>, lo que fue posible dado que en el trabajo original se presenta el error estándar asociado a la estimación del promedio de la mayoría de los parámetros (Belovsky 1986b, Tabla 3). Para definir  $P_\Theta$  se utilizó el mismo criterio que en Huggard (1994), que corresponde a una distribución de probabilidad normal e independiente con parámetros  $\mu$  y  $\sigma$  correspondientes a la media muestral y error estándar, truncada a los valores que se alejan más de 2 unidades de error estándar. Según la especie de herbívoro en cuestión, el error estándar está disponible en 4 a 6 de los parámetros del modelo.

La combinación de 14 especies de herbívoros, 2 funciones objetivo alternativas y 2 distribuciones de probabilidad de los parámetros alternativas (versiones determinista y no determinista) da como resultado un total de 56 modelos probabilísticos de optimización analizados. Para construir los modelos probabilísticos se siguió lo expuesto en el Capítulo 3. A diferencia de lo que ocurre en los modelos planteados por Belovsky (1986b), para construir los modelos probabilísticos se utilizaron la totalidad de las 3 restricciones planteadas, tanto en el caso de la función objetivo de minimización del tiempo de forrajeo como de una maximización de la energía adquirida. En ambos modelos una de las restricciones se relaciona con la función objetivo y por lo tanto no influye en el valor del óptimo global, pero sí lo hace a la hora de establecer el espacio de soluciones candidatas. Como función de normalización utilizamos la normalización cuantil, Definimos la normalización cuantil como

$$n(\omega, \theta) = c_\theta(\omega) \cdot F_{f_\theta}(\omega) \quad , \quad (4.6)$$

y como función de transferencia a la función escalonada,

$$t(n(\omega, \theta), q) = \begin{cases} 1 & \text{si } n(\omega, \theta) \geq q, \\ 0 & \text{si } n(\omega, \theta) < q. \end{cases} \quad (4.7)$$

En consecuencia, el valor del parámetro  $q$  del modelo probabilístico de optimización representa el valor del cuantil de la distribución de la función objetivo en el modelo nulo para un cierto  $\theta$  a partir del que un estado del sistema compatible con las restricciones se considera perteneciente al espacio de soluciones óptimas.

El uso de la normalización cuantil requiere de la estimación de la distribución de la función objetivo  $F_{f_\theta}$  en el modelo nulo (Ecuación 3.6.1.2). La estimación de  $F_{f_\theta}$  se reali-

<sup>1</sup> El modelo determinista corresponde a un caso particular de un modelo no determinista donde  $P_\Theta$  corresponde a una medida de Dirac concentrada en un  $\theta$  dado.

Tabla 4.1: Descripción de los componentes de los modelos de optimización de maximización de la energía adquirida y minimización del tiempo de forrajeo.

Elemento	Símbolo	Unidades
Función objetivo		
Ingesta de energía asimilable	$f$	kJ / día
Tiempo de forrajeo	$f$	min / día
Variables ( $\Omega$ )		
Ingesta de monocotiledóneas	$g$	g PS / día
Ingesta de dicotiledóneas	$d$	g PS / día
Parámetros ( $\Theta$ )		
Contenido de energía asimilable de monocotiledóneas	$a_1$	kJ / g PS
Contenido de energía asimilable de dicotiledóneas	$a_2$	kJ / g PS
Capacidad del órgano digestivo	$D_c$	g PH
Velocidad de renovación del órgano digestivo	$D_t$	día <sup>-1</sup>
Conversión peso seco a peso húmedo monocotiledóneas	$b_1$	g PH / g PS
Conversión peso seco a peso húmedo dicotiledóneas	$b_2$	g PH / g PS
Máximo tiempo de forrajeo	$T$	min / día
Velocidad de pastoreo monocotiledóneas	$c_1$	g PS / día
Velocidad de pastoreo dicotiledóneas	$c_2$	g PS / día

Nota. PS = peso seco ; PH = peso húmedo.

zó a partir de un muestreo aleatorio sobre el espacio de soluciones del modelo nulo mediante un algoritmo de Metropolis Hasting implementado por Van den Meersche y col. (2009). Este algoritmo es un tipo particular de método de Monte Carlo por cadenas de Markov y resulta adecuado para un espacio muestral euclideano y restricciones de tipo lineal.

Dado que la función objetivo se trata de una función lineal, fue posible utilizar el algoritmo de Metropolis Hasting ya descrito para obtener muestras del espacio de soluciones para diferentes valores del parámetro de posición  $q$  de la función de transferencia. Para obtener muestras del espacio de soluciones  $\Omega \times \Theta \times q$  se incorporó al modelo nulo una nueva restricción a las ya existentes. En el caso de la maximización de energía ingerida se incorporó la restricción

$$q \leq F_{\theta} a_1 \cdot g + a_2 \cdot d , \quad (4.8)$$

mientras que en el caso de la minimización del tiempo de forrajeo se incorporó la restricción

$$1 - q \leq F_{f_{\theta}} c_1 \cdot g + c_2 \cdot d . \quad (4.9)$$

La estimación numérica por los métodos antes descriptos corresponde al caso de un cierto elemento  $\theta$  del espacio de parámetros  $\Theta$ . En el caso del modelo determinista, el espacio de parámetros consiste en un único elemento  $\theta$  que, como ya se dijo, corresponde al valor promedio de los parámetros del modelo. En cambio, en el modelo no determinista existe una distribución de probabilidad  $P_{\Theta}$  definida sobre un espacio de parámetros  $\Theta$ , que consiste en un espacio  $\mathbb{R}^4$  a  $\mathbb{R}^6$  dependiendo de la especie, y por lo tanto contiene un número no contable de elementos. Por lo tanto, la estimación del modelo no determinista se realizó tomando  $10^3$  muestras al azar de acuerdo a  $P_{\Theta}$ .  $F_{f_{\theta}}$  fue estimada como la distribución acumulada empírica calculada sobre un total de  $10^7$  muestras del espacio de soluciones del modelo nulo, lo que significa que en el modelo determinista se realizó una única vez mientras que en el no determinista se hace de forma independiente  $10^3$  veces, una para cada  $\theta$ . Para estimar la distribución de las variables en el espacio de soluciones de cada modelo de optimización se utilizó una secuencia de valores de  $q$  entre 0 y 0,99 con incrementos de 0,01, sumados los valores de 0,999 y 0,9999. Para cada valor de  $q$  se obtuvieron  $10^6$  en el modelo determinista, y  $10^4$  muestras por cada uno de los  $10^3$  elementos  $\theta$  del espacio de parámetros.

La muestra del espacio de soluciones para cada valor de  $q$  se utilizó para estimar la densidad de probabilidad de la variable aleatoria observada  $X$  (Ecuación 4.5) y de esta manera permitir la aplicación de una inferencia estadística. Se utilizó a tal fin un método de suavizado por kernel gaussiano propuesto por Botev y col. (2010) que realiza un ajuste no paramétrico y permite estimar una distribución de probabilidad sin realizar suposiciones sobre la naturaleza de la distribución subyacente. Este método incorpora una selección óptima de la ventana de datos analizados y un parámetro crítico del ajuste por kernel.

Finalmente, a fin de ejemplificar las alternativas metodológicas disponibles, se realizó una inferencia estadística mediante la prueba de cociente de verosimilitud, el criterio de información de Akaike y el factor de Bayes. Para cada tipo de método se realizó una inferencia de forma independiente para cada una de las especies, y también para el conjunto de las especies, considerando tanto un mismo valor de  $q$  para cada especie como valores independientes de  $q$ . En el análisis bayesiano de todos los modelos de optimización se utilizó como distribución de probabilidad previa de  $q$  una distribución uniforme en el intervalo  $(0, 1]$ .

La implementación de los análisis descriptos fue realizada en MATLAB (Inc. 2011), a excepción del algoritmo de Metropolis Hasting utilizado para obtener muestras del espacio de soluciones, que forma parte del paquete *limSolve* (Soetaert y col. 2009) para

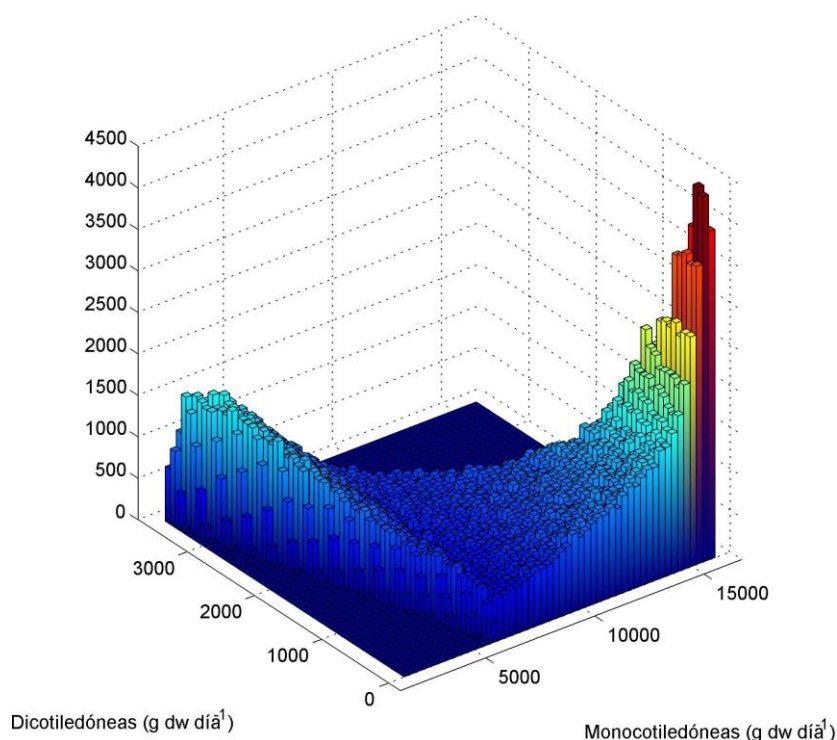


Figura 4.1: Distribución de la composición de la dieta esperada de *Bison bison* en base al modelo nulo de acuerdo a las restricciones planteadas por Belovsky (1986b). El color indica abundancia de las muestras que caen en cada intervalo con el fin de facilitar la visualización del histograma.

el software GNU R (Team 2011). El paquete R.matlab (Bengtsson 2013) se usó como interfase entre MATLAB y GNU R. Se utilizó la implementación del método de suavizado por kernel gaussiano disponible como una función de MATLAB (Botev y col. 2010).

La obtención de los resultados a través de los métodos descritos demanda un elevado poder de cómputo. En el caso no determinista, para cada modelo se debieron obtener un total de  $1,102 \cdot 10^{10}$  muestras de espacios de soluciones para cada modelo de optimización: correspondientes a  $10^7$  muestras del modelo nulo y  $10^4$  muestras por cada uno de 102 valores de  $q$  por cada uno de los  $10^3$  valores de  $\theta$ . Por lo tanto, para acelerar el cómputo se realizó una implementación de los algoritmos que presenta una paralelización en varios puntos, tanto a nivel de los diferentes  $\theta$  correspondientes a un modelo no determinista como a nivel de los valores de  $q$ . Los análisis se corrieron en un servidor con doble procesador Xeon E5-2650 de 12 núcleos y 24 subprocesos y 256 Gb de memoria RAM, lo que permitió mantener tiempos de procesamientos razonables. El código completo utilizado para reproducir los resultados presentados en este capítulo se incluye como apéndice (Apéndice A).

### 4.3 RESULTADOS

En este trabajo realizamos una primera aplicación de la metodología para la generalización y evaluación de modelos de optimización a los planteados por Belovsky (1986b). Para la estimación de la función de verosimilitud se utilizó un método de Monte Carlo para obtener muestras del espacio muestral con una densidad de Radon-Nikodym uniforme y de esta forma estimar la distribución de probabilidad de cualquier variable observable de la que se tenga información. En los modelos aquí analizados, el espacio muestral consiste en la biomasa de plantas angiospermas monocotiledóneas y dicotiledóneas que las especies de herbívoros ingieren diariamente. A modo de ejemplo, en la Figura 4.1 se presenta la distribución de la composición de la dieta de *Bison bison* esperada de acuerdo al modelo nulo, que corresponde a un valor del parámetro  $q$  de la función de transferencia igual a 0.

Dada la metodología utilizada por Belovsky (1986b) en el trabajo de campo, la variable observada con la que se evalúan empíricamente los modelos es el cociente entre la ingesta diaria de monocotiledóneas y dicotiledóneas. La estimación de la densidad de Radon-Nikodym del porcentaje de monocotiledóneas en la dieta para cada especie predicha por los modelos deterministas se presenta en la Figura 4.2. El orden en que se presentan los resultados para cada especie sigue el criterio utilizado por Belovsky (1986b), y corresponde a un tamaño corporal creciente. A diferencia del modelo nulo, los modelos de máxima ingesta de energía y de minimización del tiempo de forrajeo pueden tener un valor parámetro  $q$  en el intervalo  $(0, 1]$ , por lo que se eligieron valores de  $q = 0,5$  y  $q = 0,95$  para presentar gráficamente la densidad de probabilidad de ambos modelos. En primer lugar, cabe destacar que el análisis del modelo nulo revela que para la especie *Dissosteira carolina* la densidad de probabilidad es nula para la composición de la dieta observada en campo, lo que significa que el sistema observado *no* respeta las restricciones planteadas en el modelo. Al ser incompatible con las restricciones, los modelos de optimización planteados predicen necesariamente una densidad nula para el valor observado de la variable. En consecuencia, la especie *D. carolina* se excluyó de los análisis subsecuentes.

Dado que la variable aleatoria observada difiere de la función objetivo en ambos modelos de optimización, de acuerdo a lo tratado en la Sección 3.7.4.1 no es posible aplicar una prueba de significación por hipótesis nula, y necesariamente los métodos de inferencia estadística deben estar basados en la función de verosimilitud. Al evaluar la verosimilitud en función del parámetro  $q$  para la versión determinista de los modelos de optimización, se observó que esta presenta un máximo en torno a valores de  $q > 0,9$  para el modelo de máxima ingesta de energía en la mayoría de las especies (Figura 4.3). Belovsky (1986b) presenta solo un valor de la composición de la dieta por especie correspondiente estimado en el sitio de estudio. Sin embargo, la metodología aplicada

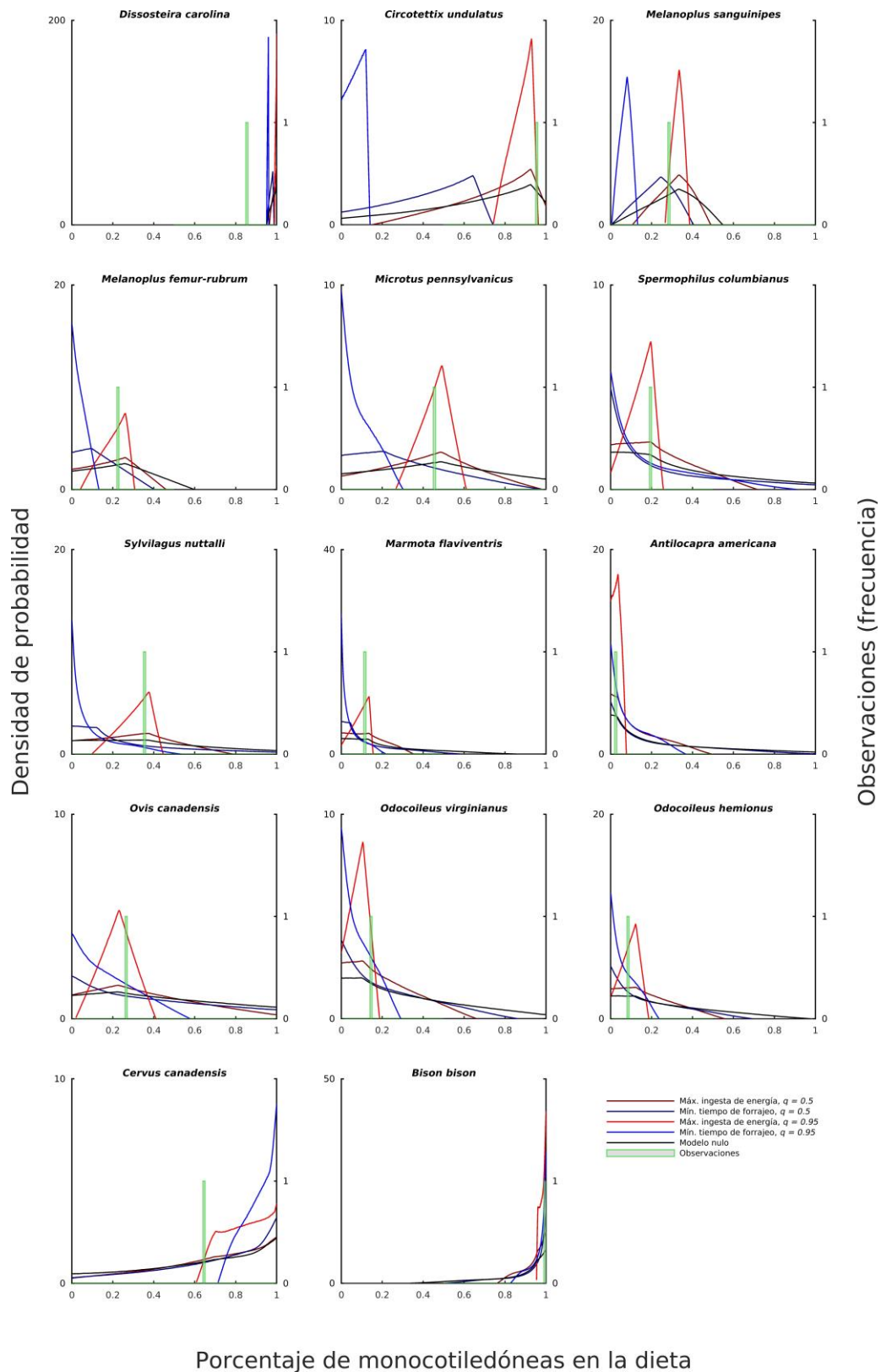


Figura 4.2: Densidad de Radon-Nikodym de la composición de dieta predicha por la versión determinista del modelo nulo y de los modelos de optimización de máxima energía ingerida y mínimo tiempo de forrajeo para valores de  $q$  de 0,5 y 0,95.

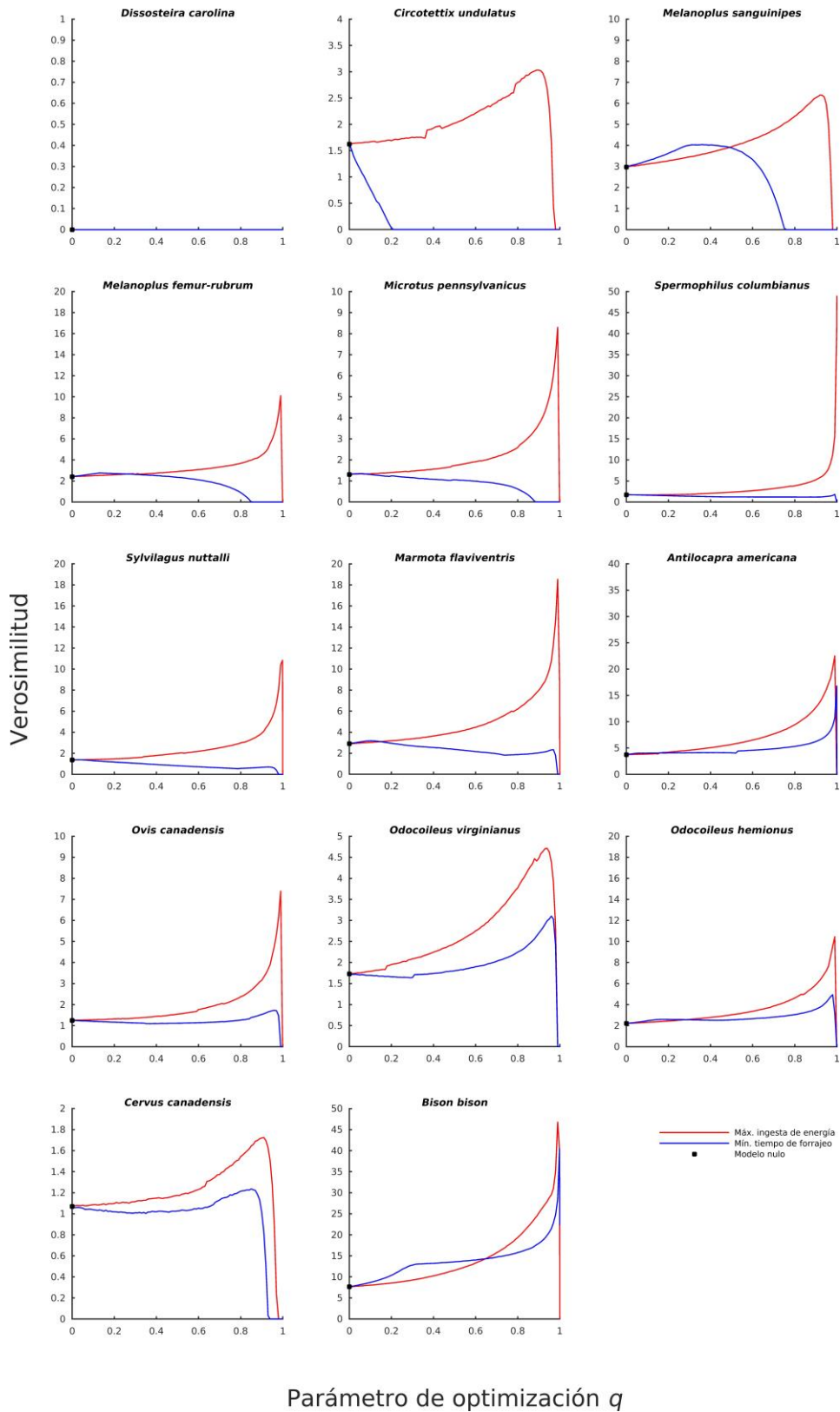


Figura 4.3: Verosimilitud de los modelos de forrajeo óptimo para cada especie analizada en función del parámetro  $q$  de la versión determinista de los modelos, donde los parámetros son constantes y tienen un valor igual al promedio de la estimación de campo. Nótese que la escala del eje vertical varía en cada gráfico.

Tabla 4.2: Resultados de los métodos de inferencia estadística aplicados a la versión determinista del modelo de máxima ingesta de energía para cada especie de herbívoro de forma independiente. Para cada método de inferencia se analiza si el resultado cumple con criterios comúnmente utilizados para establecer si existe evidencia en favor del modelo evaluado.

Especie	Máx.L	Test de cociente de verosimilitud		Criterio de información de Akaike		Factor de Bayes	
	$q$	$P$	$P < 0,05$	$\Delta AIC_{nulo}$	$\Delta AIC_{nulo} > 2$	$B$	$B > 3$
<i>Circotettix undulatus</i>	>0.99	0.264	×	0.75	×	1.26	×
<i>Melanoplus sanguinipes</i>	0.92	0.216	×	0.47	×	1.37	×
<i>Melanoplus femur-rubrum</i>	0.99	0.090	×	-0.87	×	1.36	×
<i>Microtus pennsylvanicus</i>	0.99	0.055	×	-1.70	×	1.65	×
<i>Spermophilus columbianus</i>	>0.99	0.010	C	-4.71	C	1.91	×
<i>Sylvilagus nuttalli</i>	>0.99	0.042	C	-2.13	C	1.79	×
<i>Marmota flaviventris</i>	0.99	0.054	×	-1.71	×	1.69	×
<i>Antilocapra americana</i>	0.99	0.058	×	-1.58	×	1.88	×
<i>Ovis canadensis</i>	0.99	0.059	×	-1.56	×	1.56	×
<i>Odocoileus virginianus</i>	0.94	0.157	×	0.00	×	1.56	×
<i>Odocoileus hemionus</i>	0.99	0.078	×	-1.10	×	1.63	×
<i>Cervus canadensis</i>	0.91	0.329	×	1.05	×	1.13	×
<i>Bison bison</i>	0.99	0.057	×	-1.62	×	1.86	×

Nota. L = verosimilitud;  $\Delta AIC_{nulo}$  = diferencia en el criterio de Akaike con el modelo nulo;  $B$  = factor de Bayes.

hace posible la evaluación de un modelo de optimización aún si se cuenta con un único dato de la variable observada. Por lo tanto, con un fin ilustrativo se presentan los resultados de los métodos de inferencia aplicados a la versión determinista del modelo de maximización de energía para cada especie por separado en la Tabla 4.2. Solo en el caso de *Spermophilus columbianus* y *Sylvilagus nuttalli* la prueba de verosimilitud y el criterio de información de Akaike dan soporte al modelo de optimización; el análisis bayesiano mediante el cómputo del factor de Bayes no brinda soporte al modelo en ninguna de las especies.

Al evaluar la capacidad de los modelos deterministas para predecir la composición de la dieta para las 13 especies que cumplen con las restricciones en su conjunto, los resultados mostraron soporte en favor de la hipótesis de máxima adquisición de energía y un rechazo tanto de la hipótesis de minimización del tiempo de forrajeo como del modelo nulo. La relación entre la verosimilitud de los modelos en función del parámetro  $q$ , considerando un mismo valor para todas las especies, se presenta en la

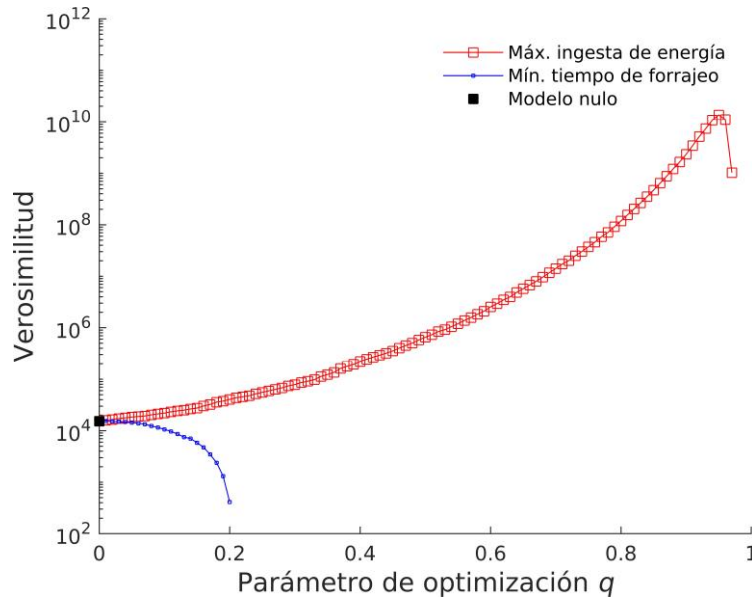


Figura 4.4: La variación de la verosimilitud de los modelos de forrajeo óptimo en función del parámetro  $q$  con un mismo valor para el conjunto de las especies muestra un fuerte soporte a la hipótesis de máxima energía ingerida. Nótese que el eje de las ordenadas posee una escala logarítmica.

Figura 4.4. Se observa que la verosimilitud del modelo de máxima ingesta de energía aumenta respecto del modelo nulo con valores crecientes de  $q$  hasta alcanzar un máximo en torno a  $q = 0,95$ . En contraste, el ajuste del modelo de minimización del tiempo de forrajeo es peor que el del modelo nulo para todos los valores de  $q$ . De hecho, las predicciones se alejan de los valores observados a medida que  $q$  aumenta, y para valores de  $q > 0,2$  la verosimilitud se hace nula debido a que la distribución de probabilidad predicha excluye a la composición de dieta observada. Los resultados de los métodos de inferencia estadística fueron en todos los casos concluyentes en favor del modelo de maximización ingesta de energía (Tabla 4.3). La prueba de cociente de verosimilitud resultó altamente significativo en favor de dicha hipótesis: se obtuvo un  $P < 10^{-6}$ , cuando el criterio de significación para rechazar una hipótesis nula usado comúnmente en biología es un valor de  $P$  menor a 0,05, mientras que para el modelo de mínimo tiempo de forrajeo se obtuvo un  $P$  cercano a la unidad. El criterio de información de Akaike, por su parte, arroja un peso del modelo de máxima ingesta de energía  $> 0,99$ , descartando los modelos alternativos. Por último, el factor de Bayes de este modelo se ubicó en el rango que es considerado de soporte decisivo en favor de la hipótesis de maximización de energía: la verosimilitud marginal de este modelo fue de casi cuarenta mil veces la del modelo nulo, y de un valor todavía mayor con respecto al modelo de mínimo tiempo de forrajeo ( $\log_{10} B = 5,49$ ). En síntesis, los resultados de la aplicación de la metodología de generalización y evaluación a los modelos de forra-

jeo óptimo son coherentes con los obtenidos por Belovsky (1986b) cuando se utilizan como parámetros los valores utilizados en el trabajo original, a excepción de la especie *D. carolina*.

Ahora bien, el desarrollo matemático presentado en el Capítulo 3 permite la evaluación de modelos donde los parámetros no poseen un valor constante, sino que se caracterizan en cambio por una cierta distribución de probabilidad que representa la información que se dispone sobre dichos parámetros. Dado que parte de las críticas al trabajo de Belovsky (1986b) se centraban en el uso de un valor promedio de los parámetros en lugar de tener en cuenta que dicha estimación presenta variabilidad, analizamos además una versión no determinista de los modelos. En esta versión los parámetros de los modelos poseen una distribución de probabilidad que representa la incertidumbre en torno al valor de la media muestral. La estimación de la densidad de Radon-Nikodym del porcentaje de monocotiledóneas en la dieta predicha por los modelos no deterministas para cada especie se presenta en la Figura 4.5.

Al evaluar la verosimilitud en función del parámetro  $q$  para la versión no determinista de los modelos de optimización se observaron diferencias cualitativas respecto de la versión determinista (Figura 4.6). Si bien para la mayoría de las especies la verosimilitud del modelo de máxima ingesta de energía aumenta con el valor de  $q$  y es superior a la de los modelos alternativos, en algunos casos sucede lo contrario. En particular, la especie para la que el resultado del modelo no determinista presenta un mayor contraste es *D. carolina*. En la versión no determinista existen combinaciones del valor de los parámetros para los que la dieta observada resulta compatible con las restricciones planteadas. En esta especie se observa que la verosimilitud del modelo de máxima ingesta de energía disminuye respecto del modelo nulo con valores crecientes de  $q$  hasta hacerse nula para valores superiores a 0,6 aprox., mientras que para el modelo de minimización de tiempo de forrajeo sucede lo contrario.

Dado que *D. carolina* fue excluida de la inferencia estadística basada en los modelos deterministas, en los análisis del modelo no determinista fue también excluida de manera de que ambos análisis fuesen comparables. En la Figura 4.7 se presenta la relación entre la verosimilitud de los modelos no deterministas en función del parámetro  $q$  para el conjunto de las especies, considerando un mismo valor para todas las especies y excluyendo a *D. carolina*). Al igual que en el caso del modelo determinista, se observa un aumento sostenido de la verosimilitud con un aumento de  $q$  hasta un máximo en torno a un valor de  $q = 0,99$  para el modelo de maximización de la energía ingerida, mientras que la verosimilitud del modelo de minimización de tiempo de forrajeo es menor a la del modelo nulo y disminuye hasta hacerse nula para valores de  $q$  mayores a 0,3. Al aplicar métodos de inferencia estadística sobre la versión no determinista de los modelos se obtuvieron resultados similares a la versión versión determinista y en soporte del modelo de maximización ingesta de energía (Tabla 4.3). Si bien en los tres métodos aplicados se observa una ligera disminución de la evidencia en favor de

Tabla 4.3: Resultados de los métodos de inferencia aplicados al reanálisis de modelos de forrajeo óptimo para el conjunto de las especies de herbívoros que cumplen con las restricciones planteadas por Belovsky (1986b).

Modelo	$k$	Máxima verosimilitud		Test cociente de verosimilitud			Criterio de información de Akaike			Factor de Bayes	
		$q$	ln L	$t$	$P$	$AIC$	$\Delta AIC$	$w$	ln L marginal	$\log_{10} B$	
Versión determinista											
Modelo nulo	0	0	9.6	-	-	-19.3	25.4	<0.01	9.6	0	
Mín. tiempo de forrajeo	1	0.01	9.6	0.99	0.89	-17.3	27.4	<0.01	7.6	-0.89	
Máx. ingesta energía	1	0.95	23.3	$1,13 \cdot 10^{-6}$	$1,7 \cdot 10^{-6}$	-44.7	0	>0.99	20.2	4.60	
Versión no determinista											
Modelo nulo	0	0	6.6	-	-	-13.2	20.1	<0.01	6.6	0	
Mín. tiempo de forrajeo	1	0	6.6	>0.99	>0.99	-11.2	22.2	<0.01	4.6	-0.91	
Máx. ingesta energía	1	0.99	17.7	$1,56 \cdot 10^{-5}$	$2,5 \cdot 10^{-6}$	-33.4	0	>0.99	14.2	3.30	

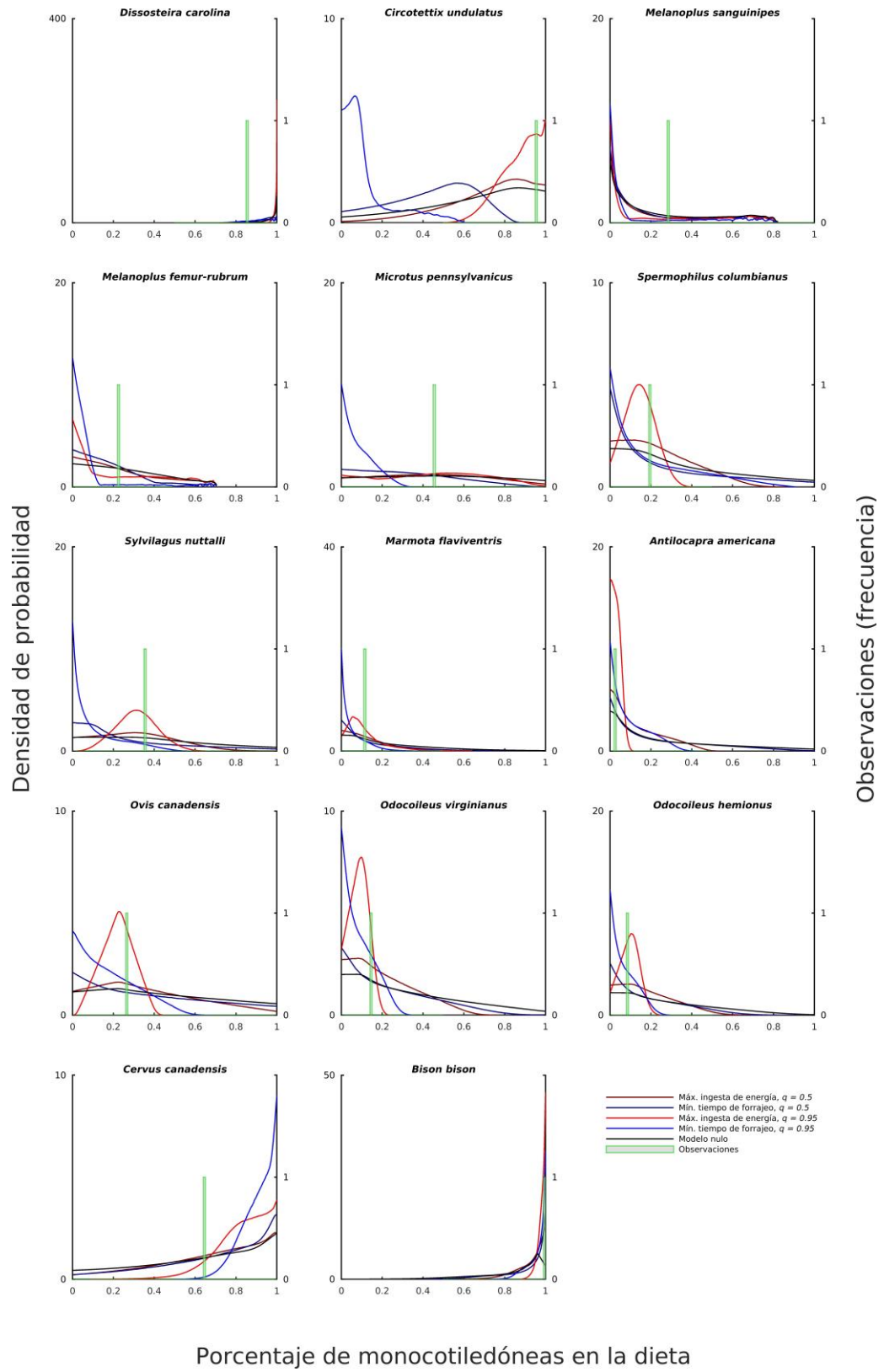


Figura 4.5: Densidad de Radon-Nikodym de la composición de dieta de la versión no determinista del modelo nulo y los modelos de forrajeo óptimo para valores de  $q$  de 0,5 y 0,95.

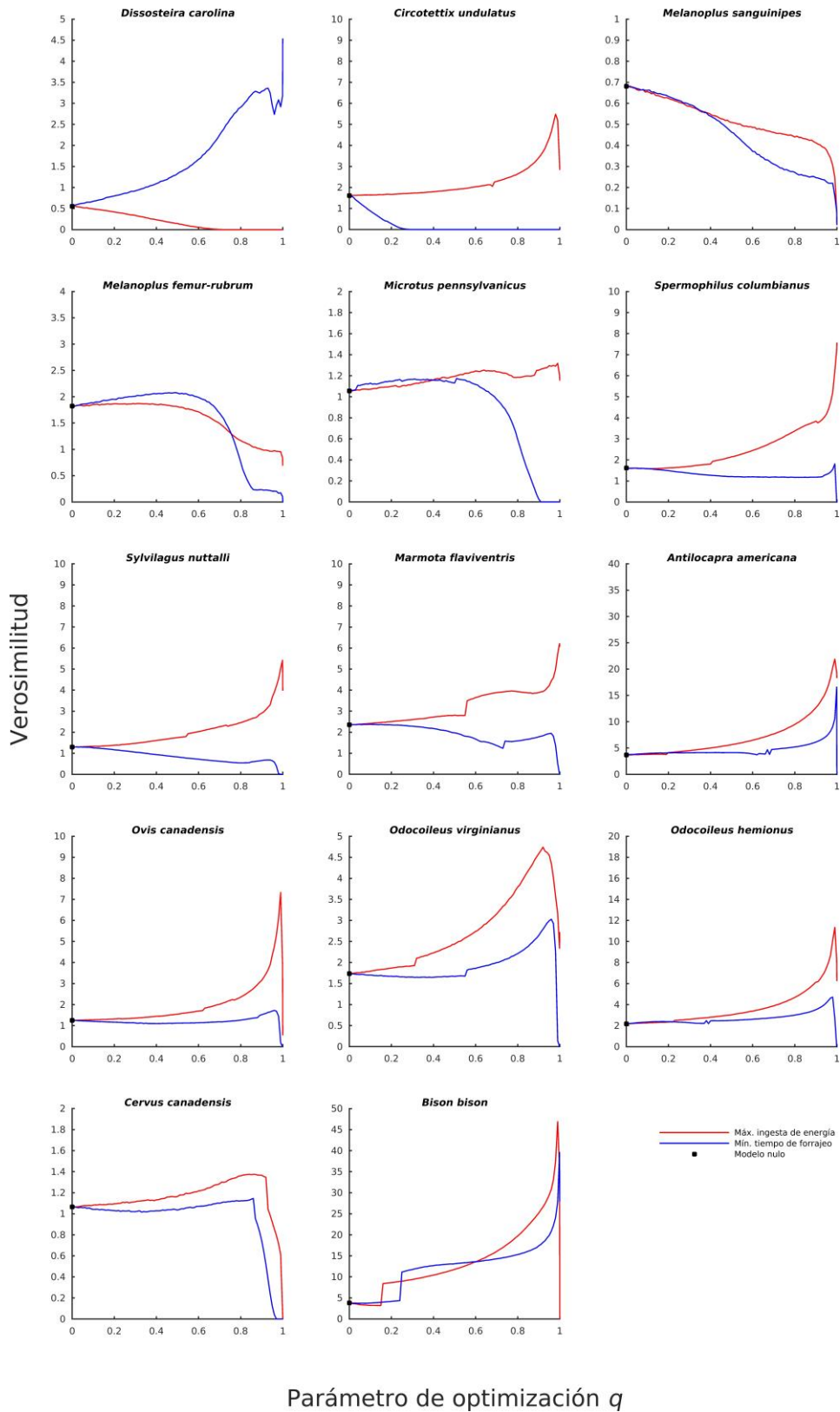


Figura 4.6: Verosimilitud de los modelos de forrajeo óptimo para cada especie analizada en función del parámetro  $q$  de la versión no determinista de los modelos, donde los parámetros poseen una distribución de probabilidad estimada a partir de la información disponible. Nótese que la escala del eje vertical varía en cada gráfico.

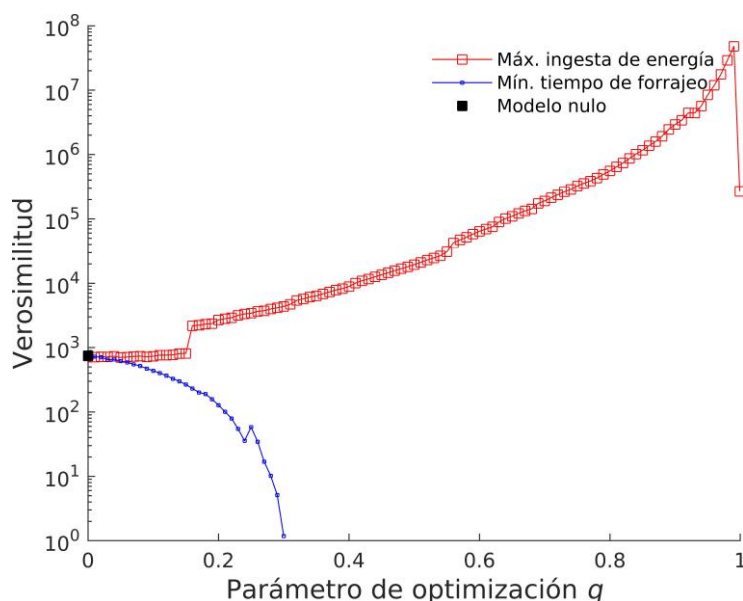


Figura 4.7: La variación de la verosimilitud de la versión no determinista de los modelos de forrajeo óptimo en función del parámetro  $q$  con un mismo valor también muestra un fuerte soporte a la hipótesis de máxima energía ingerida para el mismo grupo de especies analizado en la versión determinista. Nótese que el eje de las ordenadas posee una escala logarítmica.

dicho modelo, los resultados son igualmente concluyentes, por lo que esta pequeña diferencia no tiene un efecto cualitativo.

Por último, a fines comparativos se computó asimismo la verosimilitud de cada modelo para el conjunto de las especies, de manera de evaluar si la inclusión de *D. carolina* podría afectar de forma cualitativa las conclusiones. En la Figura 4.8 se observa que en este caso la verosimilitud del modelo de maximización de la energía ingerida es mayor que la del modelo nulo, pero alcanza un máximo de menor magnitud en torno a  $q = 0,6$ , y se hace nula a valores superiores a  $0,7$ . La aplicación de los métodos de inferencia, sin embargo, muestra igualmente apoyo al modelo de maximización de la ingesta de energía, pero los resultados no son en este caso concluyentes: en la prueba de verosimilitud se obtiene un  $P = 0,037$ , el peso de Akaike del modelo es de  $0,70$ , y el factor de Bayes resulta de  $2,93$  con respecto al modelo nulo.

#### 4.4 DISCUSIÓN

En este trabajo hemos realizado una primera aplicación de la metodología para realizar una inferencia sobre hipótesis de optimización, que en este caso intentan explicar la composición de la dieta de herbívoros en el contexto de la teoría de forrajeo. Se analizaron dos modelos que variaron en la distribución de probabilidad de los parámetros

$\Theta$  de los modelos de optimización, aquellos que se hipotetiza *no* se encuentran sujetos a una optimización (Tabla 4.1). En primer lugar discutimos los modelos de optimización determinista, donde cada parámetro es fijado al valor de la media muestral de las estimaciones experimentales, para luego analizar los resultados del caso más general de los modelos no deterministas.

La aplicación del método de inferencia a los modelos de optimización deterministas, que corresponden a los formulados originalmente por Belovsky (1986b), produjo resultados coherentes a los del trabajo original para 13 de las 14 especies analizadas. En la especie *Dissosteira carolina*, en contraste, el análisis del modelo nulo mostró que la densidad de probabilidad de la dieta observada en campo es nula. Este hecho no es mencionado por Belovsky (1986b), pero es notorio que tampoco sea señalado por Hobbs (1990) ni Huggard (1994), quienes reanalizan los modelos de optimización mediante simulaciones de Monte Carlo y realizan severas críticas al trabajo original. Este hecho puede ser explicado por el método utilizado para evaluar el ajuste de los modelos en todos los casos, que consiste en una regresión lineal sobre valores de dieta observados y predichos por los modelos de optimización, un tipo de análisis que posee una serie de deficiencias que fueron señaladas en el Capítulo 2. Este ejemplo de aplicación permite ilustrar una limitación de esta metodología de evaluación de modelos de optimización: cuando la variable aleatoria observada no coincide con el espacio muestral  $\Omega$  no es posible evaluar si es compatible con las restricciones. Esta deficiencia se relaciona con la incapacidad de estimar una distribución de probabilidad asociada a la variable aleatoria observada. Si bien esta cuestión ha sido abordada por varios autores, no hemos encontrado descrito en la literatura que uno de los riesgos de esta metodología consiste en concluir que un modelo de optimización posee sustento empírico cuando, en realidad, los datos *no* son compatibles con las restricciones planteadas y por lo tanto el modelo debería ser descartado. En suma, este resultado pone de manifiesto la ventaja de aplicar la metodología presentada en el Capítulo 3 en relación a otras metodologías tradicionalmente utilizadas.

En la versión no determinista de los modelos se utilizó como  $P_{\Theta}$  una distribución de probabilidad independiente para cada parámetro y correspondiente a una distribución normal con parámetros  $\mu$  y  $\sigma$  correspondientes a la media muestral y error estándar de las estimaciones experimentales respectivamente de acuerdo a los datos de Belovsky (1986b). Como ya se mencionó, Belovsky (1990) plantea que los modelos analizados fueron formulados con el objetivo de responder preguntas en términos del promedio de los parámetros, y que por lo tanto deben ser parametrizados con la media muestral y el error estándar como medida de precisión. Sin embargo, con posterioridad Belovsky (1994) plantea que el error estándar presentado en Belovsky (1986b) representa una sobreestimación debido a que se había asumido una distribución normal, mientras que la distribución observada se asemejaba en cambio a una distribución log-normal o exponencial negativa. Entendemos que esto último resulta

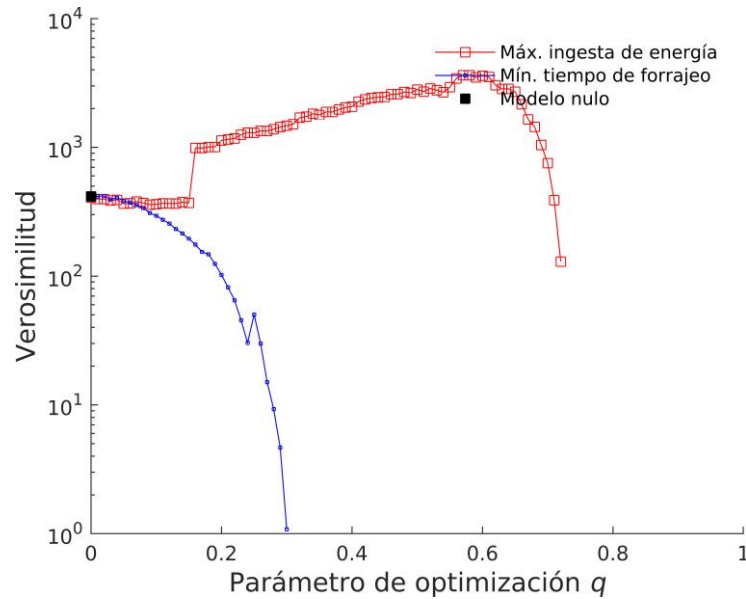


Figura 4.8: Verosimilitud de la versión no determinista de los modelos de forrajeo óptimo en función del parámetro  $q$  para la totalidad de las especies incluida *Dissosteira carolina*, para la que la composición observada de dieta no es compatible con la versión determinista del modelo nulo. Nótese que el eje de las ordenadas posee una escala logarítmica.

errado en virtud del teorema del límite central, más allá de cuál sea efectivamente la distribución de los parámetros cuantificados en campo o laboratorio. El teorema del límite central establece que la distribución de la media muestral tiende a una distribución normal con un desvío estándar igual al error estándar de la muestra para *cualquier* distribución. De acuerdo a los resultados tanto de Belovsky (1986b) como de Huggard (1994), la tasa de pastoreo es el parámetro para el que los resultados de los modelos muestran una mayor sensibilidad. El tamaño muestral de este parámetro es para todas las especies lo suficientemente grande como para asegurar que la distribución normal resulte una buena aproximación de la distribución de la media muestral. De hecho, en la mayoría de los parámetros para los que Belovsky (1986b, p. 44) presenta un error estándar asociado poseen tamaños muestrales elevados. En consecuencia, consideramos que el modelo no determinista consiste en una generalización adecuada del modelo determinista original.

La implementación del código realizada contempla el uso de funciones específicas para cada problema de optimización; las utilizadas en el presente análisis se presentan en la Sección A.1.3. En la implementación general (Sección A.1.1) se contempla la posibilidad de utilizar diferentes y múltiples variables para cada modelo de optimización, así como diferente número de datos, con el único supuesto de que todas las observaciones son independientes.

Los análisis se realizaron tanto para una de las especies de herbívoro por separado así como para el conjunto de éstas, lo que permitió comparar la evidencia en favor de las hipótesis de optimización que aporta cada caso con la obtenida al integrar toda la información disponible. En comparación con los resultados de la publicación original, los obtenidos en este trabajo presentan una mayor generalidad ya que son el resultado de métodos de inferencia estadística estándar. Se espera que el presente trabajo sirva como una referencia para ejemplificar cómo la aplicación de la metodología desarrollada al análisis de un modelo de optimización particulares permite realizar una inferencia estadística.



## DISCUSIÓN GENERAL

---

Al comienzo de este trabajo de tesis nos preguntamos si las críticas realizadas a los modelos de optimización en biología representan verdaderos impedimentos para el avance del conocimiento a través de esta metodología, o si se trata en cambio de asuntos con una importancia únicamente histórica y sin relevancia en la actualidad. A partir del trabajo de revisión bibliográfica presentado en el Capítulo 2 podemos concluir que evaluar el poder predictivo de las hipótesis de optimización de una forma objetiva es efectivamente relevante por múltiples razones. En primer lugar, parece ser la única manera de establecer si las múltiples idealizaciones utilizadas en un modelo de optimización permiten igualmente aportar información sobre el sistema. En segundo lugar, puede aportar información sobre cuáles son las restricciones que afectan el proceso evolutivo del sistema bajo estudio. Por último, permite establecer el sustento empírico que posee un determinado principio de optimización. Esto es sobreentendido en casos como el de la selección natural, dado que en este concepto se basa la teoría de la evolución de Darwin y representa una base teórica para el conjunto de la biología. Existen, sin embargo, principios de optimización de índole termodinámica que descansan en postulados teóricos pero poseen un escaso sustento empírico. Establecer pautas para un contraste riguroso de las hipótesis de optimización, en vista de lo expuesto, podría permitir un avance del conocimiento en esta área. Se espera que este intento de síntesis sobre los cuestionamientos a la aplicación de modelos de optimización en biología pueda servir como referencia a la hora de diseñar y evaluar trabajos que tengan el objetivo de analizar el sustento empírico de estos modelos. El segundo interrogante planteado como disparador del trabajo de revisión del Capítulo 2 fue si la popularidad actual de los modelos de optimización en biología es un indicativo de que se ha dado una respuesta satisfactoria a las diversas críticas realizadas a este abordaje teórico-metodológico, o si, por el contrario, éstas han sido simplemente ignoradas. En base al análisis realizado podemos concluir que la situación parece estar más cerca de la segunda opción que la primera. Si bien los trabajos actuales muestran un mayor cautela al afirmar que los sistemas bajo estudio son efectivamente óptimos, las carencias a nivel de la metodología utilizada para evaluar las hipótesis de optimización es notoria incluso en un campo de amplio desarrollo como el análisis del metabolismo celular. Esta situación sugiere que esta área del conocimiento se podría beneficiar de una metodología que no tenga las limitaciones que han sido señaladas.

La metodología presentada en el Capítulo 3 permite una evaluación de modelos de optimización que carece de las limitaciones de los métodos tradicionalmente utiliza-

dos. La metodología propuesta tiene como punto de partida una generalización de los modelos de optimización, lo que hace posible su expresión en la forma de un modelo probabilístico bajo ciertas condiciones generales y, en consecuencia, permite analizar de forma objetiva el soporte empírico de un modelo de optimización mediante una inferencia estadística. Se demostró que para que pueda ser aplicada esta metodología sobre un modelo de optimización particular se deben cumplir algunas condiciones generales. El requerimiento de mensurabilidad de la función que define al modelo general está asegurado si las funciones que caracterizan al problema de optimización son medibles, y esta es una condición que poseen las funciones típicamente utilizadas en la formulación de modelos de sistemas biológicos. El requerimiento de que el modelo nulo tenga una medida finita está asegurado en casos donde el espacio muestral tenga una medida finita, mientras que en caso de un espacio muestral de medida infinita dependerá de las restricciones planteadas que se cumpla con esta condición. El último requerimiento se aplica en caso de un espacio muestral métrico no contable, y se relaciona con la topología que adquiere el conjunto de soluciones óptimas cuando el modelo general se acerca al modelo de optimización estricta. Esta condición presenta una mayor complejidad con respecto de las condiciones anteriores, y se relaciona con una limitante inherente a los métodos de inferencia basados en la función de verosimilitud: no resultan adecuados para comparar el ajuste de modelos probabilísticos si sus medidas poseen un componente mutuamente singular no nulo.

Una de las principales ventajas de la metodología desarrollada es que su generalidad hace posible su aplicación a modelos de optimización de sistemas biológicos de jerarquía y complejidad arbitrarias. En consecuencia, la inferencia estadística sobre modelos de sistemas de mayor complejidad que los analizados en este trabajo posee el mismo fundamento, y su interpretación similar a la aquí realizada en el Capítulo 4. La disponibilidad de métodos numéricos de aplicabilidad general más allá de la formulación matemática de las restricciones resultarían por lo tanto de gran utilidad, especialmente cuando existan restricciones no lineales. Sin embargo, la demanda computacional de los métodos numéricos necesarios para estimar la distribución de probabilidad sobre el espacio muestral se incrementa a medida que el modelo del sistema se hace más complejo, por lo que se espera que la eficiencia de los métodos utilizados sea crítica para mantener tiempos de cómputo razonables. En tales casos es posible que sea necesario el uso de algoritmos específicos que sean eficientes para el tipo particular de modelo de optimización en cuestión para que sea factible la obtención de resultados en tiempos razonables de cómputo.

En el Capítulo 4 se presentó una implementación de la metodología desarrollada para la evaluación de modelos de optimización en la forma de una biblioteca de funciones en el lenguaje de programación MATLAB. Esta implementación se diseñó con el objetivo específico de que sea aplicable a cualquier sistema biológico particular para el que se puedan escribir las funciones específicas necesarias, ya que de esta forma es

que se define el análisis a realizar. Con el mismo fin se realizó una automatización de los análisis de inferencia estadística y las salidas gráficas, lo que facilita la aplicación de esta metodología a otros sistemas. El uso de esta herramienta de análisis está limitado por la disponibilidad de algoritmos capaces de realizar un muestreo de Monte Carlo sobre el espacio de soluciones óptimas de un modelo general probabilístico de optimización. Dado que el modelo analizado en el Capítulo 4 es un modelo de optimización lineal, mediante una modificación de las funciones utilizadas en este caso es posible aplicar el método de inferencia a cualquier otro sistema para el que resulte adecuado un modelo de optimización lineal. Los métodos de Monte Carlo para la obtención de muestras aleatorias del espacio de soluciones son computacionalmente intensivos. En la implementación realizada se utilizaron las capacidades de cómputo en paralelo de MATLAB, lo que permite utilizar la totalidad de los núcleos de los microprocesadores actuales. La función de estimación de  $F_{f_\theta}$  implementada se basa en la distribución empírica de las muestras aleatorias de  $\Omega$  del modelo nulo

Si bien en el diseño de la implementación de la metodología se intentó hacer énfasis en una estructura que permita realizar una inferencia estadística sobre un modelo de optimización arbitrario, en su estado actual la implementación posee algunas limitaciones. La función de transferencia no está explícitamente incluida como una entrada, sino que forma parte de la función de muestreo por Monte Carlo de  $\Omega$ , por lo que es posible utilizar cualquier otra función de transferencia. El uso de otra función de transferencia, sin embargo, requeriría una revisión del código de la función de inferencia estadística, ya que este asume un espacio de parámetros  $Q$  de una única dimensión. Otro aspecto que requiere ser modificado para poder representar el caso general es el establecimiento de las distribuciones de probabilidad previas sobre  $Q$  de los modelos de optimización. La realización de una inferencia estadística bayesiana, a diferencia de los análisis frecuentistas o basados en la teoría de la información, requieren establecer una distribución de probabilidad previa, que representa el conocimiento que se tiene a priori del espacio de parámetros sobre el que se desea realizar una inferencia. En el estado actual de la implementación la distribución previa para el modelo nulo se concentra en el valor de  $q$ , correspondiente a  $q_N \in Q$ , mientras que para los modelos de optimización es uniforme sobre  $Q$ , asumiendo como ya se dijo un espacio  $Q$  unidimensional. La distribución uniforme se eligió al ser una distribución no informativa, ya que es la de máxima entropía, pero otras distribuciones podrían emplearse. La función de transferencia y distribuciones de probabilidad previas sobre  $Q$  representan, en consecuencia, aspectos cuyo estudio e implementación de alternativas contribuiría a ampliar el menú de modelos de optimización para los que la metodología presentada en este trabajo puede ser aplicada para la evaluación de hipótesis de optimización.

Por último, creemos que la capacidad de establecer el valor empírico de modelos de optimización puede aportar información no solo sobre el sustento de modelos particulares, sino de forma más general sobre principios de optimización hipotéticos. Como

hemos mencionado, en la biología el caso más común de principio de optimización es la selección natural, cuya aceptación es unánime al ser el concepto central de la teoría evolutiva de Darwin. Cuando se evalúa una hipótesis basada en la selección natural, en consecuencia, el objetivo suele ser establecer el valor empírico del resto de los componentes del modelo utilizado, como restricciones, supuestos accesorios, y la relación entre las variables de estado cuantificadas y la aptitud biológica. Una situación contrastante se da para el caso de principios de optimización que se han postulado en el marco del estudio termodinámico de los sistemas biológicos. Si bien se ha mostrado que algunos de estos principios pueden ser complementarios, otros parecen ser excluyentes para una situación experimental dada, como el caso de una minimización de la disipación específica frente a una maximización de la producción de entropía (Fath y col. 2001). El principio de la maximización de la producción de entropía en particular ha recibido atención en los últimos años, y algunos trabajos reportan resultados en acuerdo con lo predicho (e.g. Jesus y col. 2012). Sin embargo, estos trabajos adolecen en general de las deficiencias ya señaladas al evaluar el poder predictivo de las hipótesis, por lo que el sustento empírico de este y otros postulados resta por ser esclarecido. De acuerdo a lo expuesto, puede descartarse que los autores consideren como un supuesto al principio de optimización. Más aún, el objetivo explícito de estos trabajos suele ser brindar soporte empírico al principio de optimización en cuestión. Esto indica que las deficiencias señaladas pueden tener origen en la falta de una metodología que permita evaluar hipótesis de optimización cumpliendo con los requerimientos ya mencionados. Dado el menor grado de desarrollo en el uso de principios de optimización termodinámica, la aplicación del método presentado podría representar un beneficio extra en este campo de aplicación en comparación con trabajos en el marco de la biología evolutiva clásica.

En este trabajo hemos presentado las bases de una metodología basada en la teoría axiomática de la probabilidad que permite evaluar modelos de optimización bajo ciertas condiciones generales, y que provee una solución a las críticas realizadas a los métodos comúnmente utilizados en la evaluación del sustento empírico de hipótesis de optimización al permitir la aplicación de métodos de inferencia estadística estándar. Al estar planteada de una manera general con asiento en la teoría matemática de la medida, la metodología desarrollada puede aplicarse a modelos de optimización de sistemas biológicos independientemente de su nivel de organización o complejidad del modelo matemático subyacente. La aplicación de la metodología desarrollada a modelos de sistemas de una mayor complejidad, así como la extensión a modelos de optimización multi-objetivo y multi-nivel, son líneas de investigación que se podrían desarrollar en base a lo expuesto en este trabajo de tesis. Esperamos que los aportes de este trabajo contribuyan al avance del conocimiento acerca del rol de los procesos de optimización como principios generales de organización de los sistemas biológicos.

## APÉNDICES

---

En este apéndice se incluye el código utilizado en la obtención de los resultados presentados en el cuerpo de la tesis.

### a.1 CAPÍTULO 4

#### a.1.1 Implementación metodología general

La implementación de la metodología general consta de 3 funciones básicas, y cada una de estas utiliza a su vez como entradas funciones designadas con el usuario de acuerdo al espacio muestral  $\Omega$  y el espacio de parámetros  $\Theta$  del modelo de optimización que se desee analizar. La función “mc\_sampling\_gral\_optimization\_model” realiza una obtención de muestras aleatorias del espacio de parámetros  $\Theta$ , el modelo nulo  $\Omega \times \Theta \times q_N$ , y de los espacios de soluciones óptimas  $\Omega \times \Theta \times q$  basados en un modelo general de optimización con una función de transferencia arbitraria y una función de normalización cuantil, esta última implementada como parte de la función. La función “radon\_nikodym\_density\_random\_var\_gral\_opt\_model” realiza la estimación de la densidad de Radon-Nikodym. La función “statistical\_inference\_gral\_opt\_model”, a diferencia de las funciones antes descritas, acepta como entrada múltiples modelos de optimización, de modo que es permite realizar una inferencia basada en múltiples modelos de optimización sobre sistemas independientes, siempre que dichos modelos compartan una misma función de transferencia y de normalización.

##### a.1.1.1 Muestreo espacio de soluciones

```

1 function [prob_opt_model] = mc_sampling_gral_optimization_model ( input_som )
  %%FUNDAMENTO
  % Espacio muestral  $\Omega$  de  $k$  dimensiones
  % Espacio de par metros  $\Theta$  de  $h$  dimensiones
  % Espacio de par metros de funci n de transferencia  $Q$  de  $l$  dimensiones .
6 %SUPUESTOS
  %La dimensionalidad de  $Q$  es baja
  time_ref_fun = tic ;
  perf = [] ;
  mkdir ( input_som . data_management . mc_sampling . dir_save_d_results )
11 mkdir ( input_som . data_management . mc_sampling . common_dir_saved_results ) ;
  %%MUESTREO ESPACIO DE PARAMETROS

```

```

% Se toman $m$ elementos de $\Theta$ al azar, y por cada uno de los $m$
% elementos $\theta$ se toman $n$ muestras de $\Omega$.
tref_theta_sampling = tic;
16 [mcsamples_Theta, prob_opt_model.Theta.logs] = input_som.Theta.mc_sampling.
    fun (...
        input_som.Theta.mc_sampling.m, ...
        input_som.Theta.mc_sampling.pars);
perf(1).Section = 'Theta sampling';
perf(1).Time = toc(tref_theta_sampling);
21 disp(['Theta sampling took ' num2str(toc(tref_theta_sampling)/60) ' minutes
    for ' ...
        num2str(input_som.Theta.mc_sampling.m) ' samples.'])
% MUESTREO DEL MODELO NULO ($\Omega \times \Theta \times q_N$)
% En paralelo a nivel de $\Theta$
tref_quantile_normalization = tic;
26 null_model_Omega_mc_sampling = input_som.null_model.Omega_mc_sampling;
input_quantile_normalization = input_som.quantile_normalization;
q_null_model = input_som.Q.q_null_model;
ticBytes(gcp);
parfor i_theta = 1 : input_som.Theta.mc_sampling.m
31     %% Paralell computing
    sampling_pars = null_model_Omega_mc_sampling.sampling_pars; %%ok<PFBNS>
    %% Temp directory for each worker
    sampling_pars.temp_dir = [null_model_Omega_mc_sampling.sampling_pars.
        temp_dir ...
            '/nm_' theta_ num2str(i_theta) '_' datestr(now, 'yyyy_mm_dd-HH:MM:
                SS.FFF')];
36     cd (null_model_Omega_mc_sampling.sampling_pars.temp_dir)
    if numel(dir(sampling_pars.temp_dir)) == 0
        mkdir(sampling_pars.temp_dir)
    end
    %%
41     % mcsamples_Theta: arreglo de $h \times m \times n_{cases}$ dimensiones.
    [null_Omega_mcsamples{i_theta}, null_Omega_logs{i_theta}] = ...
        null_model_Omega_mc_sampling.fun(mcsamples_Theta{i_theta}, ...
            input_som.c.as_theta_fun, ...
            q_null_model, ...
46     input_quantile_normalization.n_per_theta_null_model, ...
            sampling_pars); %%ok<PFBNS>
    % null_Omega_mcsamples: arreglo de $m$ unidades de $n \times k$
        dimensiones.
    [~,~,~] = rmdir(sampling_pars.temp_dir);
    disp([num2str(i_theta) ' null model theta completed'])
51 end
prob_opt_model.null_model.Omega.logs = null_Omega_logs;
parallel_data = tocBytes(gcp);
disp(['Null model sampling took ' num2str(toc(tref_quantile_normalization)
    /60) ' minutes for ' ...

```

```

num2str(input_som.Theta.mc_sampling.m) ' samples from Theta samples , with
'...
56 num2str( input_quantile_normalization . n_per_theta_null_model ) 'omega
samples per theta , '...
num2str( toc ( tref_quantile_normalization )/input_som.Theta.mc_sampling.m) '
s per each theta sample '])
perf(end+1).Section = 'Null_model sampling ';
perf(end).Time = toc(tref_quantile_normalization);
perf(end).Gbytes_sent_paralell_workers = sum(parallel_data(:,1))/(10^9);
61 perf(end).Gbytes_from_paralell_workers = sum(parallel_data(:,2))/(10^9);
% Save definition of null model
prob_opt_model.null_model.definition = input_som.null_model;
input_som = rmfield(input_som, 'null_model');
%%NORMALIZACION CUANTIL
66 %ESTIMACION FUNCION DE DISTRIBUCION FUNCION OBJETIVO
% Para cada uno de los $m$ elementos de $\Theta$ se debe estimar la funcion
% de distribucion de la funcion objetivo a fin de permitir la normalizacion
% cuantil.
F_f_theta_estimation = input_som.quantile_normalization.F_f_theta_estimation;
71 Q_def = input_som.Q;
for i_opt_model = 1: numel(input_som.alt_opt_models)
    tref_F_f_theta = tic;
    f = input_som.alt_opt_models(i_opt_model).f;
    ticBytes(gcp);
76 for i_theta = 1: input_som.Theta.mc_sampling.m
        [mcsamples_f_null_model] = f.fun(...
            null_Omega_mcsamples{i_theta}, ...
            mcsamples_Theta{i_theta}, f.pars); %#ok<PFBNS>
        % mcsamples_f_null_model : arreglo de $n \times m \times n_{obj\_fun}$
        dimensiones.
81 % Cada vector columna es una muestra aleatoria de la funcion objetivo
        dado
        % uno de los $m$ elemento tomados de $\Theta$.
        [F_f_theta{i_opt_model, i_theta}] = F_f_theta_estimation.fun(...
            mcsamples_f_null_model, ...
            Q_def, ...
86 f, ...
            F_f_theta_estimation.pars); %#ok<PFBNS>
        % F_f_theta : arreglo de $m$ elementos , donde cada uno es la funcion
        de
        % distribucion de la funcion objetivo para cada uno de los $m$
        elemento
        % tomados de $\Theta$.
91 end
parallel_data = tocBytes(gcp);
perf(end+1).Section = [input_som.alt_opt_models(i_opt_model).f.labels '
null_model distribution']; %#ok<*AGROW>
perf(end).Time = toc(tref_F_f_theta);
perf(end).Gbytes_sent_paralell_workers = sum(parallel_data(:,1))/(10^9);

```

```

96     perf(end).Gbytes_from_parallel_workers = sum( parallel_data (:,2) )/(10^9);
end
% Guardar resultados
if input_som . data_management . mc_sampling . data2save . null_model_samples
    if input_som . data_management . mc_sampling . write_results
101         cd ( input_som . data_management . mc_sampling . di r_s a v e d_r e s u l t s )
            filename_null_model_samples = [ input_som . data_management .
                filename_prefix '_null_model_samples' ...
                '_' datestr(now, 'yyyy-mm-dd-HH:MM:SS') '.mat' ];
            prob_opt_model . null_model . Omega . mcsamples {1} =
                save_or_load_data_opt_model (...
106                 'save', ...
                    filename_null_model_samples, ...
                    null_Omega_mcsamples );
            else
                prob_opt_model . null_model . Omega . mcsamples = null_Omega_mcsamples ;
            end
111 end
clear null_Omega_mcsamples
% MUESTREO ALEATORIO ESPACIO DE SOLUCIONES OPTIMAS
for i_opt_model = 1 : numel ( input_som . alt_opt_models )
    tref_gral_model_sampling = tic ;
116    Omega_mc_sampling = input_som . alt_opt_models ( i_opt_model ) .
        Omega_mc_sampling ;
    Omega_mc_sampling . sampling_pars . f = input_som . alt_opt_models ( i_opt_model )
        . f ;
    ticBytes ( gcp ) ;
    F_f_theta_i_opt_model = F_f_theta ( i_opt_model , : ) ;
    parfor i_q = 1 : numel ( input_som . Q . grid ) % PARFOR
121        % Declare variables so parallel matlab computation can classify
        % variables within for loop
        disp ( [ num2str ( i_q ) 'i_q iniciado ...' ] )
        general_Omega_mcsamples = { } ;
        strict_Omega_mcsamples = { } ;
126        general_Omega_logs = { } ;
        strict_Omega_logs = { } ;
        for i_theta = 1 : input_som . Theta . mc_sampling . m %#ok<PFBNS>
            % % Paralell computing
            sampling_pars = Omega_mc_sampling . sampling_pars ; %#ok<PFBNS>
            % Temp directory for each worker
131            sampling_pars . temp_dir = [ sampling_pars . temp_dir '/' ...
                'theta_' num2str ( i_theta ) 'q_' num2str ( i_q ) 'obj_' num2str (
                    i_opt_model ) ...
                '_' datestr ( now , 'yyyy_mm_dd-HH:MM:SS.FFF' ) ] ;
            cd ( Omega_mc_sampling . sampling_pars . temp_dir )
136            if numel ( dir ( sampling_pars . temp_dir ) ) == 0
                mkdir ( sampling_pars . temp_dir )
            end
            if i_q == 1

```

```

    sampling_pars.compute_optimal = 1;
141 end
    %%
    [general_Omega_mcsamples { i_theta }, ...
     general_Omega_logs { i_theta }, ...
     strict_Omega_mcsamples { i_theta }, ...
146 strict_Omega_logs { i_theta } ] = input_som.alt_opt_models (
        i_opt_model ). Omega_mc_sampling.fun (...
        mcsamples_Theta { i_theta }, ...
        input_som.c.as_theta_fun, ...
        input_som.Q.grid { i_q }, ...
        Omega_mc_sampling.n_per_theta, ...
151 sampling_pars, ...
        F_f_theta_i_opt_model { i_theta }); %#ok<PFBNS>
    % samples_model_q_omega : arreglo de $k \times n \times m$
    dimensiones
    cd ( Omega_mc_sampling.sampling_pars.temp_dir )
    [~,~,~] = rmdir ( sampling_pars.temp_dir );
156 end
    % Guardar resultados
    if input_som.data_management.mc_sampling.write_results
        cd ( input_som.data_management.mc_sampling.dir_saved_results )
        filename_omega_saved_data = [ input_som.data_management.
161 filename_prefix '_samples_q_i_' ...
            num2str ( i_q ) '_opt_model_i_' num2str ( i_opt_model ) ...
            '_' datestr ( now, 'yyyy-mm-dd-HH:MM:SS' ) '.mat' ];
        general_Omega_parallel_mcsamples { i_q } =
            save_or_load_data_opt_model (...
                'save', ...
                filename_omega_saved_data, ...
166 general_Omega_mcsamples );
    else
        general_Omega_parallel_mcsamples { i_q } = general_Omega_mcsamples ;
    end
    general_Omega_parallel_logs { i_q } = general_Omega_logs ;
171 strict_Omega_parallel_mcsamples { i_q } = strict_Omega_mcsamples ;
    strict_Omega_parallel_logs { i_q } = strict_Omega_logs ;
    disp ( [ num2str ( i_q ) ' i_q completed ' ] )
end % parfor q loop
parallel_data = tocBytes ( gcp );
176 prob_opt_model.opt_models ( i_opt_model ). general_model.Omega.mcsamples =
    general_Omega_parallel_mcsamples ;
prob_opt_model.opt_models ( i_opt_model ). general_model.Omega.logs =
    general_Omega_parallel_logs ;
prob_opt_model.opt_models ( i_opt_model ). strict_model.Omega.mcsamples =
    strict_Omega_parallel_mcsamples ;
prob_opt_model.opt_models ( i_opt_model ). strict_model.Omega.logs =
    strict_Omega_parallel_logs ;

```

```

clear general_Omega_parallel_mcsamples general_Omega_parallel_logs
strict_Omega_parallel_mcsamples strict_Omega_parallel_logs
181 disp(['General model sampling took ' num2str(toc(tref_gral_model_sampling
)/60) ' minutes for ' ...
num2str(input_som.Theta.mc_sampling.m) ' samples from Theta samples,
with ' ...
num2str(Omega_mc_sampling.n_per_theta) ' omega samples per theta and'
...
num2str(numel(input_som.Q.grid)) ' q values, for one objective
function ' ...
num2str(toc(tref_gral_model_sampling)/(input_som.Theta.mc_sampling.m*
size(input_som.Q.grid,2))) ...
186 's per each theta sample and q value.'])
perf(end+1).Section = [input_som.alt_opt_models(i_opt_model).f.labels '
general optimization modelsampling'];
perf(end).Time = toc(tref_gral_model_sampling);
perf(end).Gbytes_sent_paralell_workers = sum(parallel_data(:,1))/(10^9);
perf(end).Gbytes_from_paralell_workers = sum(parallel_data(:,2))/(10^9);
191 % Save F_f_theta
if input_som.data_management.mc_sampling.data_2_save.F_f_theta
if input_som.data_management.mc_sampling.write_results
cd(input_som.data_management.mc_sampling.dir_saved_results)
filename_F_f_theta_saved_data = [input_som.data_management.
filename_prefix '_F_f_theta_opt_model_i_' ...
196 num2str(i_opt_model) '_' datestr(now,'yyyy-mm-dd-HH:MM:SS') '
.mat'];
prob_opt_model.opt_models(i_opt_model).general_model.
normalization.F_f_theta = save_or_load_data_opt_model(...
'save',...
filename_F_f_theta_saved_data,...
F_f_theta(i_opt_model,:));
201 else
prob_opt_model.opt_models(i_opt_model).general_model.
normalization.F_f_theta = F_f_theta(i_opt_model,:);
end
end
F_f_theta(i_opt_model,:) = {0}; % Clear memory
206 end % for objective functions loop
% Save Theta samples
if input_som.data_management.mc_sampling.write_results
cd(input_som.data_management.mc_sampling.dir_saved_results)
filename_samples_theta = [input_som.data_management.filename_prefix '
211 '_samples_theta' ...
'_' datestr(now,'yyyy-mm-dd-HH:MM:SS') '.mat'];
prob_opt_model.Theta.mcsamples = save_or_load_data_opt_model(...
'save',...
filename_samples_theta,...
mcsamples_Theta);
216 else

```

```

    prob_opt_model.Theta.mcsamples = mcsamples_Theta;
end
% Save data management info
prob_opt_model.data_management = input_som.data_management;
221 input_som = rmfield(input_som, 'data_management');
% Save definition of alternative optimization models
for i = 1:numel(input_som.alt_opt_models)
    prob_opt_model.opt_models(i).definition = input_som.alt_opt_models(i);
end
226 input_som = rmfield(input_som, 'alt_opt_models');
% Save common model definition
prob_opt_model.definition = input_som;
% Save computing performance data
perf(end+1).Section = 'Other parts of code';
231 perf(end).Time = toc(time_ref_fun) - sum([perf.Time]);
cell_time_percent = num2cell(100*[perf.Time] / sum([perf.Time]));
[perf.Time_percent] = deal(cell_time_percent{:});
[perf(:).Function] = deal('Monte Carlo Sampling');
prob_opt_model.computing_performance = struct2table(perf);
236 % Write filename of masterfile
if prob_opt_model.data_management.mc_sampling.write_results
    filename_master_file = [prob_opt_model.data_management.mc_sampling.
        dir_saved_results '/' ...
        prob_opt_model.data_management.filename_prefix '
            _sampling_master_results_file '...
            '_' datestr(now, 'yyyy-mm-dd-HH:MM:SS') '.mat'];
241 prob_opt_model.data_management.mc_sampling.filename_master_file =
        filename_master_file;
    save(filename_master_file, 'prob_opt_model', '-v7.3', '-nocompression')
end

```

```

function [F_objective] = empirical_F_q_samples(samples_null_model_objective,
    Q_def, f_def, pars)
2 % Assume step transfer function
Q = cell2mat(Q_def.grid);
samples_null_model_objective = sort(samples_null_model_objective, 1);
n_null_samples = numel(samples_null_model_objective);
switch f_def.type
7     case 'max'
        quantil_index = (1:n_null_samples)/n_null_samples;
        q_F_index = sum(bsxfun(@le, quantil_index, Q), 2);
        q_F_index(1) = 1;
    case 'min'
12     quantil_index = (n_null_samples:-1:1)/n_null_samples;
        q_F_index = sum(bsxfun(@ge, quantil_index, Q), 2);
end
F_objective.f_q = samples_null_model_objective(q_F_index);
F_objective.q = Q;

```

a.1.1.2 *Estimación de densidad de Radon-Nikodym*

```

1 function [prob_opt_model]=radon_nikodym_density_random_var_gral_opt_model (
    prob_opt_model ,input_pdom)
% This function uses random samples from the solution space of a general
% optimization model to estimate the Radon–nikodym density of a random
% variable with respect to a base measure. Typically , random variables are
% real functions , and the Lebesgue measure is the base measure.
6 %%LOAD THETA DATA
if prob_opt_model . data_management . mc_sampling . write_results
    cd ( prob_opt_model . data_management . mc_sampling . dir_save_d_results )
    Theta_mcsamples = save_or_load_data_opt_model ( 'load ' ,prob_opt_model . Theta
        . mcsamples ) ;
else
11 Theta_mcsamples = prob_opt_model . Theta . mcsamples ;
end
%%FOLDER TO STORE RESULTS
prob_opt_model . data_management . radon_nikodym_density = input_pdom .
    data_management . radon_nikodym_density ;
if ~isfield ( prob_opt_model . data_management . radon_nikodym_density , '
    write_results ' )
16 prob_opt_model . data_management . radon_nikodym_density . write_results = 0 ;
end
% Create folders inside folders from previous data
mkdir ( prob_opt_model . data_management . radon_nikodym_density . dir_save_d_results )
    ;
mkdir ( prob_opt_model . data_management . radon_nikodym_density .
    common_dir_saved_results )
21 %%ESTIMATE PROBABILITY DENSITY FUNCTION OF OBSERVED RANDOM VARIABLES FOR
    EACH VALUE OF q
% Avoid passing to parallel workers all of prob_opt_model
load_mc_sampling_data = prob_opt_model . data_management . mc_sampling .
    write_results ;
mc_sampling_dir_saved_results = prob_opt_model . data_management . mc_sampling .
    dir_saved_results ;
% Reduce memory exchange between parallel workers
26 X = input_pdom . X ;
for i_model = 1 : numel ( prob_opt_model . opt_models ) +1
    % Put null_model and general_models in a same data structure
    if i_model <= numel ( prob_opt_model . opt_models )
        model_i = prob_opt_model . opt_models ( i_model ) . general_model ;
31 else
        model_i = prob_opt_model . null_model ;
    end
    % Declare variables to avoid redefinition at each loop cycle
    X_radon_nikodym_density_paralell = cell ( [ numel ( model_i . Omega . mcsamples )
        numel ( input_pdom . X ) ] ) ;
36 for i_q = 1 : numel ( model_i . Omega . mcsamples )
        X_radon_nikodym_density_i_q = cell ( [1 numel ( X ) ] ) ;

```

```

41     if load_mc_sampling_data
        cd ( mc_sampling_dir_saved_results )
        Omega_mcsamples = save_or_load_data_opt_model ( ...
            'load ', model_i.Omega.mcsamples { i_q } );
    else
        Omega_mcsamples = model_i.Omega.mcsamples { i_q };
    end
46     for i_X = 1:numel(X)
        % Compute random variable
        X_mcsamples = X(i_X).fun(Omega_mcsamples, Theta_mcsamples); %#ok<*
        PFBNS>
        % Estimate density
51     [X_radon_nikodym_density_i_q { i_X } ] = ...
        X(i_X).rn_denisty_estimation.fun(X_mcsamples, ...
        X(i_X).rn_denisty_estimation.pars);
    end
    X_radon_nikodym_density_parell(i_q, :) = X_radon_nikodym_density_i_q
        ;
    % NOTE: slicing is necessary due to parallel processing
56 end % parfor q
    for i_X = 1:numel(input_pdom.X)
        if i_model <= numel(prob_opt_model.opt_models) % Alternative
            optimization models
            prob_opt_model.opt_models(i_model).general_model.X(i_X).
                radon_nikodym_density = ...
                X_radon_nikodym_density_parell(:, i_X);
        else % Null model
61     prob_opt_model.null_model.X(i_X).radon_nikodym_density =
                X_radon_nikodym_density_parell(:, i_X);
        end
    end
end
prob_opt_model.definition.X = input_pdom.X;
66 if prob_opt_model.data_management.radon_nikodym_density.write_results
    cd(prob_opt_model.data_management.radon_nikodym_density.dir_saved_results
        )
    filename_master_file = [prob_opt_model.data_management.filename_prefix '
        _rn_density_master_results_file' ...
        '_' datestr(now, 'yyyy-mm-dd-HH:MM:SS') '.mat'];
    prob_opt_model.data_management.radon_nikodym_density.filename_master_file
        = filename_master_file;
71 save(filename_master_file, 'prob_opt_model', '-v7.3', '-nocompression')
end

```

```

function [radon_nikodym_density] = univariate_pdf_kernel_estimation(
    rand_var_samples, input_st)
rand_var_samples = cell2mat(rand_var_samples);
3 rand_var_samples = rand_var_samples(:);
[n, x_out] = hist(rand_var_samples, input_st.empirical_hist_division);

```

```

empirical_cdf=cumsum(n)/sum(n);
rand_var_samples_mod=rand_var_samples;
%%If range is bounded and known, extend histogram to improve fitting
8 if isfield(input_st, 'min_bound') && (min(rand_var_samples)<input_st.min_bound
+input_st.tolerance_to_bound)
    % seleccionar 0.5% inferior
    rep_min=rand_var_samples(rand_var_samples <x_out(find(empirical_cdf >
input_st.data_perc_to_repeat,1)));
    rep_min=bsxfun(@plus, rep_min, -(range(rep_min).*(1:round(input_st.
perc_rep_extreme_data/input_st.data_perc_to_repeat))));
    rand_var_samples_mod=[rand_var_samples_mod; rep_min(:)];
13 end
if isfield(input_st, 'max_bound') && (max(rand_var_samples)+input_st.
tolerance_to_bound >input_st.max_bound)
    % seleccionar 0.5% inferior
    rep_max=rand_var_samples(rand_var_samples >x_out(find(empirical_cdf >(1-
input_st.data_perc_to_repeat),1)));
    rep_max=bsxfun(@plus, rep_max, range(rep_max).*(1:round(input_st.
perc_rep_extreme_data/input_st.data_perc_to_repeat))));
18 rand_var_samples_mod=[rand_var_samples_mod; rep_max(:)];
end
%%Kernel density estimation function by Z. I. Botev, J. F. Grotowski, and D
. P.
% Kroese (2010) Annals of Statistics, Volume 38, Number 5, pages 2916 –2957.
[bandwidth, density, xmesh, cdf]=kde(rand_var_samples_mod, input_st.
intervals_kernel_estimation);
23 xmesh_mod=xmesh;
%%Eliminate points out of known bounds
if isfield(input_st, 'min_bound') && (min(rand_var_samples)<input_st.min_bound
+input_st.tolerance_to_bound)
    selection_min_bound=xmesh_mod>input_st.min_bound;
    xmesh_mod=xmesh_mod(selection_min_bound);
28 density=density(selection_min_bound);
else
end
if isfield(input_st, 'max_bound') && (max(rand_var_samples)+input_st.
tolerance_to_bound >input_st.min_bound)
    selection_max_bound=xmesh_mod<input_st.max_bound;
33 xmesh_mod=xmesh_mod(selection_max_bound);
    density=density(selection_max_bound);
end
%%Eliminate points out of empirical bounds
if ~isempty(find(xmesh_mod<min(rand_var_samples),1, 'last')) %Out of empirical
min
38 selection_lower=find(xmesh_mod<min(rand_var_samples),1, 'last')+1;
else
    selection_lower=1;
end

```

```

if ~isempty( find(xmesh_mod>max(rand_var_samples),1,'first')) %Out of
    empirical max
43     selection_higher=find(xmesh_mod>max(rand_var_samples),1,'first')-1;
else
    selection_higher=length(xmesh_mod);
end
xmesh_mod=xmesh_mod(selection_lower:selection_higher);
48 density=density(selection_lower:selection_higher);
%%Extrapolate (linearly) density to empirical bounds
[pd_extrapolated_min,x_extrapolated_min]=univariate_extapolate_pd(min(
    rand_var_samples),xmesh_mod,density);
[pd_extrapolated_max,x_extrapolated_max]=univariate_extapolate_pd(max(
    rand_var_samples),xmesh_mod,density);
radon_nikodym_density.grid=[x_extrapolated_min xmesh_mod x_extrapolated_max
];
53 radon_nikodym_density.density=[pd_extrapolated_min density '
    pd_extrapolated_max];
%%Set to zero if density negative values were (erroneously) estimated
radon_nikodym_density.density(radon_nikodym_density.density < 0) = 0;
%%Graphics
if isfield(input_st,'graphs_options') && input_st.graphs_options.make_graph
58     %Check constrained polynomial fit of CDF
    fig=figure('visible','off');
    left_color=[0 0 0];
    right_color=[0 0 0];
    set(fig,'defaultAxesColorOrder',[left_color;right_color]);
63     %PDF vs histogram (normalized so as to sum 1)
    yyaxis left
    bar_h=bar(x_out,n./sum(n*mean(x_out(2:end)-x_out(1:end-1))),...
        'FaceColor',[0.5 0.5 0.5]);
    hold on
68     ylabel('Frequency');
    yyaxis right
    plot(radon_nikodym_density.grid,radon_nikodym_density.density,'color','
        black')
    hold off
    ylabel('Radon-Nikodym density');
73     xlabel([input_st.graphs_options.xlabel_title])
    h_legend=legend('Empirical frequency','Kernel density estimation','
        Location','northwest');
    legend('boxoff')
    print('-djpeg',[ 'Empirical vs fitted PDF - ' input_st.graphs_options.
        filetype '.jpg'],'-r300')
    close
78 end

```

### a.1.1.3 Inferencia estadística

```

function [stat_inf,lnL] = statistical_inference_gral_opt_model(prob_opt_model
    ,x,input_siom)
2 % This function uses random samples from the solution space of a general
% optimization model to estimate the Radon–nikodym density of a random
% variable with respect to a base measure. Typically, random variables are
% real functions, and the Lebesgue measure is the base measure.
%%VALIDATE REQUIRED STRUCTURE OF MODELS TO ANALYSE
7 % Check all cases whitin prob_opt_model have equal number of alternative
models
for i_cs = 1:numel(prob_opt_model)
    % Alternative optimization models
    if ~(numel(prob_opt_model(1).opt_models) == numel(prob_opt_model(i_cs).
        opt_models))
        error('The number of alternative optimization models (prob_opt_model .
            opt_models) has to be the same for every analyzed case .')
12    end
    %Q
    if ~isequal(prob_opt_model(1).definition.Q,prob_opt_model(i_cs).
        definition.Q)
        error('Definition of Q (prob_opt_model.definition.Q) must be the same
            for all study cases .')
17    end
    %Random variables
    for i_X = 1:numel(prob_opt_model(i_cs).definition.X)
        try
            if ~isequal(prob_opt_model(1).definition.X(i_X),prob_opt_model(i
                cs).definition.X(i_X))
                error('Definition of X (prob_opt_model.definition.X) must be
                    the same for all study cases .')
22            end
        catch
            error('De finition of X (prob_opt_model.definition.X) must be the
                same for all study cases .')
        end
    end
27    % Shared results folder
    if ~isequal(prob_opt_model(1).data_management.mc_sampling.
        common_dir_saved_results ,...
        prob_opt_model(i_cs).data_management.mc_sampling.
            common_dir_saved_results)
        error(['The shared results folder must be the same in all study cases
            '...
            '(prob_opt_model.data_management.mc_sampling.
                common_dir_saved_results)'])
32    end
end
if numel(input_siom.inference.q_to_compare) > 3

```

```

    error('Choose at most 3 values of q to compare optimization models -
          numel(input_siom.inference.q_to_compare) > 3')
end
37 n_alt_opt_models = numel(prob_opt_model(1).opt_models);
   %%FOLDER TO STORE RESULTS
shared_id_stamp = input_siom.data_management.statistical_inference.
    shared_id_stamp;
common_dir_saved_results = [prob_opt_model(1).data_management.
    radon_nikodym_density.common_dir_saved_results ...
    '/Statistical_inference_' shared_id_stamp];
42 % Create folders inside folders from previous data
for i_case_study = 1:numel(prob_opt_model)
    % Not(necesarely) shared
    prob_opt_model(i_case_study).data_management.statistical_inference.
        dir_saved_results = ...
        [prob_opt_model(i_case_study).data_management.radon_nikodym_density.
        dir_saved_results ...
47    '/Statistical_inference_' shared_id_stamp];
    mkdir(prob_opt_model(i_case_study).data_management.statistical_inference.
        dir_saved_results);
    % Shared
    prob_opt_model(i_case_study).data_management.statistical_inference.
        common_dir_saved_results = ...
        common_dir_saved_results;
52 end
mkdir(prob_opt_model(1).data_management.statistical_inference.
    common_dir_saved_results);
   %%ESTIMATE LIKELIHOOD FUNCTION OF q BASED ON OBSERVED RANDOM VARIABLES x FOR
   EACH MODEL
n_x_i = cellfun(@numel,x);
X_def = prob_opt_model(1).definition.X; %NOTE: X are the same for all studied
    cases
57 for i_model = 1:n_alt_opt_models+1
    if i_model <= n_alt_opt_models
        lnL(i_model).raw = NaN([numel(prob_opt_model(1).definition.Q.grid)
            numel(prob_opt_model) size(x,2) max(n_x_i(:))]);
    else % Null model
        lnL(i_model).raw = NaN([1 numel(prob_opt_model) size(x,2) max(n_x_i
            (:))]);
62 end
for i_case_study = 1:numel(prob_opt_model)
    % Put null_model and general_models in a same data structure
    if i_model <= n_alt_opt_models
        model_i = prob_opt_model(i_case_study).opt_models(i_model).
            general_model;
67 else
        model_i = prob_opt_model(i_case_study).null_model;
    end
    % Reduce memory exchange to parallel workers

```

```

72     radon_nikodym_density_i_model = [ model_i.X.radon_nikodym_density ];
    prob_opt_model( i_case_study ).x = x( i_case_study ,: );
    n_x_i_case_study = cellfun(@numel,x(i_case_study ,:));
    x_i_model = x(i_case_study ,:);
    ln_likelihood_i_model = NaN( [ numel( model_i.Omega.mcsamples) ...
77     numel(X_def) max( n_x_i_case_study ) ] );
    for i_q = 1:numel( model_i.Omega.mcsamples )
        likelihood_i_q = NaN( [ numel(X_def) max( n_x_i_case_study ) ] );
        % Loop over random variables available for this study case
        for i_X = 1:numel(X_def)
            % Loop over independent data for each random variable
            % available for this study case
82         for i_x = 1:numel(x_i_model{i_X}) %#ok<PFBNS>
                likelihood_i_q( i_X, i_x ) = log( X_def( i_X ).
                    rn_density_at_value . fun ( ...
                        radon_nikodym_density_i_model { i_q , i_X } , ...
                        x_i_model { i_X } { i_x } , ...
                        X_def( i_X ). rn_density_at_value . pars ) );
87         end % i_x
        end % i_X
        ln_likelihood_i_model( i_q , : , : ) = likelihood_i_q ;
    end % parfor q
    if i_model <= n_alt_opt_models
92        prob_opt_model( i_case_study ). opt_models( i_model ). general_model .
            likelihood = ...
            ln_likelihood_i_model ;
    else
        prob_opt_model( i_case_study ). null_model . likelihood =
            ln_likelihood_i_model ;
    end
97    % Aggregate likelihood data
    lnL( i_model ). raw( 1 : numel( model_i.Omega.mcsamples ) , i_case_study , : , 1 :
        size( ln_likelihood_i_model , 3 ) ) ...
        = ln_likelihood_i_model ; % Dims: 1: Q ; 2: i_case ; 3: i_X 4: i_x
        ;
    end
end
102 %%DEFINE ANALISIS TO BE MADE
%Any analysis defined share the same basic structure: an array of ln
% likelihood
%SAME q MODELS
for i_opt = 1:n_alt_opt_models+1
107    lnL( i_opt ). size = size( lnL( i_opt ). raw );
    if numel( lnL( i_opt ). size ) < 3
        lnL( i_opt ). size( 3 ) = 1 ;
    end
    if numel( lnL( i_opt ). size ) < 4
112    lnL( i_opt ). size( 4 ) = 1 ;
    end
end

```

```

lnL(i_opt).nan = isnan(lnL(i_opt).raw);
lnL(i_opt).zeroed = lnL(i_opt).raw;
lnL(i_opt).zeroed(lnL(i_opt).nan) = 0;
117 % % Global results (all cases)
% Considering all available data
stat_inf(i_opt).global_same_q.lnL = sum(reshape(lnL(i_opt).zeroed,lnL(
    i_opt).size(1),prod(lnL(i_opt).size(2:end))),2);% lnL
stat_inf(i_opt).global_same_q.n = sum(~lnL(i_opt).nan(1,:)); % n
% Considering separately each random variable
122 if numel(X_def) > 1 % Only if there is more than 1 random variable
    for i_X = 1:numel(X_def)
        stat_inf(i_opt).per_X_same_q(i_X).lnL = ...
            sum(reshape(lnL(i_opt).zeroed(:, :, i_X, :), lnL(i_opt).size(1),
                prod(lnL(i_opt).size([2 4]))), 2); % lnL
        stat_inf(i_opt).per_X_same_q(i_X).n = sum(reshape(~lnL(i_opt).nan
            (1, :, i_X, :), 1, prod(lnL(i_opt).size([2 4])))); % n
127    end
end
% % Results per case
if numel(prob_opt_model) > 1 % Only if there is more than 1 studied
system
    for i_case_study = 1:numel(prob_opt_model)
132 % Considering all available data
stat_inf(i_opt).per_case(i_case_study, 1).lnL = ...
    sum(reshape(lnL(i_opt).zeroed(:, i_case_study, :, :), lnL(i_opt).
        size(1), prod(lnL(i_opt).size([3 4]))), 2); % lnL
stat_inf(i_opt).per_case(i_case_study, 1).n = ...
    sum(reshape(~lnL(i_opt).nan(1, i_case_study, :, :), 1, prod(lnL(
        i_opt).size([3 4])))); % n
137 if numel(X_def) > 1 % Only if there is more than 1 random
variable
    for i_X = 1:numel(X_def)
        % Considering separately each random variable
stat_inf(i_opt).per_case_and_X(i_case_study, i_X).lnL = ...
    sum(reshape(lnL(i_opt).zeroed(:, i_case_study, i_X, :),
        lnL(i_opt).size(1), prod(lnL(i_opt).size([4]))), 2);
        % lnL
142 stat_inf(i_opt).per_case_and_X(i_case_study, i_X).n = sum
    (~lnL(i_opt).nan(1, i_case_study, i_X, :)); % n
        end
    end
end
end
147 end
% MAXIMUM LIKELIHOOD
Q = prob_opt_model(1).definition.Q; % Note prob_opt_model(i).definition.Q are
equal for all i
analyses = fields(stat_inf);
for i_model = 1:n_alt_opt_models+1

```

```

152 for i_analysis = 1:numel( analyses ) %
    for i_instance = 1:numel( stat_inf(i_model).( analyses { i_analysis } ))
        lnL_level = stat_inf(i_model).( analyses { i_analysis } )( i_instance ) .
            lnL;
        if i_model <= n_alt_opt_models
            % Model degrees of freedom
157 stat_inf(i_model).( analyses { i_analysis } )( i_instance ).k = Q.
                dimensionality;
            % Maximum likelihood
            [maxL.lnL, maxL.i_q] = max(lnL_level);
            maxL.q = Q.grid {maxL.i_q};
            stat_inf(i_model).( analyses { i_analysis } )( i_instance ).maxL =
162 maxL;
        else
            % Null model
            stat_inf(i_model).( analyses { i_analysis } )( i_instance ).k = 0;
            maxL.lnL = lnL_level;
            maxL.i_q = Q.i_q_null_model;
167 maxL.q = Q.q_null_model;
            stat_inf(i_model).( analyses { i_analysis } )( i_instance ).maxL =
                maxL;
        end
    end
end
172 end
% Variable q models
[ var_q_analysis (1:2) .reference ] = deal ( 'per_case ', 'per_case_and_X' );
[ var_q_analysis (1:2) .field ] = deal ( 'global_var_q ', 'per_X_var_q ' );
for i_model = 1:n_alt_opt_models+1
177 for i_analysis = 1:numel( var_q_analysis )
    if isfield( stat_inf, var_q_analysis ( i_analysis ) .reference )
        ref_st = stat_inf( i_model ) .( var_q_analysis ( i_analysis ) .reference )
            ;
        size_st = size( ref_st );
        maxL = reshape ( [ ref_st .maxL ], size_st (1) , size_st (2) );
182 k = reshape ( [ ref_st .k ], size_st (1) , size_st (2) );
        n = reshape ( [ ref_st .n ], size_st (1) , size_st (2) );
        for i_instance = 1:size_st (2)
            maxL_new.lnL = sum ( [ maxL (: , i_instance ) .lnL ] );
            if i_model <= n_alt_opt_models
187 maxL_new.i_q = [ maxL (: , i_instance ) .i_q ];
                maxL_new.q = [ maxL (: , i_instance ) .q ];
            else
                maxL_new.i_q = Q.i_q_null_model ;
                maxL_new.q = Q.q_null_model ;
192 end
            stat_inf(i_model).( var_q_analysis ( i_analysis ) .field )(1,
                i_instance) .maxL = maxL_new;

```

```

stat_inf(i_model).(var_q_analysis(i_analysis).field)(1,
    i_instance).k = sum(k(:, i_instance), 1);
stat_inf(i_model).(var_q_analysis(i_analysis).field)(1,
    i_instance).n = sum(n(:, i_instance), 1);
197 if i_model <= n_alt_opt_models
    stat_inf(i_model).(var_q_analysis(i_analysis).field)(1,
        i_instance).lnL = [ref_st(:, i_instance).lnL];
    else % Null model
        stat_inf(i_model).(var_q_analysis(i_analysis).field)(1,
            i_instance).lnL = sum([ref_st(:, i_instance).lnL]);
    end
    end
202 end
end
end
end
%%STATISTICAL INFERENCE: MODEL SELECTION COMPARING NULL AND ALL ALTERNATIVE
OPTIMIZATION MODELS
analyses = fields(stat_inf);
207 if size(Q.grid, 1) < size(Q.grid, 2)
    Q.grid = Q.grid';
end
for i_analysis = 1:numel(analyses)
    level = [stat_inf(:).(analyses{i_analysis})];
212 numel_instances = numel(stat_inf(1).(analyses{i_analysis}));
    level = reshape(level, numel_instances, numel(level)/numel_instances);
    for i_instance = 1:numel_instances
        maxL_models = [level(i_instance, :).maxL];
        k_models = [level(i_instance, :).k];
217 n_models = [level(i_instance, :).n];
        %%Likelihood ratio test
        null_model_lnL = stat_inf(end).(analyses{i_analysis})(i_instance).
            maxL.lnL;
        clear Lratio_test
        for i_model = 1:numel(maxL_models)
222 if i_model < numel(maxL_models) % Null model is the last
            Lratio_test.Lratio_statistic = exp(null_model_lnL -
                maxL_models(i_model).lnL);
            Lratio_test.X2_alpha_95_cut = chi2inv(0.95, Q.dimensionality);
            Lratio_test.significant = - 2 * log(Lratio_test.
                Lratio_statistic) >= Lratio_test.X2_alpha_95_cut;
            Lratio_test.P = 1 - chi2cdf(- 2 * log(Lratio_test.
                Lratio_statistic), Q.dimensionality);
227 else
            Lratio_test = 'Analysis not applicable to the null model.';
        end
        stat_inf(i_model).(analyses{i_analysis})(i_instance).Lratio_test
            = Lratio_test;
    end
232 %%Information theory

```

```

for i_model = 1:numel(maxL_models)
    k = k_models(i_model);
    n = n_models(i_model);
    inf_theory.AIC(i_model).IC = 2 * k - 2 * maxL_models(i_model).lnL
    ;
237     if n-k-1 > 0
        inf_theory.AICc(i_model).IC = 2 * k - 2 * maxL_models(i_model)
            .lnL + (k^2+2*k)/(n-k-1);
    else
        inf_theory.AICc(i_model).IC = NaN;
    end
242 end
it_fields = fields(inf_theory);
%Compute delta IC and weights
for i_it = 1:numel(it_fields)
    IC_st = [inf_theory.(it_fields{i_it})];
247     IC = [IC_st.IC];
    delta = num2cell(IC - min(IC));
    weight = num2cell(exp(- cell2mat(delta)/2)/sum(exp(- cell2mat(delta)/2)));
    [IC_st.delta] = deal(delta{:});
    [IC_st.weight] = deal(weight{:});
252     [inf_theory.(it_fields{i_it})] = IC_st;
    for i_model = 1:numel(maxL_models)
        stat_inf(i_model).(analyses{i_analysis})(i_instance).(
            it_fields{i_it}) = ...
            inf_theory.(it_fields{i_it})(i_model);
    end
257 end
%%Bayesian inference
% Bayes factor: is equal to the ratio of marginal likelihood, which
    is obtained by
% integrating the prior times the likelihood over the whole free
    parameter space
% COMPUTE MARGINAL LIKELIHOOD OF EACH MODEL
262 ln_marginal_L_array = NaN(1, numel(maxL_models));
for i_model = 1:numel(maxL_models)
    switch level(i_instance, i_model).k
        case 0
            % Zero dimensional case: only one possible value
            % Null model or fixed-q optimization model
            ln_marginal_L = level(i_instance, i_model).lnL;
        case 1
            % One dimensional case
            % Marginal likelihood is obtained as the weighted
            % average of an extensive grid over Q
            % Assume Q grid is sorted
            Qgrid = cell2mat(Q.grid);
            Qlimits = [0 ; (Qgrid(2:end) + Qgrid(1:end-1))/2 ; 1];
272

```

```

277         Qweights = Qlimits (2:end) - Qlimits (1:end-1);
        ln_marginal_L = log (sum(Qweights .* exp (level (i_instance ,
            i_model) .lnL)));
        otherwise
            % Multidimensional case
            % Marginal likelihood is obtained by MC integration ,
            % uniform sampling of Q
282         Qgrid = cell2mat (Q.grid);
            Qlimits = [0 ; (Qgrid (2:end) + Qgrid (1:end-1)) / 2];
            input_siom.inference.bayes_factor .
                n_mc_samples_marginal_likelihood = 10 ^ 1;
            n_samples = input_siom.inference.bayes_factor .
                n_mc_samples_marginal_likelihood ;
            unif_random = rand (n_samples , level (i_instance , i_model) .k)
                ;
287         q_mc_samples = sum (bsxfun (@ge , unif_random , permute (Qlimits
            , [3 2 1])) , 3);
            lnL_mc_samples = ...
                sum (level (i_instance , i_model) .lnL (bsxfun (@ plus ,
                    q_mc_samples , (0 : size (q_mc_samples , 2) - 1) .* numel (Q.
                        grid))) , 2);
            ln_marginal_L = log (mean (exp (lnL_mc_samples))) ;
        end
292         stat_inf (i_model) .( analyses {i_analysis} ) (i_instance) .
            ln_marginal_L = ln_marginal_L ;
            ln_marginal_L_array (i_model) = ln_marginal_L ;
        end
        %COMPUTE BAYES FACTOR
        log10_bayes_factor_array = log10 (exp (bsxfun (@minus ,
            ln_marginal_L_array ' , ln_marginal_L_array))) ;
297         for i_model = 1 : numel (maxL_models)
            stat_inf (i_model) .( analyses {i_analysis} ) (i_instance) .
                log10_bayes_factor = log10_bayes_factor_array (i_model , :) ;
        end
    end
end
302 %%PLOTS
%Plots are implemented as custom functions
plots = input_siom.inference.plots ;
cd (prob_opt_model (i_case_study) .data_management .statistical_inference .
    dir_saved_results)
for i_plot = 1 : numel (plots.plot_def)
307     % Which kind of plot?
    per_case_and_individual_plot = plots.plot_def (i_plot) .is_per_case && ...
        plots.plot_def (i_plot) .individual_plots ;
    per_case_and_collective_plot = plots.plot_def (i_plot) .is_per_case && ...
        plots.plot_def (i_plot) .individual_plots == 0 ;
312     global_plot = plots.plot_def (i_plot) .is_per_case == 0 ;
    if plots.plot_def (i_plot) .is_per_X

```

```

        % Different plots for each X
        n_X_plot = numel(prob_opt_model(i_cs).definition.X);
else
317     % Same plot for all X
        n_X_plot = 1;
end
for i_X = 1:n_X_plot
    % Label
322     if plots.plot_def(i_plot).is_per_X
        i_X_label = prob_opt_model(i_cs).definition.X(i_X).label;
    else
        % Same plot for all X
        i_X_label = '';
327     end
    if plots.plot_def(i_plot).is_per_case
        % Different plots for each X
        n_case_plot = numel(prob_opt_model);
    else
332     % Same plot for all X
        n_case_plot = 1;
    end
    for i_cs = 1:n_case_plot
        %%Init figure (or determine subplot)
337     if i_cs == 1 || per_case_and_individual_plot
        close all
        fig_h = figure('visible','off');
        left_color = [0 0 0];
        right_color = [0 0 0];
342     set(fig_h, 'defaultAxesColorOrder', [left_color; right_color]);
    end
    if per_case_and_collective_plot
        n_rows_plot = ceil((numel(prob_opt_model)+1) / plots.plot_def(i_plot).n_column_subplots);
        subplot_h = subplot(n_rows_plot, plots.plot_def(i_plot).n_column_subplots, i_cs);
347     end
        %%Call plot fun
        handles = plots.plot_def(i_plot).fun(prob_opt_model, input_siom, x, stat_inf, i_X, i_cs, i_plot);
        %%Apply format to plot (or subplot)
    if global_plot
352     filename_plot_wo_fmt = [prob_opt_model(i_cs).data_management.analysis_filename_prefix ...
        plots.plot_def(i_plot).filename_string ...
        i_X_label];
    end
    if per_case_and_individual_plot
357     % Individual plots

```

```

        filename_plot_wo_fmt = [prob_opt_model(i_cs).data_management.
            filename_prefix ...
            plots.plot_def(i_plot).filename_string ...
            i_X_label];
    end
362 if global_plot || per_case_and_individual_plot
        axes_h = gca;
        axes_h.FontSize = plots.format.ind.axis_font_size;
        axes_h.XLabel.FontSize = plots.format.xylabls.font_size;
        if plots.plot_def(i_plot).secondYaxis
367             yyaxis left
                axes_h.YLabel.FontSize = plots.format.xylabls.font_size;
                yyaxis right
            end
            axes_h.YLabel.FontSize = plots.format.xylabls.font_size;
372 handles.null_model.MarkerSize = plots.format.ind.marker_size
            ;
            for i_om_plot = 1:numel(handles.null_model)
                handles.alt_opt_model(i_om_plot).MarkerSize = plots.
                    format.ind.marker_size;
            end
        end
    end
377 if per_case_and_collective_plot
        % Unified plot
        axis square
        %Add title to the subplot
        title(['\it{' prob_opt_model(i_cs).data_management.
            studied_system_id_string '}', 'FontSize',6)
382 % Eliminate labels of axes and legend
        labels.XLabel = subplot_h.XLabel.String;
        subplot_h.XLabel.delete
        if plots.plot_def(i_plot).secondYaxis
387             yyaxis left
                labels.YLabel_left = subplot_h.YLabel.String;
                subplot_h.YLabel.delete
                yyaxis right
                labels.YLabel_right = subplot_h.YLabel.String;
                subplot_h.YLabel.delete
392         else
            labels.YLabel_left = subplot_h.YLabel.String;
            subplot_h.YLabel.delete
            labels.YLabel_right = '';
        end
397 %Raw
        set(gca, 'FontSize', 4)
        p = get(subplot_h, 'position');
        p(4) = p(4) * 1.10; %Add 10 percent to height
        set(subplot_h, 'position', p);
402 if i_cs == numel(prob_opt_model)

```

```

legend ( 'Location ' , 'northeast ' )
% Modify paper size
hg_empty = subplot ( n_rows_plot , plots.plot_def ( i_plot ) .
    n_column_subplots , i_cs + 1 );
axis square
407 xlim ([ 0 1 ])
set ( gca , 'FontSize ' , 4 )
p = get ( hg_empty , ' position ' );
p ( 4 ) = p ( 4 ) * 1.10 ; % Add 10 percent to height
set ( hg_empty , ' position ' , p );
412 set ( gcf , ' PaperUnits ' , ' centimeters ' )
xSize = 15 ; ySize = 24 ;
xLeft = ( 21 - xSize ) / 2 ; yTop = ( 30 - ySize ) / 2 ;
if isempty ( labels . YLabel_right )
    set ( gcf , ' PaperPosition ' , [ xLeft yTop xSize ySize ] )
417 else
    % Make room for right common label
    set ( gcf , ' PaperPosition ' , [ xLeft yTop xSize ySize ] )
end
set ( gcf , ' Position ' , [ 0 0 xSize * 50 ySize * 50 ] )
422 % Move legend to empty subplot
pos_subplot = subplot_h . Position ;
pos_empty_subplot = hg_empty . Position ;
diff_pos_plot = pos_empty_subplot - pos_subplot ;
handles . legend . Position ( 1 ) = handles . legend . Position ( 1 ) ...
427 + diff_pos_plot ( 1 ) ;
set ( hg_empty , ' vis ' , ' off ' )
% Labels
[ ax , h1 ] = suplabel ( labels . XLabel ) ;
[ ax , h2 ] = suplabel ( labels . YLabel_left , ' y ' ) ;
432 set ( h1 , ' FontSize ' , 10 )
set ( h2 , ' Font Size ' , 10 )
h1 . Position ( 2 ) = 0 ;
h2 . Position ( 1 ) = 0 ;
if ~isempty ( labels . YLabel_right )
437 [ ax , h3 ] = suplabel ( labels . YLabel_right , ' yy ' ) ;
set ( h3 , ' FontSize ' , 10 )
h3 . Position ( 1 ) = 1 ;
end
filename_plot_wo_fmt = [ prob_opt_model ( i_cs ) .
    data_management . analysis_filename_prefix ...
442 plots . plot_def ( i_plot ) . filename_string ...
    i_X_label ] ;
else
    % Delete legend
    if isfield ( handles , ' legend ' )
447 handles . legend . delete ;
    end
end
end

```

```

end % per_case_and_collective_plot
%%Print plot
452 if per_case_and_collective_plot == 0 || (
    per_case_and_collective_plot && i_cs == n_case_plot)
    % Loop over output formats
    for i_print = 1:numel(plots.print_pars)
        print(plots.print_pars(i_print).file_format_str,...
            [filename_plot_wo_fmt input_siom.data_management.
            statistical_inference.id_inference_suffix ...
457         '.' plots.print_pars(i_print).file_format],...
            plots.print_pars(i_print).resolution_dpi_str);
    end
end % Print
end % Case loop
462 end % X loop
end % Plots loop
%%SAVE DATA
% Append input to results structure
% opt_inference.input_siom = input_siom;
467 % if input_siom.save.write_results
%     cd(input_siom.save.dir_saved_results)
%     % theta samples
%     filename_master_file = [input_siom.save.filename_prefix '
        _inference_master_results_file'...
%         '_' datestr(now,'yyyy-mm-dd-HH:MM:SS') '.mat'];
472 %     opt_inference.filename_master_file = filename_master_file;
%     save(filename_master_file,'opt_inference','-v7.3','-nocompression')
% end

```

## a.1.2 Específicas para problemas de programación lineal

### a.1.2.1 Muestreo espacio de soluciones

```

1 function [samples_model_omega_given_theta , sampling_output , optimum_value ,
    optimum_output] ...
    = sampling_linear_opt_model(theta , linear_constraints_as_theta_fun , q , n ,
        input_struct , varargin)
linear_constraints = linear_constraints_as_theta_fun(theta);
%TRANSFER FUNCTION AS A CONSTRAINT
if q == 0
6     % Null model, just constraints
else
    %Add linear constraint
    F_f_theta = varargin{1};
    linear_constraints = input_struct.linear_objective_as_constraint_fun(
        theta , q , F_f_theta , linear_constraints , input_struct);
11 end

```

```

%CONSTRAINTS ON STANDARD FORM
[ linear_constraints_mat_form ] = linear_constraints_to_matrix(
    linear_constraints);
%SAMPLING CANDIDATE SOLUTION SPACE
% Obtain samples from null model using MCMC (xsample function of limsolve R
package)
16 if isfield(input_struct, 'compute_optimal') && input_struct.compute_optimal
    F_f_theta = varargin{1};
    linear_constraints_t = input_struct.linear_objective_as_constraint_fun(
        theta, q, F_f_theta, linear_constraints, input_struct);
    switch linear_constraints_t.type {end}
        case 'inequality - min'
21         input_struct.objective_fun_vector = - linear_constraints_t.
            equation(end,:);
        case 'inequality - max'
            input_struct.objective_fun_vector = linear_constraints_t.equation
                (end,:);
    end
end
26 [ samples_model_omega_given_theta, sampling_output, optimum_value, optimum_output
    ] = ...
    sampling_linear_with_xsample_R(linear_constraints_mat_form, n, input_struct
    );

```

```

function [linear_mat_form] = linear_constraints_to_matrix(constraints)
% Transforms a linear constraint problem into a standard formulation
3 linear_mat_form.A = [];
  linear_mat_form.b = [];
  linear_mat_form.sdb = [];
  linear_mat_form.E = [];
  linear_mat_form.f = [];
8 linear_mat_form.G = [];
  linear_mat_form.h = [];
  for i=1:length(constraints.value)
    switch constraints.type {i}
        case 'equality - with error'
13         linear_mat_form.A = [linear_mat_form.A ; constraints.equation(i,:);
            ];
            linear_mat_form.b = [linear_mat_form.b ; constraints.value(i)];
            linear_mat_form.sdb = [linear_mat_form.sdb ; constraints.sd(i)];
        case 'equality'
            linear_mat_form.E = [linear_mat_form.E ; constraints.equation(i,:);
                ];
18         linear_mat_form.f = [linear_mat_form.f; constraints.value(i)];
        case 'inequality - max'
            linear_mat_form.G = [linear_mat_form.G ; -constraints.equation(i,:);
                ];
            linear_mat_form.h = [linear_mat_form.h; -constraints.value(i)];
        case 'inequality - min'

```

```

23         linear_mat_form .G=[linear_mat_form .G ; constraints .equation ( i ,:)
           ];
           linear_mat_form .h=[linear_mat_form .h ; constraints .value ( i )];
       end
end
% Constraints on parameter search region
28 for i=1:size ( constraints .equation ,2)
       switch constraints .parameter_search_region { i}
           case 'positive '
               line_G=zeros (1 ,size ( constraints .equation ,2));
               line_G ( i)=1;
33         linear_mat_form .G=[linear_mat_form .G ; line_G];
           linear_mat_form .h=[linear_mat_form .h ;0];
           case 'negative '
               line_G=zeros (1 ,size ( constraints .equation ,2));
               line_G ( i)=-1;
38         linear_mat_form .G=[linear_mat_form .G ; line_G];
           linear_mat_form .h=[linear_mat_form .h ;0];
           case 'proportion '
               line_G_max=zeros (1 ,size ( constraints .equation ,2));
               line_G_max ( i)=-1;
43         line_G_min=zeros (1 ,size ( constraints .equation ,2));
               line_G_min ( i)=1;
               linear_mat_form .G=[linear_mat_form .G ; line_G_max ; line_G_min];
               linear_mat_form .h=[linear_mat_form .h ;0 ;0];
           end
       end
48 end

```

```

function [ samples , sampling_output , optimum_value , optimum_output ] ...
2     = sampling_linear_with_xsample_R (linear_mat_form , max_iter , pars) %#ok<
       INUSL>
% This function is just a link to R (GNU R), where the xsample function of
% the limSolve R package is used to uniformly sample the available
% parameter subspace of solutions.
% saveR, by Jeroen Janssens (http://www.jeroenjanssens.com), is used to
7 % save Matlab data in R format, and R.matlab R library to save data back
% from R in a Matlab format
%%SAVE DATA FOR R
cd ( pars .temp_dir )
warning ( 'off' )
12 delete ( 'nmot_results .mat', 'optimization .R' );
info_to_pass = [];
mat_fields = fields (linear_mat_form );
for i =1:length ( mat_fields )
       if isempty ( eval ( ['linear_mat_form .' mat_fields { i}]))
17         eval ( [mat_fields { i} '=NaN;' ] );
       else
           eval ( [ mat_fields { i} '=linear_mat_form .' mat_fields { i} ';' ] );
       end
end

```

```

end
22 if isfield(pars, 'compute_optimal')
    compute_optimal = pars.compute_optimal;
else
    compute_optimal = 0;
end
27 if compute_optimal
    objective_fun_vector = pars.objective_fun_vector;
else
    objective_fun_vector = NaN; %#ok<*NASGU>
end
32 saveR('optimization.R', 'A', 'b', 'sdb', 'E', 'f', 'G', 'h', 'max_iter', '
    compute_optimal', 'objective_fun_vector');
%%Run R funtion based on xsample from limSolve package
% Get the current directory and switch slashes (R file pathes require '\')
pars.code_dir=strrep(pars.code_dir, '\', '/');
% Run script in R (note that the file path cannot contain any spaces),
37 % the paths could also be written out directly for each script
[~, sampling_output] = system(['pars.R_path_dir /Rscript' pars.code_dir '/
    xsample_linear_sampling.R']);
% Check if xsample run as expected
if numel(dir(['nmot_results.mat']))==0
42     samples = [];
    optimal_value = [];
else
    load('nmot_results.mat') %#ok<LOAD>
    %%Checking that samples fullfil constraints
    if isnumeric(E) && (size(E,1)+size(E,2))==2 && E==0
47     else
        checking.equality_constraints=samples*(E'); %#ok<NODEF>
        checking.equality_constraints=bsxfun(@eq, checking.equality_constraints,
            f');
    end
    if isnumeric(G) && (size(G,1)+size(G,2))==2 && G==0
52     else
        checking.inequality_constraints=samples*(G');
        checking.inequality_constraints=bsxfun(@ge, checking.
            inequality_constraints, h');
    end
    if (sum(checking.inequality_constraints==0)>0)
57     error('sampling_linear_with_xsample_R failed checking constraints
        satisfaction of candidate solution space samples.')
    end
    % else
    %     error('xsample call from sampling_linear_with_xsample_R failed to
        produce expected results.')
    % end
62 end
delete('nmot_results.mat', 'optimization.R');

```

```

if ~isempty ( optimal_value )
    optimum_value = ( optimal_value .X)';
    optimum_output = optimal_value ;
67 else
    optimum_value = [];
    optimum_output = [];
end

#Load, or previously install required packages
if (!require ('R.matlab')) install.packages ('R.matlab');
if (!require ('limSolve')) install.packages ('limSolve');
library (R.matlab)
5 library (limSolve)
#Load data saved from Matlab
source ('optimization .R')
#Make empty variables NULL
if ( is.na (A[1])) {A = NULL}
10 if ( is.na (b [1])) {b = NULL}
if ( is.na (E [1])) {E = NULL}
if ( is.na (f [1])) {f = NULL}
if ( is.na (G [1])) {G = NULL}
if ( is.na (h [1])) {h = NULL}
15 # Uniformly sample the available parameter subspace according to constraints
using xsample
nmot_samples <- xsample (A=A,B=b,E=E,F=f,G=G,H=h,iter=max_iter)
samples <- nmot_samples$X
if (compute_optimal) {
    optimal_value <- linp (E=E,F=f,G=G,H=h, Cost = objective_fun_vector)
20 } else {
    optimal_value = NULL
}
#Save results to Matlab data file
writeMat ('nmot_results .mat', samples=samples , optimal_value=optimal_value)

```

### a.1.3 Específicas del análisis de Belovsky (1986b)

```

1 %%SCRIPT CAPITULO 4 – Belovsky (1986)
%%LISTA DE FUNCIONES ESPECIFICAS Belovsky (1986)
% belovsky_1986_theta_sampling
% belovsky_1986_linear_constraints_fun
% belovsky_1986_objective_function
6 % belovsky_1986_objective_as_constraint_fun
% belovsky_1986_observed_variable_fun
clearflow_control
dbstop iferror
warning ('off')
11 n_Theta_prob_model = 2;

```

```

define_level_analysis.fun = {@mc_sampling_gral_optimization_model @
    radon_nikodym_density_random_var_gral_opt_model };
define_level_analysis.field_id_string = { 'mc_sampling '
    Radon_Nikodym_density_estimation_prob_opt_model ' };
define_level_analysis.id_string = { 'MC_sampling_prob_opt_mod '
    Radon_Nikodym_density_estimation_prob_opt_model ' };
define_level_analysis.id_display_string = { 'Monte Carlo sampling of general
16    optimization model ' ...
    'Radon Nikodym density estimation of reandom variable(s) of general
    optimization model ' };
flow_control.name_filelist_saved_data_string = { '
    Mc_sampling_gral_opt_model_filelist_saved_data . csv ' , ...
    ' Random_nykodym_density_estimation_gral_opt_model_filelist_saved_data . csv
    ' };
flow_control.re_do_analysis_from_level = 999;
flow_control.start_from_saved_analysis.dir = ...
21    '/mnt/2018_9tb_raid1_x2/Santiago/opt_big_data_files/Belovsky_19862019
    -01 -17 -19:25:03/Radon-Nikodym_density_X_2019 -01 -22 -07:35:03 ' ;
flow_control.start_from_saved_analysis.level_num = 2;
% DIRS
% Server 256Gb
base_dir= '/home/santiago/Dropbox/ Tesis/Tesis_repo_SRD_MAC/Contenido/
    Cap_4_ap_forrajeo ' ;
26 pars.dir.results = [base_dir '/Productos/Preliminares ' ] ;
pars.dir.save_big_data_files = '/mnt/2018_9tb_raid1_x2/Santiago/
    opt_big_data_files ' ;
pars.dir.code = [base_dir '/Codigo/Funciones ' ] ;
pars.dir.R_path = '/usr/bin ' ;
pars.dir.data = [base_dir '/Datos/Publicaciones/Belovsky1986 ' ] ;
31 pars.dir.temp = '/home/santiago/Tmp ' ;
m_theta = 10^3;
[P_Theta_model(1:n_Theta_prob_model).m_Theta_analisis] = deal(1,m_theta);
[P_Theta_model(1:n_Theta_prob_model).n_per_theta_null_model] = deal
    (10^7,10^6);
Q.grid = num2cell([0:0.01:0.99 0.999 0.9999]);
36 [P_Theta_model(1:n_Theta_prob_model).n_per_theta_opt_model] = deal(10^7,10^4)
    ;
addpath(genpath(pars.dir.code))
%Nombre de las especies
cd(pars.dir.data)
Belovsky_1986_table_data = readtable('Belovsky_1986_average_values.csv','
    Delimiter',';');
41 cases_label_cell = Belovsky_1986_table_data.Especie;
observations_cell = num2cell(Belovsky_1986_table_data.monocot_perc);
theta_prob_model_string = { 'Promedio ' 'Z_indep ' };
analysis_id_prefix = 'Belovsky_1986 ' ;
label_now_dirs = datestr(now,'yyyy-mm-dd-HH:MM:SS');
46 %%CHECK WHETHER THERE IS AVAILABLE SAVED DATA FROM PREVIOUS RUNS
flow_control.zero_level_dir = pars.dir.save_big_data_files;

```



```

%MONTE CARLO SAMPLING INPUT DATA
%MANEJO DE DATOS
input_som ( i_sc , i_TPm ) . data_management . studied_system_id_string =
    cases_label_cell { i_sc }; %ok<SAGROW>
input_som ( i_sc , i_TPm ) . data_management . filename_prefix = ...
91     [ analysis_id_prefix '_' cases_label_cell { i_sc } '_' ...
        '_Theta_prob_model_' theta_prob_model_string { i_TPm } ];
input_som ( i_sc , i_TPm ) . data_management . analysis_filename_prefix = ...
    [ analysis_id_prefix '_Theta_prob_model_' theta_prob_model_string {
        i_TPm } ];
input_som ( i_sc , i_TPm ) . data_management . mc_sampling . write_results = 1;
96 input_som ( i_sc , i_TPm ) . data_management . mc_sampling . dir_saved_results =
    flow_control . dir_saved_results { 1 };
input_som ( i_sc , i_TPm ) . data_management . mc_sampling .
    common_dir_saved_results = ...
    input_som ( i_sc , i_TPm ) . data_management . mc_sampling .
        dir_saved_results ;
input_som ( i_sc , i_TPm ) . data_management . mc_sampling . data2save . F_theta
    = 1;
input_som ( i_sc , i_TPm ) . data_management . mc_sampling . data2save .
    null_model_samples = 1;
101 input_som ( i_sc , i_TPm ) . data_management . mc_sampling . omega_samples_data .
    format = 'numeric_array' ;
input_som ( i_sc , i_TPm ) . data_management . mc_sampling . omega_samples_data .
    precision = 'double' ;
%Omega (ESPACIO MUESTREAL) MODELO NULO
input_som ( i_sc , i_TPm ) . null_model . Omega_mc_sampling . fun = @
    sampling_linear_opt_model ;
input_som ( i_sc , i_TPm ) . null_model . Omega_mc_sampling . n_per_theta = ...
106 P_theta_model ( i_TPm ) . n_per_theta_opt_model ;
% Parametros específicos de la función de muestreo por Monte Carlo
input_som ( i_sc , i_TPm ) . null_model . Omega_mc_sampling . sampling_pars .
    code_dir = pars . dir . code ;
input_som ( i_sc , i_TPm ) . null_model . Omega_mc_sampling . sampling_pars .
    temp_dir = pars . dir . temp ;
input_som ( i_sc , i_TPm ) . null_model . Omega_mc_sampling . sampling_pars .
    R_path_dir = pars . dir . R_path ;
111 %c (RESTRICCIONES)
input_som ( i_sc , i_TPm ) . c . as_theta_fun = @
    belovsky_1986_linear_constraints_fun ;
%Q (ESPACIO DE PARAMETROS FUNCION DE TRANSFERENCIA)
input_som ( i_sc , i_TPm ) . Q = Q;
input_som ( i_sc , i_TPm ) . Q . q_null_model = 0; % Funcion escalonada , $q_N
    = 0$
116 input_som ( i_sc , i_TPm ) . Q . i_q_null_model = 1;
input_som ( i_sc , i_TPm ) . Q . dimensionality = 1;
%Theta (ESPACIO DE PARAMETROS)
input_som ( i_sc , i_TPm ) . Theta . mc_sampling . fun = @
    belovsky_1986_theta_sampling ;

```



```

% %DENSIDAD DE PROBABILIDAD FUNCION ALEATORIA OBSERVADA
% Data management
input_pdom(i_sc,i_TPm).data_management.radon_nikodym_density.
    dir_saved_results = ...
    flow_control.dir_saved_results{2};
156 input_pdom(i_sc,i_TPm).data_management.radon_nikodym_density.
    common_dir_saved_results = ...
    flow_control.dir_saved_results{2};
input_pdom(i_sc,i_TPm).data_management.radon_nikodym_density.
    write_results = 1;
% Funcion aleatoria observada
input_pdom(i_sc,i_TPm).X(1).fun = @
    belovsky_1986_observed_variable_fun;
161 input_pdom(i_sc,i_TPm).X(1).label = 'Porcentaje de monocotiled neas
    en la dieta';
input_pdom(i_sc,i_TPm).X(1).variable_class = 'continuous';
input_pdom(i_sc,i_TPm).X(1).dimensionality = 'univariate';
input_pdom(i_sc,i_TPm).X(1).XLim = num2cell([0 1]);
input_pdom(i_sc,i_TPm).X(1).rn_denisty_estimation.fun = @
    univariate_pdf_kernel_estimation;
166 % Parametros estimacion
input_pdom(i_sc,i_TPm).X(1).rn_denisty_estimation.pars.min_bound=0;
input_pdom(i_sc,i_TPm).X(1).rn_denisty_estimation.pars.max_bound=1;
input_pdom(i_sc,i_TPm).X(1).rn_denisty_estimation.pars.
    tolerance_to_bound = 10^-3;
input_pdom(i_sc,i_TPm).X(1).rn_denisty_estimation.pars.
    intervals_kernel_estimation = 2^7;
171 input_pdom(i_sc,i_TPm).X(1).rn_denisty_estimation.pars.
    empirical_hist_division = 10^3;
input_pdom(i_sc,i_TPm).X(1).rn_denisty_estimation.pars.
    perc_rep_extreme_data = 0.03;
input_pdom(i_sc,i_TPm).X(1).rn_denisty_estimation.pars.
    data_perc_to_repeat = 0.001;
% Estimacion verosimilitud
input_pdom(i_sc,i_TPm).X(1).rn_density_at_value.fun = @
    univariate_pdf_interpolation;
176 input_pdom(i_sc,i_TPm).X(1).rn_density_at_value.pars = [];
end
end
% %OBTENCION MUESTRAS POR MONTE CARLO DEL ESPACIO DE SOLUCIONES MODELO
GENERAL DE OPTIMIZACION
% %Check whether analysis are to be done, loaded from saved date , or skipped
since
181 % subsequent data of posterior analysis is available or data is already in
% workspace memory
for i_study_case = 1:numel(cases_label_cell)
    for i_Theta_prob_model = 1:numel(P_Theta_model)
        for i_step_analysis = 1:numel(define_level_analysis.fun)

```

```

186 display ([ 'Analysis #' num2str(i_step_analysis) ': '
           define_level_analysis.id_display_string{i_step_analysis}])
% Check if analysis are already on workspace memory
valid_analysis_in_memory = 0;
191 if exist('prob_opt_model','var') && ...
           size(prob_opt_model,1) >= numel(cases_label_cell) && ...
           size(prob_opt_model,2) >= numel(P_Theta_model)
           % See to what directory the data on workspace memory is
           % linked to
           field_level_string = define_level_analysis.field_id_string{
           i_step_analysis};
           if isfield(prob_opt_model(i_study_case,i_Theta_prob_model).
           data_management,field_level_string)
196             current_level_case_in_memory_dir = ...
                 prob_opt_model(i_study_case,i_Theta_prob_model).
                 data_management.(field_level_string).
                 dir_saved_results;
                 valid_analysis_in_memory = strcmp(
                 current_level_case_in_memory_dir,flow_control .
                 dir_saved_results{i_step_analysis});
                 if ~ valid_analysis_in_memory
                 disp('Case in memory but does not match directory
                 specified by user or automatically assigned by
                 script')
201             end
           end
           end
           % Determine action to be made
           if valid_analysis_in_memory
206             action_for_current_analysis = 'on_memory';
           else
           % Check last saved version
           last_saved_step = flow_control.scanned_saved_data .
           last_saved_data_num(i_study_case,i_Theta_prob_model);
           if last_saved_step > i_step_analysis
211             action_for_current_analysis = 'skip';
           else
           if last_saved_step == i_step_analysis
                 action_for_current_analysis = 'load';
           else
216             action_for_current_analysis = 'compute';
           end
           end
           end
           % Do action ...
221 disp(['Model ' num2str(i_study_case) '-' num2str(
           i_Theta_prob_model) ': action ' action_for_current_analysis])
           string_disp_pref = ['Model ' num2str(i_study_case) '-' num2str(
           i_Theta_prob_model) ...

```

```

    ' ' define_level_analysis.id_display_string{i_step_analysis}
    ''];
switch action_for_current_analysis
case 'on_memory'
226     disp([string_disp_pref 'already on workspace memory,
        nothing needs to be done.'])
case 'skip'
    disp([string_disp_pref 'not on workspace memory, but
        saved data of analysisf higher hierarchy are
        available(so nothing needs to be done)'])
case 'load'
    disp([string_disp_pref 'loading from save data from
        previous analysisinitiated...'])
231     init = toc;
        flow_control.current_case_subindex = [i_study_case
            i_Theta_prob_model];
        flow_control.current_level_num = i_step_analysis;
        prob_opt_model(i_study_case, i_Theta_prob_model) =
            save_or_load_hier_analyzes('load', flow_control);
        ended = toc;
236     disp([string_disp_pref 'loading from save data from
        previous analysis finished, took ' num2str(ended-
            init) 'seconds.'])
case 'compute'
    % EJECUCION FUNCION GENERAL ESPACIO DE PARAMETROS THETA Y
    MUESTREO ESPACIO MUESTRAL OMEGA
    disp([string_disp_pref ' computing initiated ... ' ])
    [prob_opt_model(i_study_case, i_Theta_prob_model)] = ...
241     mc_sampling_gral_optimization_model(input_som(
        i_study_case, P_Theta_model));
    disp([string_disp_pref ' computing finished!' ])
    %GUARDAR INFORMACION PARA RETOMAR ANALISIS
    list_of_masterfiles_mc_sampling = flow_control.
        scanned_saved_data.
        file_list_saved_data_array_within_cell{1};
    if isfield(prob_opt_model(i_study_case, i_Theta_prob_model)
246     ).data_management.mc_sampling, 'filename_master_file')
        master_mc_sampling_filenames_list{i_study_case,
            i_Theta_prob_model} = ...
            prob_opt_model(i_study_case, i_Theta_prob_model).
                data_management.mc_sampling.
                    filename_master_file;
        cd(pars.dir.save_big_data_files)
        mc_sampling_master_list_file = [prob_opt_model(
            i_study_case, i_Theta_prob_model).data_management.
                mc_sampling.common_dir_saved_results ...
                    '/Sampling_gral_opt_model_global_results_list.mat
                    '];
251     % Save list

```

```

save ( mc_sampling_master_list_file , '
        master_mc_sampling_filenames_list ','-v7.3','-
        nocompression ');
% Save list filename
save ( ' mc_sampling_master_list_file .mat' , '
        mc_sampling_master_list_file ','-v7.3','-
        nocompression ');
end
256     end % switch
        end % step
    end
end
%%INFERENCIA ESTADISTICA
261 %Supuesto : las observaciones son independientes
cd ( pars . dir . data )
Belovsky_1986_table_data = readtable ( ' Belovsky1986_average_values . csv ' , '
    Delimiter ' , ';' );
x = cellfun (@ num2cell , num2cell ( Belovsky_1986_table_data . monocot_perc ) , '
    UniformOutput ' , false );
input_siom . data_management . statistical_inference . shared_id_stamp = datestr (
    now , ' yyyy-mm-dd-HH:MM:SS ' );
266 input_siom . inference . q_value_to_index_fun = @ q_value_to_index ;
input_siom . inference . q_to_compare = { 0.5 0.95 };
input_siom . inference . q_to_label_fun = @( q ) ( [ '\ it { q } = ' num2str ( q ) ] );
input_siom . inference . likelihood_ratio_test . q_fixed = 0.95 ;
%% %% %% P%OTS
271 % Print formats
plots . print_pars ( 1 ) . file_format_str = '-djpeg ' ;
plots . print_pars ( 1 ) . file_format = ' jpg ' ;
plots . print_pars ( 1 ) . resolution_dpi_str = '-r1200 ' ;
plots . print_pars ( 1 ) . file_format_str = '-dpng ' ;
276 plots . print_pars ( 1 ) . file_format = ' png ' ;
plots . print_pars ( 1 ) . resolution_dpi_str = '-r1200 ' ;
plots . print_pars ( 2 ) . file_format_str = '-depsc ' ;
plots . print_pars ( 2 ) . file_format = ' eps ' ;
plots . print_pars ( 2 ) . resolution_dpi_str = '' ;
281 % Definition of plots
plot_def = struct ( [] ) ;
% Radon_Nikodym density of models vs observed
plot_def ( numel ( plot_def ) + 1 ) . fun = @
    plot_gral_opt_model_plot_per_X_per_case_RN_density_vs_observed ;
plot_def ( end ) . filename_string = '- Radon_Nikodym density of models vs
    observed ' ;
286 plot_def ( end ) . is_per_X = 1 ;
plot_def ( end ) . is_per_case = 1 ;
plot_def ( end ) . individual_plots = 0 ; % Only relevant if is_per_case == 1
plot_def ( end ) . n_column_subplots = 3 ; % Only relevant if is_per_case == 1 AND
    individual_plots = 0
plot_def ( end ) . secondYaxis = 1 ;

```

```

291 plot_def(end).style.LineStyle = {'-' '-' '-'};
plot_def(end).style.color = {[1 0 0] [0 0 1] [0 0 0]};
plot_def(end).style.color_attenuation_q_to_compare = cell 2 mat(input_siom.
    inference.q_to_compare);
plot_def(end).style.Marker = {'none' 'none' 'none'};
plot_def(end).style.MarkerFaceColor = {'none' 'none' 'none'};
296 plot_def(end).style.MarkerSize = {2 2 2};
plot_def(end).observed.style.color = {[0 0.8 0]};
plot_def(end).observed.XLim = [0 1];
% Global likelihood
plot_def(numel(plot_def)+1).fun=@plot_gral_opt_model_plot_likelihood;
301 plot_def(end).filename_string = '- Global same q Likelihood';
plot_def(end).is_per_X = 0;
plot_def(end).is_per_case = 0;
plot_def(end).YScale = 'log';
plot_def(end).secondYaxis = 0;
306 plot_def(end).style.LineStyle = {'-' '-' 'none'};
plot_def(end).style.color = {[1 0 0] [0 0 1] [0 0 0]};
plot_def(end).style.Marker = {'s' 's' 's'};
plot_def(end).style.MarkerFaceColor = {'none' 'none' [0 0 0]};
plot_def(end).style.MarkerSize = {2 2 3};
311 % Likelihood per case
plot_def(numel(plot_def)+1).fun=@plot_gral_opt_model_plot_likelihood;
plot_def(end).filename_string = '- per case Likelihood';
plot_def(end).is_per_X = 0;
plot_def(end).is_per_case = 1;
316 plot_def(end).individual_plots=0;
plot_def(end).n_column_subplots=3;
plot_def(end).YScale = 'linear';
plot_def(end).secondYaxis = 0;
plot_def(end).style.LineStyle = {'-' '-' 'none'};
321 plot_def(end).style.color = {[1 0 0] [0 0 1] [0 0 0]};
plot_def(end).style.Marker = {'none' 'none' 's'};
plot_def(end).style.MarkerFaceColor = {'none' 'none' [0 0 0]};
plot_def(end).style.MarkerSize = {2 2 2};
%%
326 plots.plot_def = plot_def;
%%Format of plots
plots.format.ind.axis_font_size = 10;
plots.format.ind.marker_size = 6;
plots.format.xylabls.font_size = 13;
331 % Labels
options.language = 'Castellano';
switch options.language
    case 'Castellano'
        labels.q = 'Par metro de optimizaci n \it{q}';
336 labels.likelihood='Verosimilitud';
labels.radon_nikodym_density = 'Densidad de probabilidad';
labels.null_model='Modelo nulo';

```

```

    labels.alternative_optimization_models = {'M x . ingesta de energia '
        'M n . tiempo de forrajeo '};
    labels.observed = 'Observaciones ';
341 labels.observed_histogram = 'Observaciones (frecuencia) ';
    case 'English '
        labels.q = '\it{q} optimization parameter ';
        labels.likelihood = 'Likelihood ';
        labels.radon_nikodym_density = 'Probability density ';
346 labels.null_model = 'Null model ';
        labels.alternative_optimization_models = {'M x . ingesta de energia '
            'M n . tiempo de forrajeo '};
        labels.observed = 'Observations ';
        labels.observed_histogram = 'Observations (frequency) ';
    end
351 plots.labels = labels ;
    input_siom.inference.plots = plots ;
    % Inferencia sobre todos los casos
    input_siom.data_management.statistical_inference.id_inference_suffix = '
        _completo ' ;
    for i_Theta_prob_model = 1:numel(P_Theta_model)
356 [stats_def {1}{i_Theta_prob_model} , lnL {i_Theta_prob_model}] = ...
        statistical_inference_gral_opt_model (...
            prob_opt_model (: , i_Theta_prob_model) , ...
            x (: , : ) , ...
            input_siom ) ;
361 end
    % Inferencia excluyendo la especie Dissoteria carolina , ya que la densidad de
    % probabilidad
    % del modelo nulo es nula para la variable aleatoria observada
    input_siom.data_management.statistical_inference.id_inference_suffix = '
        _exc_esp_1 ' ;
    for i_Theta_prob_model = 1:numel(P_Theta_model)
366 [stats_def {2}{i_Theta_prob_model} , lnL {i_Theta_prob_model}] = ...
        statistical_inference_gral_opt_model (...
            prob_opt_model (2 : end , i_Theta_prob_model) , ...
            x (2 : end , : ) , ...
            input_siom ) ;
371 end

```

```

function [samples_theta , info_output] = belovsky_1986_theta_sampling (m, pars )
cd (pars . data . dir )
% Mean
4 data_mean_values_table = readtable ('Belovsky 1986_average_values . csv ', '
    Delimiter ' , ';' );
mean_theta = permute (table2struct (data_mean_values_table) , [2 3 1]);
mean_theta = mean_theta (: , : , pars . i_species );
% Standard error
data_standard_error_values_table = readtable ('
    Belovsky 1986_standar_errors_values . csv ', 'Delimiter ' , ';' );

```

```

9  standard_error_theta = permute( table2struct( data_standard_error_values_table )
    ,[2 3 1]);
standard_error_theta = standard_error_theta (:, :, pars.i_species);
% Preallocate
valid_theta = zeros ([1 m]);
if ~isempty( pars . prob_dist . parameter_groups . fun_per_group )
14  m_draws = m;
    cont = 0;
    while sum( valid_theta ) < m
        cont = cont + 1;
        candidate_samples_theta ( 1:m_draws ) = mean_theta ;
19  id_groups = unique ( pars . prob_dist . parameter_groups . definition_array );
        for i_g = 1:numel(id_groups) %loop over parameters groups
            dist_fun = pars . prob_dist . parameter_groups . fun_per_group {i_g};
            if ~isempty( dist_fun )
24  % Sample from specified univariate distribution and retain
                % indicated proportion of central distribution samples
                if m_draws > 50
                    m_mc_samples = max(12, ceil (m_draws./ pars . prob_dist .
                        parameter_groups . retain_central_prop));
                    mc_dist = dist_fun ([ m_mc_samples 1]);
                    [~, index_sorted_mc_dist] = sort(mc_dist, 1);
29  i_init = floor(m_mc_samples*(1 - pars . prob_dist .
                        parameter_groups . retain_central_prop)/2);
                    mc_dist (index_sorted_mc_dist ([1:i_init-1 i_init+m_draws:
                        m_mc_samples])) = [];
                else
                    mc_dist = dist_fun ([m_draws 1]);
                end
34  pars_in_group = pars . prob_dist . parameters ( find ( pars . prob_dist
                    . parameter_groups . definition_array )); %#ok<FNDSB>
                for i_par = 1:numel( pars_in_group ) %loop over parameters
                    within group
                    mc_par_values = mean_theta .( pars_in_group {i_par} ) + ...
                        mc_dist * standard_error_theta .( pars_in_group {i_par} )
                        ;
                    if sum( mc_par_values < 0 ) > 0
39  error ( 'Parameters out of range ' )
                    end
                    mc_par_cell = num2cell ( mc_par_values );
                    [ candidate_samples_theta (:). ( pars_in_group {i_par} ) ] =
                        deal ( mc_par_cell { : } );
                end
44  end
            end
        end
        valid_candidate = false ( size ( candidate_samples_theta ) );
        % Check if theta produce a Omega \times theta set with non null
        % measure
49  parfor i = 1:numel( candidate_samples_theta )

```

```

    % % Paralell computing
    sampling_pars = pars.check.sampling_pars; %#ok<PFBNS>
    %Temp directory for each worker
    sampling_pars.temp_dir = [sampling_pars.temp_dir ' /theta ' num2str
54     (i) '_' datestr(now,'yyyymmdd-HH:MM:SS.FFF')];
    cd(pars.check.sampling_pars.temp_dir)
    if numel(dir(sampling_pars.temp_dir))==0
        mkdir(sampling_pars.temp_dir)
    end
    [samples_omega_for_theta {i}, a {i}, b {i}] = pars.check .
59     sampling_model.fun (...
        candidate_samples_theta (i), ...
        pars.check.c.as_theta_fun, ...
        pars.check.q_null_model, ...
        10, ...
        sampling_pars);
64     valid_candidate (i) = ~isempty (samples_omega_for_theta {i});
    rmdir(sampling_pars.temp_dir)
end
% Store correct results
if cont == 1
69     samples_theta = candidate_samples_theta (valid_candidate (:));
    else
        samples_theta = [samples_theta candidate_samples_theta (
            valid_candidate)];
    end
    valid_theta (1:numel(samples_theta)) = true;
74     if sum(valid_theta) < m
        m_draws = ceil((m - sum(valid_theta)) * ((numel(valid_candidate)
            / sum(valid_candidate)) * 1.2));
    else
        samples_theta = samples_theta (1:m);
    end
79     clear candidate_samples_theta
end % while
else
    samples_theta (1:m) = mean_theta;
end
84 info_output = [];
samples_theta_st = samples_theta;
samples_theta = cell (size(samples_theta_st));
for i = 1:numel(samples_theta)
    samples_theta {i} = samples_theta_st (i);
89 end
1 function [constraints] = belovsky_1986_linear_constraints_fun (theta)
% Constraints values
% Digestive capacity
constraints.value (1) = theta.digestive_v .* theta.digestive_t;

```

```

% Feeding time
6 constraints.value(2)=theta.feeding_time;
% Energy requirement
constraints.value(3)=theta.energy_requirement;
% Type of constraints
constraints.type={'inequality - max' , 'inequality - max' , 'inequality -
    min'};
11 % Constraints on parameters values
constraints.parameter_search_region={'positive' 'positive'};
% Constraints equation: each column represents a parameter to be varied
constraints.equation=[theta.bulk_monocot , ...
    theta.bulk_dicot ;
16 theta.cropping_monocot , ...
    theta.cropping_dicot ;...
    theta.energy_content_monocot , ...
    theta.energy_content_dicot ];

1 function [obj_fun_samples_model] = belovs_ky_1986_objective_function (
    samples_null_model_omega , samples_theta , input_struct)
switch input_struct
    case 'max_energy'
        %%INGESTED ENERGY
        parameters = ([[ samples_theta.energy_content_monocot ]; [
            samples_theta.energy_content_dicot ]])';
6    case 'min_time'
        %%FORAGING TIME
        parameters = ([[ samples_theta.cropping_monocot ]; [ samples_theta .
            cropping_dicot ]])';
end
obj_fun_samples_model = sum(bsxfun(@times , samples_null_model_omega , parameters
    ),2);

function [linear_constraints] = belovsky_1986_objective_as_constraint_fun (
    theta , q , F_f_theta , linear_constraints , input_struct)
% Find objective function value that correspond to the value of parameter q
f_q = F_f_theta.f_q;
Q_grid = F_f_theta.q;
5 f_for_q = f_q ( q == Q_grid );
switch input_struct.f.pars
    case 'max_energy'
        objective.type = 'inequality - min';
        objective.equation = [theta.energy_content_monocot , theta .
            energy_content_dicot ];
10    case 'min_time'
        objective.type = 'inequality - max';
        objective.equation = [theta.cropping_monocot , theta.cropping_dicot ];
end
% Add to linear constraints struct
15 linear_constraints.value(end+1) = f_for_q ;

```

```
linear_constraints.type(end+1) = { objective.type };
linear_constraints.equation(end+1,:) = objective.equation;
```

```
function [observed_variable_samples] = belovsky_1986_observed_variable_fun (
    samples_model_omega, samples_theta)
% Cell 2 numeric array to avoid slow computing due loops
3 samples_model_omega = cell2mat ( permute ( samples_model_omega , [ 1 3 2 ] ) );
% Parameter array
observed_variable_samples = squeeze ( samples_model_omega (:, 1, :) ./ sum(
    samples_model_omega, 2 ) );
% Retain original structure of samples_model_omega
observed_variable_samples = mat2cell ( observed_variable_samples , size (
    observed_variable_samples , 1 ) , ones ( [ numel ( samples_theta ) 1 ] ) );
```



## BIBLIOGRAFÍA

---

- Agutter, P. S., y J. A. Tuszynski. 2011. «Analytic theories of allometric scaling». *Journal of Experimental Biology* 214 (7): 1055-1062.
- Agutter, P. S., y D. N. Wheatley. 2004. «Metabolic scaling: consensus or controversy?» *Theoretical Biology and Medical Modelling* 1 (1): 13.
- Akaike, H. 1974. «A new look at the statistical model identification». *IEEE Transactions on Automatic Control* 19 (6): 716-723.
- Albeverio, S., y G. Torbin. 2004. «Image measures of infinite product measures and generalized Bernoulli convolutions». *Transactions of the National Pedagogical University (Phys.-Math. Sci.)* 228-241.
- Allsop, D. J., y S. A. West. 2003. «Life history: Changing sex at the same relative body size». *Nature* 425 (6960): 783.
- Amundson, R. 1994. «Two concepts of constraint: Adaptationism and the challenge from developmental biology». *Philosophy of Science* 61 (4): 556-578.
- Anderson, R. M. 1982. «Star-finite representations of measure spaces». *Transactions of the American Mathematical Society* 271 (2): 667-687.
- Ando, T. 2010. *Bayesian Model Selection and Statistical Modeling*. CRC Press.
- Andrews, P. W., S. W. Gangestad y D. Matthews. 2002a. «Adaptationism—how to carry out an exaptationist program». *Behavioral and Brain Sciences* 25 (4): 489-504.
- Andrews, P., S. Gangestad y D. Matthews. 2002b. «Adaptationism, exaptationism, and evolutionary behavioral science». *Behavioral and Brain Sciences* 25 (4): 534-553.
- Anten, N. P. R., y H. J. During. 2011. «Is analysing the nitrogen use at the plant canopy level a matter of choosing the right optimization criterion?» *Oecologia* 167 (2): 293-303.
- Anten, N. 2005. «Optimal characteristics of individual plants in vegetation stands and implications for species coexistence». *Annals of Botany* 95:495-506.
- Arnold, T. W. 2010. «Uninformative parameters and model selection using Akaike's Information Criterion». *Journal of Wildlife Management* 74 (6): 1175-1178.
- Arthur, W. 2001. «Developmental drive: an important determinant of the direction of phenotypic evolution». *Evolution & Development* 3 (4): 271-278.
- Athreya, K. B., y S. N. Lahiri. 2006. *Measure Theory and Probability Theory*. Springer Science & Business Media.

- Bacigalupe, L. D., y F. Bozinovic. 2002. «Design, limitations and sustained metabolic rate: lessons from small mammals». *Journal of Experimental Biology* 205 (19): 2963-2970.
- Ballesteros, F. J., y col. 2018. «On the thermodynamic origin of metabolic scaling». *Scientific Reports* 8 (1): 1448.
- Banavar, J. R., A. Maritan y A. Rinaldo. 1999. «Size and form in efficient transportation networks». *Nature* 399 (6732): 130-132.
- Batten, D., S. Salthe y F. Boschetti. 2008. «Visions of evolution: Self-organization proposes what natural selection disposes». *Biological Theory* 3 (1): 17-29.
- Batterman, R. W. 2000. «Multiple realizability and universality». *The British Journal for the Philosophy of Science* 51 (1): 115-145.
- . 2001. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford Studies in Philosophy of Science. Oxford University Press, USA.
- Beatty, J. 1987. «Natural selection and the null hypothesis». En *The Latest on the Best: Essays on Evolution and Optimality*, ed. por J. Dupre, 53-75. MIT Press.
- Belovsky, G. E. 1978. «Diet optimization in a generalist herbivore: The moose». *Theoretical Population Biology* 14 (1): 105-134.
- . 1981a. «Food plant selection by a generalist herbivore: The moose». *Ecology* 62 (4): 1020-1030.
- . 1981b. «Optimal activity times and habitat choice of moose». *Oecologia* 48 (1): 22-30.
- Belovsky, G. E. 1984a. «Herbivore Optimal Foraging: A Comparative Test of Three Models» [inlangEnglish]. *The American Naturalist* 124 (1):
- Belovsky, G. E. 1984b. «Snowshoe hare optimal foraging and its implications for population dynamics». *Theoretical Population Biology* 25 (3): 235-264.
- . 1984c. «Summer diet optimization by beaver». *American Midland Naturalist*: 209-222.
- . 1986a. «Generalist herbivore foraging and its role in competitive interactions». *American Zoologist* 26 (1): 51-69.
- . 1986b. «Optimal foraging and community structure: Implications for a guild of generalist grassland herbivores». *Oecologia* 70 (1): 35-52.
- . 1990. «A reply to Hobbs». En *Behavioural Mechanisms of Food Selection*, ed. por R. N. Hughes, 415-422. Springer.
- . 1994. «How good must models and data be in ecology?» *Oecologia* 100 (4): 475-480.
- Belovsky, G. E., y P. A. Jordan. 1978. «The time-energy budget of a moose». *Theoretical Population Biology* 14 (1): 76-104.

- Belovsky, G. E., y O. J. Schmitz. 1993. «Owen-Smith's evaluation of herbivore foraging models: What is constraining?» *Evolutionary Ecology* 7 (5): 525-529.
- Bengtsson, H. 2013. «R.matlab: Read and write of MAT files together with R-to-MATLAB connectivity. R package version 2.1.0.»
- Benjamini, Y. 2010. «Simultaneous and selective inference: Current successes and future challenges». *Biometrical Journal* 52 (6): 708-721.
- Berger, J. O., y R. L. Wolpert. 1988. *The Likelihood Principle (Second Edition)*. Institute of Mathematical Statistics.
- Berger, S. A., y col. 2006. «Light supply, plankton biomass, and seston stoichiometry in a gradient of lake mixing depths». *Limnology and Oceanography* 51 (4): 1898-1905.
- Biegler, L. T. 2010. *Concepts of Constrained Optimization*. Philadelphia, PA: Society for Industrial / Applied Mathematics.
- Birch, J. 2016. «Natural selection and the maximization of fitness». *Biological Reviews* 91 (3): 712-727.
- Boltzmann, L. 1886. «Der zweite Hauptsatz der mechanischen Wärmetheorie: Vortrag gehalten in der feierlichen Sitzung der Kaiserlichen Akademie der Wissenschaften am XXIX Mai MDCCCLXXXVI». 36:225-259.
- Bonarius, H. P., y col. 1996. «Metabolic flux analysis of hybridoma cells in different culture media using mass balances». *Biotechnology and bioengineering* 50 (3): 299-318.
- Botev, Z., J. Grotowski y D. Kroese. 2010. «Kernel density estimation via diffusion». *The Annals of Statistics* 38 (5): 2916-2957.
- Brown, J. H., G. B. West y B. J. Enquist. 2005. «Yes, West, Brown and Enquist's model of allometric scaling is both mathematically correct and biologically relevant». *Functional Ecology* 19 (4): 735-738.
- Brown, J. H., P. A. Marquet y M. L. Taper. 1993. «Evolution of body size: Consequences of an energetic definition of fitness». *The American Naturalist* 142 (4): 573-584.
- Bull, J., e I. Wang. 2010. «Optimality models in the age of experimental evolution and genomics». *Journal of Evolutionary Biology* 23 (9): 1820-1838.
- Burgard, A. P., y C. D. Maranas. 2003. «Optimization-based framework for inferring and testing hypothesized metabolic objective functions». *Biotechnology and Bioengineering* 82 (6): 670-677.
- Burnham, K., y D. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Verlag.
- Buston, P. M., P. L. Munday y R. R. Warner. 2004. «Evolutionary biology: Sex change and relative body size in animals». *Nature* 428 (6983).

- Cai, T. T., C. L. Montague y J. S. Davis. 2006. «The maximum power principle: An empirical investigation». *Ecological Modelling* 190 (3-4): 317-335.
- Callahan, M., R. L. Wears y E. Weber. 2002. «Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals». *Jama* 287 (21): 2847-2850.
- Callebaut, W. 1998. «Self-organization and optimization: conflicting or complementary approaches?» En *Evolutionary Systems: Biological and Epistemological Perspectives on Selection and Self-Organization*, ed. por G. van de Vijver, S. N. Salthe y M. Delpos, 79-100. Dordrecht: Springer Netherlands.
- Carlin, N. F., y D. S. Gladstein. 1992. «Reply to Houston, McNamara and Thompson». *Oikos* 63 (3): 518.
- Casella, G., y R. L. Berger. 2002. *Statistical Inference*. Vol. 2. Duxbury.
- Cashdan, E. 1983. «Comment». *Current Anthropology* 24 (5): 642-643.
- Cavanaugh, J. E., y col. 1997. «Unifying the derivations for the Akaike and corrected Akaike information criteria». *Statistics & Probability Letters* 33 (2): 201-208.
- Chamberlin, T. C. 1890. «The method of multiple working hypotheses». *Science* 15 (366): 92-96.
- Chandrasekaran, S., y N. D. Price. 2010. «Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*». *Proceedings of the National Academy of Sciences* 107 (41): 17845-17850.
- Chang, R. L., y col. 2011. «Metabolic network reconstruction of *Chlamydomonas* offers insight into light-driven algal metabolism». *Molecular Systems Biology* 7 (1): 518.
- Chapman, E. J., D. L. Childers y J. J. Vallino. 2015. «How the second law of thermodynamics has informed ecosystem ecology through its history». *BioScience* 66 (1): 27-39.
- Charnov, E. L. 1976a. «Optimal foraging: Attack strategy of a mantid». *The American Naturalist* 110 (971): 141-151.
- . 1976b. «Optimal foraging, the marginal value theorem». *Theoretical Population Biology* 9 (2): 129-136.
- Charnov, E. L. 1993. *Life History Invariants*. Oxford University Press.
- Charnov, E. L., y J. F. Downhower. 1995. «A trade-off-invariant life-history rule for optimal offspring size». *Nature* 376 (6539): 418-419.
- Charnov, E. L., y S. M. Ernest. 2006. «The offspring-size/clutch-size trade-off in mammals». *The American Naturalist* 167 (4): 578-582.
- Charnov, E. L., T. F. Turner y K. O. Winemiller. 2001. «Reproductive constraints and the evolution of life histories with indeterminate growth». *Proceedings of the National Academy of Sciences of the United States of America* 98 (16): 9460-9464.

- Chow, S. L. 1998. «Précis of statistical significance: Rationale, validity, and utility». *Behavioral and Brain Sciences* 21 (2): 169-194.
- Cichon, M., y J. Kozłowski. 2000. «Ageing and typical survivorship curves result from optimal resource allocation». *Evolutionary Ecology Research* 2 (7): 857-870.
- Çınlar, E. 2011. *Probability and Stochastics*. Vol. 261. Springer Science & Business Media.
- Cody, M. L. 1974. «Optimization in ecology». *Science* 183 (4130): 1156-1164.
- Colijn, C., y col. 2009. «Interpreting expression data with metabolic flux models: predicting Mycobacterium tuberculosis mycolic acid production». *PLoS computational biology* 5 (8): e1000489.
- Collette, Y., y P. Siarry. 2013. *Multiobjective optimization: principles and case studies*. Springer Science & Business Media.
- Collin, R. 2006. «Sex ratio, life-history invariants, and patterns of sex change in a family of protandrous gastropods». *Evolution* 60 (4): 735-745.
- Cover, T. M., y J. A. Thomas. 2012. *Elements of Information Theory*. John Wiley & Sons.
- Covert, M. W., y col. 2008. «Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli». *Bioinformatics* 24 (18): 2044-2050.
- Cox, D. R. 2006. *Principles of Statistical Inference*. Cambridge University Press.
- Cox, R. T. 1961. *Algebra of Probable Inference*. JHU Press.
- Coyne, J. A., N. H. Barton y M. Turelli. 1997. «Perspective: A critique of Sewall Wright's shifting balance theory of evolution». *Evolution* 51 (3): 643-671.
- Cutland, N. J. 2000. *Loeb Measures in Practice: Recent Advances*. Springer.
- Darwin, C. R. 1859. «On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life». *Murray, London*.
- Dawkins, R. 1976. «The Selfish Gene».
- Dekel, E., y U. Alon. 2005. «Optimality and evolutionary tuning of the expression level of a protein». *Nature* 436 (7050): 588-592.
- DeLong, J. P. 2008. «The maximum power principle predicts the outcomes of two-species competition experiments». *Oikos* 117 (9): 1329-1336.
- Depew, D. J., y B. H. Weber. 1996. *Darwinism Evolving: Systems Dynamics and the Genealogy of Natural Selection*. The MIT Press.
- Dewar, R. C. 2010. «Maximum entropy production and plant optimization theories». *Philosophical Transactions of the Royal Society B: Biological Sciences* 365 (1545): 1429-1435.
- Dodds, P. S., D. H. Rothman y J. S. Weitz. 2001. «Re-examination of the "3/4-law" of metabolism». *Journal of Theoretical Biology* 209 (1): 9-27.

- Dupré, J. E. 1987. *The Latest on the Best: Essays on Evolution and Optimality*. The MIT Press.
- Ebenhöh, O., y R. Heinrich. 2001. «Evolutionary optimization of metabolic pathways. Theoretical reconstruction of the stoichiometry of ATP and NADH producing systems». *Bulletin of mathematical biology* 63 (1): 21-55.
- Edgar, G. A. 2013. *Integral, Probability, and Fractal Measures*. Springer Science & Business Media.
- Edwards, J. S., M. Covert y B. Palsson. 2002. «Metabolic modelling of microbes: The flux-balance approach». *Environmental Microbiology* 4 (3): 133-140.
- Eldakar, O. T., y D. S. Wilson. 2011. «Eight criticisms not to make about group selection». *Evolution* 65 (6): 1523-1526.
- Elliott, L. P., y B. W. Brook. 2007. «Revisiting Chamberlin: Multiple working hypotheses for the 21st century». *BioScience* 57 (7): 608-614.
- Emlen, J. M. 1966. «The role of time and energy in food preference». *The American Naturalist* 100 (916): 611-617.
- Engle, R. F. 1984. «Wald, likelihood ratio, and Lagrange multiplier tests in econometrics». *Handbook of Econometrics* 2:775-826.
- Falconer, K. 2004. *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons.
- Fan, J., y col. 2014. «Quantitative flux analysis reveals folate-dependent NADPH production». *Nature* 510 (7504): 298.
- Fath, B., B. Patten y J. Choi. 2001. «Complementarity of ecological goal functions». *Journal of Theoretical Biology* 208 (4): 493-506.
- Feist, A. M., y B. O. Palsson. 2010. «The biomass objective function». *Current opinion in microbiology* 13 (3): 344-349.
- Feist, A. M., y col. 2007. «A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information». *Molecular Systems Biology* 3 (1): 121.
- Fetzer, J. H. 2010. «Is evolution an optimizing process?» En *The Place of Probability in Science*, ed. por E. Eells y J. H. Fetzer, 163-177. Springer.
- Fisher, R. 1922. «On the mathematical foundations of theoretical statistics». *Philosophical Transactions of the Royal Society A* 222 (594-604): 309-368.
- . 1934. «Professor Wright on the theory of dominance». *The American Naturalist* 68 (717): 370-374.
- Fisher, R. A. 1930. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press.

- Fisher, R. A., E. B. Ford y col. 1950. «The Sewall Wright effect.» *Heredity* 4:117-19.
- Flatt, T., y A. Heyland. 2011. *Mechanisms of Life History Evolution: The Genetics and Physiology of Life History Traits and Trade-offs*. Oxford University Press.
- Folland, G. B. 2013. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons.
- Frost, P. C., y col. 2005. «Are you what you eat? Physiological constraints on organismal stoichiometry in an elementally imbalanced world». *Oikos* 109 (1): 18-28.
- Gardner, A. 2009. «Adaptation as organism design». *Biology Letters* 5 (6): 861-864.
- Gardner, A., y A. Grafen. 2009. «Capturing the superorganism: A formal theory of group adaptation». *Journal of Evolutionary Biology* 22 (4): 659-671.
- Garland Jr, T., y P. Carter. 1994. «Evolutionary physiology». *Annual Review of Physiology* 56 (1): 579-621.
- Getty, T. 2000. «A constrained view of constraints». *Trends in Ecology & Evolution* 15 (6): 249.
- Gill, J. 1999. «The insignificance of null hypothesis significance testing». *Political Research Quarterly* 52 (3): 647-674.
- Gladstein, D. S., y col. 1991. «The need for sensitivity analyses of dynamic optimization models». *Oikos* 60 (1): 121-126.
- Glazier, D. S. 2005. «Beyond the '3/4-power law': variation in the intra-and interspecific scaling of metabolic rate in animals». *Biological Reviews of the Cambridge Philosophical Society* 80 (4): 611-662.
- Glazier, D. S. 2014. «Metabolic scaling in complex living systems». *Systems* 2 (4): 451-540.
- Godfrey-Smith, P. 2001. «Three kinds of adaptationism». En *Adaptationism and Optimality*, ed. por S. H. Orzack y E. Sober, 122:335-357. England.
- . 2008. «Varieties of population structure and the levels of selection». *The British Journal for the Philosophy of Science* 59 (1): 25-50.
- Goodnight, C. J. 2000. «Heritability at the ecosystem level». *Proceedings of the National Academy of Sciences of the United States of America* 97 (17): 9365-9366.
- . 2005. «Multilevel selection: The evolution of cooperation in non-kin groups». *Population Ecology* 47 (1): 3-12.
- Gotelli, N. J. 2001. «Research frontiers in null model analysis». *Global ecology and biogeography* 10 (4): 337-343.
- Gotelli, N. J., y W. Ulrich. 2012. «Statistical challenges in null model analysis». *Oikos* 121 (2): 171-180.
- Gotelli, N., y G. Graves. 1996. *Null Models in Ecology*. Smithsonian Institution.

- Gould, S. J. 1984. «Challenges to Neo-Darwinism and their meaning for a revised view of human consciousness». *The Tanner lectures on human values* 6:53-74.
- Gould, S. J., y R. C. Lewontin. 1979. «The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme». *Proceedings of the Royal Society of London. Series B. Biological Sciences* 205 (1161): 581-598.
- Grafen, A. 1984. «Natural selection, kin selection and group selection». En *Behavioural Ecology: An Evolutionary Approach*, ed. por J. R. Davies y N. B. Krebs, 2:62-84. Sinauer Associates.
- Gray, R. D. 1987. «Faith and foraging: A critique of the “paradigm argument from design”». En *Foraging Behavior*, ed. por A. C. Kamil, J. R. Krebs y H. R. Pulliam, 69-140. Boston, MA: Springer US.
- Greeff, J. M., y D. V. K. Newman. 2011. «Testing models of facultative sex ratio adjustment in the pollinating fig wasp *Platyscapa awekei*». *Evolution* 65 (1): 203-219.
- Haldane, J. 1957. «The cost of natural selection». *Journal of Genetics* 55 (3): 511-524.
- Hamilton, W. 1964a. «The genetical evolution of social behaviour. I». *Journal of Theoretical Biology* 7 (1): 1-16.
- . 1964b. «The genetical evolution of social behaviour. II». *Journal of Theoretical Biology* 7 (1): 17-52.
- Harcombe, W. R., y col. 2013. «The ability of flux balance analysis to predict evolution of central metabolism scales with the initial distance to the optimum». *PLoS computational biology* 9 (6): e1003091.
- Harcombe, W. R., y col. 2014. «Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics». *Cell Reports* 7 (4): 1104-1115.
- Hatchwell, B. J., y J. Komdeur. 2000. «Ecological constraints, life history traits and the evolution of cooperative breeding». *Animal Behaviour* 59 (6): 1079-1086.
- Heinrich, B. 1983. «Do bumblebees forage optimally, and does it matter?» *American Zoologist* 23 (2): 273-281.
- Hill, W. G. 2014. «Applications of population genetics to animal breeding, from Wright, Fisher and Lush to genomic prediction». *Genetics* 196 (1): 1-16.
- Hobbs, N. T. 1990. «Diet selection by generalist herbivores: a test of the linear programming model». En *Behavioural Mechanisms of Food Selection*, ed. por R. N. Hughes, 395-414. Springer.
- Hoelzer, G., E. Smith y J. Pepper. 2006. «On the logical relationship between natural selection and self-organization». *Journal of Evolutionary Biology* 19 (6): 1785-1794.
- Horn, H. S. 1971. *The Adaptive Geometry of Trees*. Vol. 3. Princeton University Press.

- Huggard, D. J. 1994. «A linear programming model of herbivore foraging: Imprecise, yet successful?» *Oecologia* 100 (4): 470-474.
- Hutchins, B. I., y col. 2016. «Relative Citation Ratio (RCR): A new metric that uses citation rates to measure influence at the article level». *PLoS Biology* 14 (9): e1002541.
- Inc., T. M. 2011. *version 7.12.0.635 (R2011a)*. Natick, Massachusetts: The MathWorks Inc.
- Irvine, E. 2015. «Models, robustness, and non-causal explanation: A foray into cognitive science and biology». *Synthese* 192 (12): 3943-3959.
- Jaynes, E. T. 2003. *Probability theory: The logic of science*. Cambridge University Press.
- Jeffreys, H. 1998. *The theory of probability*. Oxford University Press.
- Jensen, A. 1996. «Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival». *Canadian Journal of Fisheries and Aquatic Sciences* 53 (4): 820-822.
- Jensen, K., y col. 2012. «Optimal foraging for specific nutrients in predatory beetles». *Proceedings of the Royal Society of London B: Biological Sciences* 279 (1736): 2212-2218.
- Jesus, M. del, y col. 2012. «Maximum entropy production, carbon assimilation, and the spatial organization of vegetation in river basins». *Proceedings of the National Academy of Sciences* 109 (51): 20837-20841.
- Johnson, D. H. 1999. «The insignificance of statistical significance testing». *The Journal of Wildlife Management* 63 (3): 763-772.
- Johnson, J. B., y K. S. Omland. 2004. «Model selection in ecology and evolution». *Trends in Ecology & Evolution* 19 (2): 101-108.
- Jorgensen, S., y H. Mejer. 1979. «A holistic approach to ecological modelling». *Ecological Modelling* 7 (3): 169-189.
- Jørgensen, S. E., y B. D. Fath. 2004. «Application of thermodynamic principles in ecology». *Ecological Complexity* 1 (4): 267-280.
- Kamil, A. C. 1983. «Optimal foraging theory and the psychology of learning». *American Zoologist* 23 (2): 291-302.
- Kass, R. E., y A. E. Raftery. 1995. «Bayes factors». *Journal of the American Statistical Association* 90 (430): 773-795.
- Katul, G., y col. 2010. «A stomatal optimization theory to describe the effects of atmospheric CO<sub>2</sub> on leaf photosynthesis and transpiration». *Annals of Botany* 105 (3): 431.
- Katzourakis, N., y E. Varvaruca. 2018. *An Illustrative Introduction to Modern Analysis*. Chapman / Hall/CRC.
- Kauffman, K. J., P. Prakash y J. S. Edwards. 2003. «Advances in flux balance analysis». *Current opinion in biotechnology* 14 (5): 491-496.

- Kechris, A. 2012. *Classical descriptive set theory*. Vol. 156. Springer Science & Business Media.
- Khalil, E. L. 1997. «Economics, biology, and naturalism: Three problems concerning the question of individuality». *Biology and Philosophy* 12 (2): 185-206.
- Kimura, M. 1962. «On the probability of fixation of mutant genes in a population». *Genetics* 47 (6): 713-719.
- King, B. H. 1987. «Offspring sex ratios in parasitoid wasps». *The Quarterly Review of Biology* 62 (4): 367-396.
- King, D. A. 1990. «The adaptive significance of tree height». *The American Naturalist* 135 (6): 809-828.
- Kitcher, P. 1987. «Why Not the Best?» En *The Latest on the Best: Essays on Evolution and Optimality*, ed. por J. Dupre, 77-102. MIT Press.
- Kleidon, A., Y. Malhi y P. M. Cox. 2010. «Maximum entropy production in environmental and ecological systems». *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, n.º 1545 (1): 1297-1302.
- Kolmogorov, A. N. 1956. *Foundations of the Theory of Probability: Second English Edition*. Chelsea Publishing Company.
- Koziowski, J., y J. Weiner. 1997. «Interspecific allometries are by-products of body size optimization». *American Naturalist* 149 (2): 352-380.
- Kozłowski, J., M. Czarnołęski y M. Dańko. 2004. «Can optimal resource allocation models explain why ectotherms grow larger in cold?» *Integrative and Comparative Biology* 44 (6): 480-493.
- Kozłowski, J., y M. Konarzewski. 2004. «Is West, Brown and Enquist's model of allometric scaling mathematically correct and biologically relevant?» *Functional Ecology* 18 (2): 283-289.
- Kozłowski, J., M. Konarzewski y A. Gawelczyk. 2003. «Cell size as a link between noncoding DNA and metabolic rate scaling». *Proceedings of the National Academy of Sciences* 100 (24): 14080-14085.
- Krebs, J. R., A. Kacelnik y P. Taylor. 1978. «Test of optimal sampling by foraging great tits». *Nature* 275 (5675): 27.
- Kuhn, T. S. 1962. *The structure of scientific revolutions*. University of Chicago Press.
- LaMunyon, C. W., y T. K. Shackelford. 2002. «Evolutionary analyses should include pluralistic and falsifiable hypotheses». *Behavioral and Brain Sciences* 25 (4): 522-523.
- Lange, M. 2017. *Because Without Cause: Non-causal Explanations in Science and Mathematics*. 1.<sup>a</sup> ed. Oxford Studies in Philosophy of Science. Oxford University Press.

- Lauder, G. V. 1996. «The argument from design». En *Adaptation*, ed. por M. R. Rose y G. V. Lauder, 55-91. Academic Press.
- Lawrence, C. E., y R. F. Gariepy. 2015. *Measure theory and fine properties of functions. Revised edition*.
- Lee, J. M., E. P. Gianchandani y J. A. Papin. 2006. «Flux balance analysis in the era of metabolomics». *Briefings in Bioinformatics* 7 (2): 140-150.
- Lehmann, E. L., y J. P. Romano. 2006. *Testing statistical hypotheses*. Springer Science & Business Media.
- Leigh, E. J. 2010. «The group selection controversy». *Journal of Evolutionary Biology* 23 (1): 6-19.
- Levins, R., y R. C. Lewontin. 1985. *The dialectical biologist*. Harvard University Press.
- Lewis, N. E., y col. 2010. «Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models». *Molecular Systems Biology* 6 (1): 390.
- Lewontin, R. C. 1970. «The Units of Selection» [inlangEnglish]. *Annual Review of Ecology and Systematics* 1:
- Lewontin, R. C. 1978. «Adaptation». *Scientific American* 239 (3): 212-231.
- Linde, M., y M. Palmer. 2008. «Testing Allsop and West's size at sex change invariant within a fish species: a spurious ratio or a useful group descriptor?» *Journal of Evolutionary Biology* 21 (3): 914-917.
- Linde-Medina, M., y R. Diogo. 2014. «Do correlation patterns reflect the role of development in morphological evolution?» *Evolutionary Biology* 41 (3): 494-502.
- Loeb, P. A. 1975. «Conversion from nonstandard to standard measure spaces and applications in probability theory». *Transactions of the American Mathematical society* 211:113-122.
- Lotka, A. J. 1922a. «Contribution to the energetics of evolution». *Proceedings of the National Academy of Sciences of the United States of America* 8 (6): 147-151.
- . 1922b. «Natural selection as a physical principle». *Proceedings of the National Academy of Sciences of the United States of America* 8 (6): 151-154.
- . 1944. «Evolution and thermodynamics». *Science & Society* 8 (2): 161-171.
- MacArthur, R. H., y E. R. Pianka. 1966. «On optimal use of a patchy environment». *American Naturalist* 100 (916): 603-609.
- Machado, D., y M. Herrgård. 2014. «Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism». *PLoS Computational Biology* 10 (4): e1003580.
- Makela, A., y col. 2002. «Challenges and opportunities of the optimality approach in plant ecology». *Silva Fennica* 36 (3): 605-614.

- Mangel, M. 1996. «Life history invariants, age at maturity and the ferox trout». *Evolutionary Ecology* 10 (3): 249-263.
- Mangel, M., y C. W. Clark. 1986. «Towards a unified foraging theory». *Ecology* 67 (5): 1127-1138.
- Margalef, R. 1963. «On certain unifying principles in ecology». *The American Naturalist* 97 (897): 357-374.
- Martyushev, L. M., y V. D. Seleznev. 2006. «Maximum entropy production principle in physics, chemistry and biology». *Physics Reports* 426 (1): 1-45.
- Martyushev, L. M. 2010. «The maximum entropy production principle: Two basic questions». *Philosophical Transactions of the Royal Society B: Biological Sciences* 365 (1545): 1333-1334.
- Massimo Pigliucci, G. B. M. 2010. *Evolution - the Extended Synthesis*. The MIT Press.
- Matthen, M., y A. Ariew. 2002. «Two ways of thinking about fitness and natural selection». *The Journal of Philosophy* 99 (2): 55-83.
- Mattila, P. 1999. *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*. Vol. 44. Cambridge university press.
- Mayhew, P. J. 1997. «Adaptive patterns of host-plant selection by phytophagous insects». *Oikos* 79 (3): 417-428.
- Maynard Smith, J. 1964. «Group selection and kin selection». *Nature* 201 (4924): 1145-1147.
- Mayr, E. 1983. «How to carry out the adaptationist program?» *The American Naturalist* 121 (3): 324-334.
- Meysman, F. J. R., y S. Bruers. 2007. «A thermodynamic perspective on food webs: Quantifying entropy production within detrital-based ecosystems». *Journal of Theoretical Biology* 249 (1): 124-139.
- . 2010. «Ecosystem functioning and maximum entropy production: a quantitative test of hypotheses». *Philosophical Transactions of the Royal Society B: Biological Sciences* 365 (1545): 1405-1416.
- Migdalas, A., P. M. Pardalos y P. Värbrand. 2013. *Multilevel optimization: algorithms and applications*. Vol. 20. Springer Science & Business Media.
- Milo, R., y R. L. Last. 2012. «Achieving diversity in the face of constraints: Lessons from metabolism». *Science* 336 (6089): 1663-1667.
- Mitchell, S. D. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge University Press.
- Moran, P. 1964. «On the nonexistence of adaptive topographies». *Annals of human genetics* 27 (4): 383-393.

- Nee, S., y col. 2005. «The illusion of invariant quantities in life histories». *Science* 309 (5738): 1236-1239.
- Neyman, J., y E. Pearson. 1933. «On the problem of the most efficient tests of statistical hypotheses». *Philosophical Transactions of the Royal Society* 231:289-337.
- Nicholls, N. 2001. «The insignificance of significance testing». *Bulletin of the American Meteorological Society* 82 (5): 981-986.
- Nogales, J., y col. 2012. «Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis». *Proceedings of the National Academy of Sciences of the United States of America* 109 (7): 2678-2683.
- Nonacs, P., y L. M. Dill. 1990. «Mortality risk vs. food quality trade-offs in a common currency: Ant patch preferences». *Ecology* 71 (5): 1886-1892.
- Oberhardt, M. A., y col. 2008. «Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PA01». *Journal of Bacteriology* 190 (8): 2790-2803.
- Öckinger, E., y col. 2010. «Life-history traits predict species responses to habitat area and isolation: a cross-continental synthesis». *Ecology Letters* 13 (8): 969-979.
- O'Connor, M. P., y col. 2007. «Reconsidering the mechanistic basis of the metabolic theory of ecology». *Oikos* 116 (6): 1058-1072.
- Odum, H. T., y R. C. Pinkerton. 1955. «Time's speed regulator: the optimum efficiency for maximum power output in physical and biological systems». *Amer. Sci.* 43:331-343.
- Odum, H. 1988. «Self-organization, transformity and information.» *Science* 242:1132-1139.
- Okasha, S. 2009. *Evolution and the Levels of Selection*. Oxford University Press.
- Olson, M. E., y A. Arroyo-Santos. 2015. «How to study adaptation (and why to do it that way)». *The Quarterly Review of Biology* 90 (2): 167-191.
- Orr, H. A. 2005. «The genetic theory of adaptation: A brief history». *Nature Reviews. Genetics* 6 (2): 119.
- Orth, J. D., I. Thiele y B. Ø. Palsson. 2010. «What is flux balance analysis?» *Nature Biotechnology* 28 (3): 245-248.
- Orzack, S. H. 2002. «Using sex ratios: The past and the future». Ed. por I. C. W. Hardy. *Sex Ratios: Concepts and Research Methods*: 383-398.
- . 2012. «The philosophy of modelling or does the philosophy of biology have any use?» *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1586): 170-180.
- Orzack, S. H., y E. Sober. 1996. «How to formulate and test adaptationism». *The American Naturalist* 148 (1): 202-210.

- Orzack, S. H., y E. Sober. 1993. «A critical assessment of Levins's *The Strategy of Model Building in Population Biology* (1966)». *The Quarterly Review of Biology* 68 (4): 533-546.
- . 1994a. «How (not) to test an optimality model». *Trends in Ecology & Evolution* 9 (7): 265-267.
- . 1994b. «Optimality models and the test of adaptationism». *The American Naturalist* 143 (3): 361-380.
- . 2001. *Adaptationism and Optimality*. 1.<sup>a</sup> ed. Cambridge Studies in Philosophy and Biology. Cambridge University Press.
- Oster, G. F., y E. O. Wilson. 1979. *Caste and ecology in the social insects*. Princeton University Press.
- Owen-Smith, N. 1993. «Assessing the constraints for optimal diet models». *Evolutionary Ecology* 7 (5): 530-531.
- . 1996. «Circularity in linear programming models of optimal diet». *Oecologia* 108 (2): 259-261.
- Parker, G. A., y J. Maynard Smith. 1990. «Optimality theory in evolutionary biology». *Nature* 348 (6296): 27-33.
- Pellis, S. M. 2002. «When is a trait an adaptation?» *Behavioral and Brain Sciences* 25 (4): 524-524.
- Pennell, M., y M. O'Connor. 2017. «A modest proposal for unifying macroevolution and ecosystem ecology.» *The American naturalist* 189 (6): ii-iii.
- Pérez-Escudero, A., M. Rivera-Alba y G. G. de Polavieja. 2009. «Structure of deviations from optimality in biological systems». *Proceedings of the National Academy of Sciences of the United States of America* 106 (48): 20544-20549.
- Perry, G., y E. R. Pianka. 1997. «Animal foraging: past, present and future». *Trends in Ecology & Evolution* 12 (9): 360-364.
- Pierce, G., y J. Ollason. 1987. «Eight reasons why optimal foraging theory is a complete waste of time». *Oikos* 48 (1): 111-118.
- Pigliucci, M. 2008. «Is evolvability evolvable?» *Nature Reviews Genetics* 9 (1): 75.
- Pigliucci, M., y J. Kaplan. 2000. «The fall and rise of Dr Pangloss: Adaptationism and the *Spandrels* paper 20 years later». *Trends in Ecology & Evolution* 15 (2): 66-70.
- Pike, N. 2011. «Using false discovery rates for multiple comparisons in ecology and evolution». *Methods in Ecology and Evolution* 2 (3): 278-282.
- Pincock, C. 2015. «Abstract explanations in science». *The British Journal for the Philosophy of Science* 66 (4): 857-882.

- Plank, M., y A. James. 2008. «Optimal foraging: Levy pattern or process?» *Journal of The Royal Society Interface* 5 (26): 1077-1086.
- Plath, K., y M. Boersma. 2001. «Mineral limitation of zooplankton: Stoichiometric constraints and optimal foraging». *Ecology* 82 (5): 1260-1269.
- Platt, J. R. 1964. «Strong inference». *Science* 146 (3642): 347-353.
- Pollard, D. 2002. *A user's guide to measure theoretic probability*. Vol. 8. Cambridge University Press.
- Poolman, M. G., y col. 2009. «A genome-scale metabolic model of Arabidopsis and some of its properties». *Plant Physiology* 151 (3): 1570-1581.
- Potgieter, P. 2009. «Nonstandard analysis, fractal properties and Brownian motion». *Fractals* 17 (01): 117-129.
- Potochnik, A. 2007. «Optimality modeling and explanatory generality». *Philosophy of Science* 74 (5): 680-691.
- . 2009. «Optimality modeling in a suboptimal world». *Biology & Philosophy* 24 (2): 183.
- . 2010. «Explanatory independence and epistemic interdependence: A case study of the optimality approach». *The British Journal for the Philosophy of Science* 61 (1): 213-233.
- Price, N. D., y col. 2003. «Genome-scale microbial in silico models: The constraints-based approach». *Trends in Biotechnology* 21 (4): 162-169.
- Prigogine, I., y J. Wiame. 1946. «Biologie et thermodynamique des phénomènes irréversibles». *Experientia* 2 (11): 451-453.
- Pyke, G. H., H. R. Pulliam y E. L. Charnov. 1977. «Optimal foraging: A selective review of theory and tests». *The Quarterly Review of Biology* 52 (2): 137-154.
- Pyke, G. H. 1984. «Optimal foraging theory: A critical review». *Annual Review of Ecology and Systematics* 15:
- Raman, K., y N. Chandra. 2009. «Flux balance analysis of biological systems: Applications and challenges». *Briefings in Bioinformatics* 10 (4): 435-449.
- Raubenheimer, D., y S. J. Simpson. 2004. «Organismal stoichiometry: quantifying non-independence among food components». *Ecology* 85 (5): 1203-1216.
- Ray, S. 2006. «Extremal principles with special emphasis on exergy and ascendancy — The modern approach in theoretical ecology». *Journal of Biological Systems* 14 (02): 255-273.
- Real Academia Española. 2017. «Diccionario de la lengua española».
- Reed, J. L., y B. Ø. Palsson. 2003. «Thirteen years of building constraint-based in silico models of *Escherichia coli*». *Journal of Bacteriology* 185 (9): 2692-2699.

- Reeve, H. K., y P. W. Sherman. 1993. «Adaptation and the goals of evolutionary research». *The Quarterly Review of Biology* 68 (1): 1-32.
- Reuters, T. 2017 (accedido el 25 de julio de 2017). *Web of Science*.
- . 2019. *Web of Science*.
- Reutlinger, A. 2017a. «Does the counterfactual theory of explanation apply to non-causal explanations in metaphysics?» *European Journal for Philosophy of Science* 7 (2): 239-256.
- . 2017b. «Explanation beyond causation? New directions in the philosophy of scientific explanation». *Philosophy Compass* 12 (2): e12395-e12395.
- Rice, C. 2012. «Optimality explanations: A plea for an alternative approach». *Biology & Philosophy* 27 (5): 685-703.
- . 2015. «Moving beyond causes: Optimality models and scientific explanation». *Noûs* 49 (3): 589-615.
- Richardson, R. C. 1994. «Optimization in evolutionary ecology». *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1994:13-21.
- Robinson, B. W., y D. S. Wilson. 1998. «Optimal foraging, specialization, and a solution to Liem's paradox». *The American Naturalist* 151 (3): 223-235.
- Roff, D. A. 2002. *Life History Evolution*. Vol. 7. Sinauer Associates Sunderland.
- Rosen, R. 1967. *Optimality Principles in Biology*. Springer.
- Rytwinski, T., y L. Fahrig. 2012. «Do species life history traits explain population responses to roads? A meta-analysis». *Biological Conservation* 147 (1): 87-98.
- Saha, R., P. F. Suthers y C. D. Maranas. 2011. «Zea mays iRS1563: a comprehensive genome-scale metabolic reconstruction of maize metabolism». *PloS one* 6 (7): e21784.
- Sánchez, C. E. G., R. G. T. Sáez y col. 2014. «Comparison and analysis of objective functions in flux balance analysis». *Biotechnology Progress* 30:985-991.
- Savage, V. M., E. J. Deeds y W. Fontana. 2008. «Sizing up allometric scaling theory». *PLoS Computational Biology* 4 (9): e1000171.
- Savage, V. M., y col. 2006. «Comment on "The illusion of invariant quantities in life histories"». *Science* 312 (5771): 198-198.
- Schneider, C. A., W. S. Rasband y K. W. Eliceiri. 2012. «NIH Image to ImageJ: 25 years of image analysis». *Nature Methods* 9 (7): 671-675.
- Schneider, E. D., y J. J. Kay. 1994. «Life as a manifestation of the second law of thermodynamics», *Mathematical and Computer Modelling* 19 (6-8): 25-48.

- Schoemaker, P. J. 1991. «The quest for optimality: A positive heuristic of science?» *Behavioral and Brain Sciences* 14 (2): 205-215.
- Schoener, T. W. 1969. «Models of optimal size for solitary predators». *The American Naturalist* 103 (931): 277-313.
- Schuetz, R., L. Kuepfer y U. Sauer. 2007. «Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*». *Molecular Systems Biology* 3 (1): 119.
- Schuetz, R., y col. 2012. «Multidimensional optimality of microbial metabolism». *Science* 336 (6081): 601-604.
- Selvarasu, S., y col. 2012. «Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture». *Biotechnology and bioengineering* 109 (6): 1415-1429.
- Sih, A., y B. Christensen. 2001. «Optimal diet theory: When does it work, and when and why does it fail?» *Animal Behaviour* 61 (2): 379-390.
- Simon, H. A. 1956. «Rational choice and the structure of the environment.» *Psychological Review* 63 (2): 129.
- Simpson, S. J., y col. 2004. «Optimal foraging when regulating intake of multiple nutrients». *Animal Behaviour* 68 (6): 1299-1311.
- Skene, K. R. 2017. «Thermodynamics, ecology and evolutionary biology: A bridge over troubled water or common ground?» *Acta Oecologica* 85:116-125.
- Smith, C. C., y S. D. Fretwell. 1974. «The optimal balance between size and number of offspring». *The American Naturalist* 108 (962): 499-506.
- Sober, E. 2008. *Evidence and Evolution: The Logic Behind the Science*. Vol. Volume 5. Cambridge University Press.
- Soetaert, K., K. Van den Meersche y D. van Oevelen. 2009. «limSolve: Solving linear inverse models». *R package version 1* (1).
- Speakman, J., y J. McQueenie. 1996. «Limits to sustained metabolic rate: the link between food intake, basal metabolic rate, and morphology in reproducing mice, *Mus musculus*». *Physiological Zoology* 69 (4): 746-769.
- Stephens, D. W., J. S. Brown y R. C. Ydenberg. 2007. *Foraging: behavior and ecology*. University of Chicago Press.
- Stephens, D. W., y J. R. Krebs. 1986. *Foraging theory*. Princeton University Press.
- Stoltzfus, A. 2012. «Constructive neutral evolution: eExploring evolutionary theory's curious disconnect». *Biology Direct* 7 (1): 35.
- Stolyar, S., y col. 2007. «Metabolic modeling of a mutualistic microbial community». *Molecular systems biology* 3 (1): 92.

- Stone, M. 1977. «An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion». *Journal of the Royal Statistical Society. Series B (Methodological)* 39 (1): 44-47.
- Sugiura, N. 1978. «Further analysts of the data by Akaike's information criterion and the finite corrections: Further analysts of the data by akaike's». *Communications in Statistics-Theory and Methods* 7 (1): 13-26.
- Sutherland, W. J. 2005. «The best solution». *Nature* 435 (7042): 569-569.
- Svensson, E., y R. Calsbeek. 2012. *The Adaptive Landscape in Evolutionary Biology*. Oxford University Press.
- Swenson, W., J. Arendt y D. S. Wilson. 2000a. «Artificial selection of microbial ecosystems for 3-chloroaniline biodegradation». *Environmental Microbiology* 2 (5): 564-571.
- Swenson, W., D. S. Wilson y R. Elias. 2000b. «Artificial ecosystem selection». *Proceedings of the National Academy of Sciences of the United States of America* 97 (16): 9110-9114.
- Team, R. D. C. 2011. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Traulsen, A., y M. A. Nowak. 2006. «Evolution of cooperation by multilevel selection». *Proceedings of the National Academy of Sciences of the United States of America* 103 (29): 10952-10955.
- Ulanowicz, R. E., y B. M. Hannon. 1987. «Life and the production of entropy». *Proceedings of the Royal Society of London. Series B, Biological Sciences* 232 (1267): 181-192.
- Uthman, O. A., y col. 2013. «Citation classics in systematic reviews and meta-analyses: Who wrote the top 100 most cited articles?» *PLoS One* 8 (10): e78517.
- Valentine, H. T., y A. Mäkelä. 2012. «Modeling forest stand dynamics from optimal balances of carbon and nitrogen». *New Phytologist* 194 (4): 961-971.
- Vallino, J. J. 2010. «Ecosystem biogeochemistry considered as a distributed metabolic network ordered by maximum entropy production». *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, n.º 1545 (1): 1417-1427.
- Vallino, J. J., y C. K. Algar. 2016. «The thermodynamics of marine biogeochemical cycles: Lotka revisited». *Annual Review of Marine Science* 8:333-356.
- Van den Meersche, K., K. Soetaert y D. Van Oevelen. 2009. «xsample (): An R function for sampling linear inverse problems». *Journal of Statistical Software* 30 (Code Snippet 1): HASH-0x57696b8.
- Vazquez, A., y col. 2010. «Catabolic efficiency of aerobic glycolysis: the Warburg effect revisited». *BMC systems biology* 4 (1): 58.
- Veech, J. A. 2012. «Significance testing in ecological null models». *Theoretical Ecology* 5 (4): 611-616.

- Velenich, A., y J. Gore. 2012. «Synthetic approaches to understanding biological constraints». *Current Opinion in Chemical Biology* 16 (3): 323-328.
- Volk, T., y O. Pauluis. 2010. «It is not the entropy you produce, rather, how you produce it». *Philosophical Transactions of the Royal Society B: Biological Sciences* 365 (1545): 1317-1322.
- Wade, M. J., y C. J. Goodnight. 1998. «Perspective: The theories of Fisher and Wright in the context of metapopulations: When nature does many small experiments». *Evolution* 52 (6): 1537-1553.
- Wang, H., I. Prentice y T. Davis. 2014. «Biophysical constraints on gross primary production by the terrestrial biosphere». *Biogeosciences* 11 (20): 5987-6001.
- Ward, D. 1992. «The role of satisficing in foraging theory». *Oikos* 63 (2): 312-317.
- . 1993. «Foraging theory, like all other fields of science, needs multiple working hypotheses». *Oikos* 67 (2): 376-378.
- Watt, W. B. 2000. «Avoiding paradigm-based limits to knowledge of evolution». En *Evolutionary Biology: Limits to Knowledge in Evolutionary Genetics*, ed. por M. T. Clegg, M. K. Hecht y R. J. Macintyre, 73-96. Boston, MA: Springer US.
- Weber, B., y col. 1989. «Evolution in thermodynamic perspective: An ecological approach». *Biology and Philosophy* 4 (4): 373-405.
- Weber, B. H., y D. J. Depew. 1996. «Natural selection and self-organization». *Biology and Philosophy* 11 (1): 33-65.
- Webster, J., y R. T. Watson. 2002. «Analyzing the past to prepare for the future: Writing a literature review». *MIS quarterly*: xiii-xxiii.
- Werner, E. E., y B. R. Anholt. 1993. «Ecological consequences of the trade-off between growth and mortality rates mediated by foraging activity». *The American Naturalist* 142 (2): 242-272.
- Werner, E. E., y D. J. Hall. 1974. «Optimal foraging and the size selection of prey by the bluegill sunfish (*Lepomis macrochirus*)». *Ecology* 55 (5): 1042-1052.
- Werner, J., y E. M. Griebeler. 2014. «Allometries of maximum growth rate versus body mass at maximum growth indicate that non-avian dinosaurs had growth rates typical of fast growing ectothermic sauropsids». *PloS one* 9 (2): e88834.
- West, G., J. Brown y B. Enquist. 1997. «A general model for the origin of allometric scaling laws in biology». *Science* 276 (5309): 122-126.
- . 1999. «The fourth dimension of life: Fractal geometry and allometric scaling of organisms». *Science* 284 (5420): 1677-1679.
- Whitham, T. G., y col. 2003. «Community and ecosystem genetics: A consequence of the extended phenotype». *Ecology* 84 (3): 559-573.

- Whitham, T. G., y col. 2006. «A framework for community and ecosystem genetics: from genes to ecosystems». *Nature Reviews Genetics* 7 (7): 510-523.
- Wilks, S. S. 1938. «The large-sample distribution of the likelihood ratio for testing composite hypotheses». *The Annals of Mathematical Statistics* 9 (1): 60-62.
- Williams, G. C. 1966. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*.
- . 1972. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton University Press.
- Williams, H. T. P., y T. M. Lenton. 2007. «Artificial selection of simulated microbial ecosystems». *Proceedings of the National Academy of Sciences of the United States of America* 104, n.º 21 (1): 8918-8923.
- Williams, N. M., y col. 2010. «Ecological and life-history traits predict bee species responses to environmental disturbances». *Biological Conservation* 143 (10): 2280-2291.
- Wilson, D. S. 1975. «A theory of group selection». *Proceedings of the National Academy of Sciences of the United States of America* 72 (1): 143-146.
- . 1983. «The group selection controversy: History and current status». *Annual Review of Ecology and Systematics* 14 (1): 159-187.
- Wilson, D. S., y E. Sober. 1994. «Reintroducing group selection to the human behavioral sciences». *Behavioral and Brain Sciences* 17 (4): 585-608.
- Wilson, R. A. 2007. «Levels of selection». En *Philosophy of Biology*, ed. por D. M. Gabbay, P. Thagard y J. Woods, 141-162. Elsevier.
- Wintermute, E. H., y P. A. Silver. 2010. «Emergent cooperation in microbial metabolism». *Molecular systems biology* 6 (1): 407.
- Wintermute, E., T. Lieberman y P. Silver. 2013. «An objective function exploiting suboptimal solutions in metabolic networks». *BMC Systems Biology* 7 (1): 98.
- Wright, S. 1929. «Fisher's theory of dominance». *The American Naturalist* 63 (686): 274-279.
- . 1931. «Evolution in Mendelian populations». *Genetics* 16 (2): 97-159.
- . 1951. «Fisher and Ford on "the Sewall Wright effect"». *American Scientist* 39 (3): 452-479.
- . 1982. «The shifting balance theory and macroevolution». *Annual Review of Genetics* 16 (1): 1-20.
- Wynne-Edwards, V. C. 1962. *Animal Dispersion: In Relation to Social Behaviour*. Oliver / Boyd.

- Yekutieli, D., e Y. Benjamini. 1999. «Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics». *Journal of Statistical Planning and Inference* 82 (1-2): 171-196.
- Yen, J. D., y col. 2014. «Thermodynamic extremization principles and their relevance to ecology». *Austral Ecology* 39 (6): 619-632.
- Yoshimura, J., y W. Shields. 1987. «Probabilistic optimization of phenotype distributions: A general solution for the effects of uncertainty on natural selection?» *Evolutionary Ecology* 1 (2): 125-138.
- Zhao, Q., y col. 2016. «Mapping the landscape of metabolic goals of a cell». *Genome Biology* 17 (1): 109.
- Zomorodi, A. R., y C. D. Maranas. 2012. «OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities». *PLoS computational biology* 8 (2): e1002363.