



UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**ESTUDIO DEL DISEÑO DE PRODUCTOS
BAJO PREFERENCIAS NO COMPENSATORIAS**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

IVÁN ALEJANDRO PEREIRA SÁNCHEZ

PROFESOR GUÍA:

RICARDO MONTOYA MOREIRA

MIEMBROS DE LA COMISIÓN:

ANDRÉS MUSALEM SAID

ALEJANDRA PUENTE CHANDÍA

SANTIAGO DE CHILE

2019

**RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE:** Ingeniero Civil Industrial
POR: Iván Alejandro Pereira Sánchez
FECHA: 02/01/2019
PROFESOR GUÍA: Ricardo Montoya

ESTUDIO DEL DISEÑO DE PRODUCTOS BAJO PREFERENCIAS NO COMPENSATORIAS

El diseño de productos, como la selección de surtido de productos, representan decisiones críticas a las que se ven enfrentados los retailers. La variedad y tipo de productos que se disponen condicionan de manera sustancial el comportamiento de compra de los consumidores y, por ende, las utilidades que reciben las compañías de la industria. En un mercado creciente, en que la cantidad de nuevos productos introducidos aumenta año tras año, resulta natural suponer que los individuos no consideran, ni evalúan todos los atributos de las alternativas a las que se exponen. De esta forma, se espera que se centren en aquellas que le son más importantes, haciendo más fácil el proceso de elección. No obstante, la investigación en materia de diseño y planificación de surtido de productos se ha centrado tradicionalmente en el uso de preferencias compensatorias, según las cuales un consumidor considera todas las alternativas y atributos disponibles al momento de tomar la decisión de compra.

El principal objetivo del presente trabajo es estudiar el diseño de productos cuando los consumidores utilizan reglas de decisión no compensatorias. Para ello, se estiman preferencias sobre un conjunto de datos simulados utilizando un modelo lexicográfico, el que se compara con un modelo compensatorio. Luego, empleando una heurística de programación dinámica, se determina el perfil de producto óptimo para cada modelo, cuyo desempeño se mide a través de un índice de participación de mercado y utilidad esperada, obtenida según una función de costos definida según el atractivo de los niveles de atributo considerados. A la vez, se replica este procedimiento sobre un conjunto de datos reales derivados de una encuesta de elección de cámaras fotográficas.

Uno de los principales resultados es que se aprecian diferencias entre el tipo de producto obtenido con cada enfoque. Así, se observa que en la función de utilidad que tiene la compañía diseñando en base a un modelo no compensatorio parece importar más el factor demanda, obteniéndose mayores índices de participación de mercado respecto a un modelo compensatorio, donde es más relevante la estructura de costos. Por otro lado, al aplicar la metodología a un set de datos reales y comparar la utilidad obtenida con ambos enfoques, se distinguen mayores diferencias en el desempeño cuando el producto óptimo es evaluado asumiendo preferencias compensatorias versus no compensatorias, reafirmando el hecho que el modelo compensatorio es robusto en múltiples escenarios. Por último, para futuros estudios en la materia, se propone explorar modelos no compensatorios en el diseño de surtido de productos, junto a la aplicación de métodos de optimización alternativos a la programación dinámica, como el uso de algoritmos genéticos y estudiar la interacción de ambos enfoques, incluyendo condiciones y supuestos que permitan aproximar la investigación a un escenario más realista.

“Stay hungry. Stay foolish.”



Agradecimientos

Llegar a esta instancia significa un logro muy importante para mí, pues representa la culminación de una de las etapas más relevantes de mi vida. Resulta difícil plasmar en pocas palabras el profundo sentimiento de agradecimiento que tengo por todos aquellos que me han acompañado a lo largo de este camino. En primer lugar, quiero agradecer a mi familia por animarme y brindarme su apoyo incondicional en cada instante de mi vida. A mi mamá Cecilia y a mi papá Iván, a quienes quiero mucho, por siempre incentivar e inculcar en mí el sentido de perseverancia y responsabilidad; y por darme las habilidades y los cimientos que me convierten en quien soy. A mi hermano Cristián, por darme ánimo e inspirarme a confiar en mis propias habilidades. A mi tía, la Eme, a quien quiero igual que a mis padres, por su cariño, incondicionalidad y apoyo en cada etapa de mi crecimiento. A mis abuelos, los que están y quienes han partido, por su afecto y las incontables enseñanzas que guardo para siempre en mi memoria. Al resto de mis tíos, tías, primas y primos, por el aprecio que me han manifestado. Sepan que valoro mucho cada una de sus muestras de cariño y estaré para ustedes siempre que Dios me lo permita.

A mis amigos de universidad, por hacer de mi permanencia en la facultad una de las etapas más felices de mi vida. A mis mejores amigos Cris y Mario, por siempre contagiarme con su entusiasmo y optimismo. Gracias por incentivar a mostrar mi perfil más innovador y por estar constantemente allí para mí cada vez que lo necesito. A mis amigos Álvaro, Luis, Iván y Osvaldo por todos los instantes que hemos pasado juntos y por contribuir día a día a reforzar la amistad de nuestro grupo. A mis amigas y hermanas Pame y Coni por estar presentes desde las etapas tempranas de mi vida universitaria y por todos los grandes momentos que hemos compartido. A Cata y Caro por la amistad que mantenemos desde tercer año, cuando este proyecto adquiriría mejor forma con Taller I. Quiero decirles que estoy muy orgulloso de ser su amigo y espero que nuestra amistad siga prosperando y se nutra de todas las aventuras que vendrán en el futuro. A todos ustedes, muchas gracias.

A mi profesor guía Ricardo, por orientar mi trabajo y ayudarme en este proceso, tolerando mi frustración y enseñándome que con paciencia y abstracción de los detalles puedo adquirir una comprensión global para abordar de mejor forma los problemas. A mis profesores Andrés y Alejandra, cuyo interés y participación en esta investigación fueron muy útiles para dar mayor fundamento y entendimiento al estudio realizado. A mi equipo de trabajo en Falabella por el ánimo brindado durante este proceso, especialmente a Caro y Cami por depositar en mí su confianza y darme la oportunidad de desafiarme profesionalmente al afrontar los retos del mundo laboral.

Mi gratitud a todos por haberme acompañado a cumplir este gran propósito, espero que esta etapa que comienza sea el inicio de nuevos desafíos que involucren tantos aprendizajes y éxitos como lo ha sido hasta el momento. Gracias totales.

Tabla de contenido

1. Introducción.....	1
1.1 Motivación	3
1.2 Objetivos	5
1.2.1 Objetivo general	5
1.2.2 Objetivos específicos	5
1.3 Metodología	6
1.3.1 Revisión bibliográfica	6
1.3.2 Estimación de preferencias no compensatorias.....	6
1.3.3 Diseño del problema de optimización	6
1.3.4 Optimización del modelo	6
1.3.5 Validación.....	6
1.4 Alcances.....	7
1.5 Resultados esperados	7
1.6 Estructura del trabajo de título	8
2 Antecedentes generales	10
2.1 Estado del arte	10
2.1.1 Diseño de nuevos productos	10
2.1.2 Assortment planning.....	11
2.1.3 Reglas de decisión	13
2.2 Marco conceptual.....	14
2.2.1 Análisis conjunto.....	14
2.2.2 Reglas de decisión	15
2.2.3 Métodos de estimación y caracterización de preferencias	16
2.2.4 Heurística Greedy	18
3 Estimación de preferencias no compensatorias	20
3.1 Representación de modelos de preferencia lexicográficos	20
3.1.1 Regla lexicográfica	20
3.1.2 Utilidad lexicográfica	22
3.1.3 Tipos de representaciones lexicográficas	23
3.2 Metodología de inferencia de modelos de preferencia lexicográficos	26
3.3 Inferencia de modelos de preferencia lexicográficos sobre datos simulados.....	30

3.3.2	Estimación variando el número de niveles de atributos.....	33
3.3.3	Estimación variando el número de preguntas.....	34
3.3.4	Estimación variando el número de alternativas por pregunta.	35
3.3.5	Estimación variando el coeficiente lexicográfico w	36
3.4	Conclusiones	37
4	Estimación de preferencias compensatorias	38
4.1	Representación de modelos de preferencia compensatorios.....	38
4.2	Simulación de preferencias compensatorias.....	39
4.3	Metodología de estimación de preferencias compensatorias	40
4.4	Inferencia de modelos de preferencia compensatorios sobre datos simulados. .	42
4.4.1	Estimación variando el número de atributos	42
4.4.2	Estimación variando el número de niveles.....	43
4.4.3	Estimación variando el número de preguntas.....	44
4.4.4	Estimación variando el número de alternativas.....	45
4.5	Conclusiones	46
5	Diseño de producto óptimo bajo preferencias simuladas	47
5.1	Metodologías de diseño de producto óptimo	48
5.2	Diseño de producto óptimo con programación dinámica	50
5.3	Diseño de producto óptimo con búsqueda exhaustiva	53
5.4	Comparación de producto óptimo entre modelos de preferencia compensatorios y no compensatorios	53
5.4.1	Datos simulados de manera compensatoria.....	55
5.4.2	Datos simulados de manera no compensatoria	59
5.4.3	Variación de la condición de heterogeneidad.....	61
5.5	Conclusiones	66
6	Diseño de producto óptimo bajo preferencias reales	67
6.1	Estructura y características del set de datos	68
6.2	Análisis descriptivo	69
6.2.1	Distribución de la elección por nivel de precio	69
6.2.2	Distribución de la elección por nivel de resolución.....	69
6.2.3	Distribución de la elección por nivel de batería	70
6.2.4	Distribución de la elección por nivel de zoom.....	70
6.2.5	Distribución de la elección por nivel de tamaño	71
6.3	Resultados.....	73
6.3.1	Estimación producto óptimo asumiendo preferencias compensatorias	73
6.3.2	Estimación producto óptimo asumiendo preferencias no compensatorias .	75

6.4	Conclusiones	78
7	Conclusiones	79
7.1	Conclusiones generales.....	79
7.2	Discusión general y trabajos futuros	81
	Apéndice A.....	83
	Apéndice B.....	87
	Bibliografía	88

Índice de tablas

Tabla 1: Set de datos de elecciones simuladas.....	31
Tabla 2: Estructura de costos para diseño de producto óptimo de datos simulados.....	54
Tabla 3: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=1$).	55
Tabla 4: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=1$).	55
Tabla 5: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=1$).	56
Tabla 6: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=1$).	56
Tabla 7: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=3$).	57
Tabla 8: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=3$).	57
Tabla 9: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=3$).	58
Tabla 10: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=3$).	58
Tabla 11: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera no compensatoria.	59
Tabla 12: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera no compensatoria.	59

Tabla 13: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera no compensatoria.	60
Tabla 14: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera no compensatoria.	60
Tabla 15: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=1$) y 2 segmentos de individuos.....	61
Tabla 16: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=1$) y 2 segmentos de individuos.	62
Tabla 17: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=1$) y 10 segmentos de individuos.....	62
Tabla 18: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=1$) y 10 segmentos de individuos.....	62
Tabla 19: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=3$) y 2 segmentos de individuos.....	63
Tabla 20: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=3$) y 2 segmentos de individuos.	63
Tabla 21: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=3$) y 10 segmentos de individuos.	63
Tabla 22: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=3$) y 10 segmentos de individuos.	64
Tabla 23: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera no compensatoria y 2 segmentos de individuos.	64
Tabla 24: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera no compensatoria y 2 segmentos de individuos.	65

Tabla 25: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera no compensatoria y 10 segmentos de individuos.	65
Tabla 26: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera no compensatoria y 10 segmentos de individuos.	65
Tabla 27: Atributos y niveles por atributo del set de datos empírico.....	68
Tabla 28: Producto óptimo obtenido con programación dinámica a partir de estimación con modelo compensatorio para conjunto de datos reales asumiendo preferencias compensatorias.....	73
Tabla 29: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación con modelo compensatorio para conjunto de datos reales asumiendo preferencias compensatorias.....	73
Tabla 30: Producto óptimo obtenido con programación dinámica a partir de estimación con modelo no compensatorio para conjunto de datos reales asumiendo preferencias compensatorias.....	74
Tabla 31: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación con modelo no compensatorio para conjunto de datos reales asumiendo preferencias compensatorias.....	74
Tabla 32: Producto óptimo obtenido con programación dinámica a partir de estimación con modelo no compensatorio para conjunto de datos reales asumiendo preferencias no compensatorias.....	75
Tabla 33: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación con modelo no compensatorio para conjunto de datos reales asumiendo preferencias no compensatorias.....	75
Tabla 34: Producto óptimo obtenido con programación dinámica a partir de estimación con modelo compensatorio para conjunto de datos reales asumiendo preferencias no compensatorias.....	76
Tabla 35: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación con modelo compensatorio para conjunto de datos reales asumiendo preferencias no compensatorias.....	76
Tabla 36: Estructura de costos para diseño de producto óptimo con datos reales de elección de cámaras digitales.	87

Índice de figuras

Gráfico 1: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 32 preguntas y 4 alternativas por pregunta.	32
Gráfico 2: Hit rate por individuo para simulación con 8 atributos, 4 niveles por atributo, 32 preguntas y 4 alternativas por pregunta.	32
Gráfico 3: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 32 preguntas y 4 alternativas por pregunta.	32
Gráfico 4: Tasa de recuperación del orden de atributos para simulación con 8 atributos, 4 niveles por atributo, 32 preguntas y 4 alternativas por pregunta.	32
Gráfico 5: Hit rate por individuo para simulación con 5 atributos, 3 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.	33
Gráfico 6: Hit rate por individuo para simulación con 5 atributos, 6 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.	33
Gráfico 7: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 3 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.	33
Gráfico 8: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 6 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.	33
Gráfico 9: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 4 alternativas por pregunta.	34
Gráfico 10: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.	34
Gráfico 11: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 4 alternativas por pregunta.	34
Gráfico 12: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.	34
Gráfico 13: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 3 alternativas por pregunta.	35
Gráfico 14: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 5 alternativas por pregunta.	35
Gráfico 15: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 3 alternativas por pregunta.	35

Gráfico 16: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 5 alternativas por pregunta.	35
Gráfico 17: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y $w=1/3$	36
Gráfico 18: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y $w=1/10$	36
Gráfico 19: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y $w=1/3$	36
Gráfico 20: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y $w=1/10$	36
Gráfico 21: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 32 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta =3$).	42
Gráfico 22: Hit rate por individuo para simulación con 8 atributos, 4 niveles por atributo, 32 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta =3$).	42
Gráfico 23: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 32 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta =1$).	42
Gráfico 24: Hit rate por individuo para simulación con 8 atributos, 4 niveles por atributo, 32 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta =1$).	42
Gráfico 25: Hit rate por individuo para simulación con 5 atributos, 3 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta =3$).	43
Gráfico 26: Hit rate por individuo para simulación con 5 atributos, 6 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta =3$).	43
Gráfico 27: Hit rate por individuo para simulación con 5 atributos, 3 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta =1$).	43
Gráfico 28: Hit rate por individuo para simulación con 5 atributos, 6 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta =1$).	43
Gráfico 29: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta =3$).	44
Gráfico 30: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta =3$).	44
Gráfico 31: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta =1$).	44
Gráfico 32: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta =1$).	44

Gráfico 33: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 3 alternativas por pregunta y condición de bajo ruido ($\beta = 3$).	45
Gráfico 34: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 5 alternativas por pregunta y condición de bajo ruido ($\beta = 3$).	45
Gráfico 35: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 3 alternativas por pregunta y condición de alto ruido ($\beta = 1$).....	45
Gráfico 36: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 5 alternativas por pregunta y condición de alto ruido ($\beta = 1$).....	45
Gráfico 37: Distribución de la elección por nivel de precio.....	69
Gráfico 38: Distribución de la elección por nivel de precio y por pregunta.....	69
Gráfico 39: Distribución de la elección por nivel de resolución.....	69
Gráfico 40: Distribución de la elección por nivel de resolución y por pregunta.....	69
Gráfico 41: Distribución de la elección por nivel de batería.....	70
Gráfico 42: Distribución de la elección por nivel de batería y por pregunta.	70
Gráfico 43: Distribución de la elección por nivel de zoom.	70
Gráfico 44: Distribución de la elección por nivel de zoom y por pregunta.	70
Gráfico 45: Distribución de la elección por nivel de tamaño.	71
Gráfico 46: Distribución de la elección por nivel de tamaño y por pregunta.	71
Gráfico 47: Elección de tipo de batería según nivel de precio.....	72
Gráfico 48: Elección de tipo de resolución según nivel de precio.	72
Gráfico 49: Elección de tipo de zoom según nivel de precio.	72
Gráfico 50: Elección de tipo de tamaño según nivel de precio.....	72

Capítulo 1

Introducción

Actualmente, el retail enfrenta un estado de fuertes cambios desde el punto de vista del consumidor, propiciado por nuevas generaciones cada vez más exigentes y dependientes de la tecnología. Dicha situación está afectando la variedad y surtido de productos, junto al diseño de los espacios comerciales y las estrategias de marketing. De esta forma, ante la gran variedad y oferta de bienes y servicios disponible, la construcción de ventajas competitivas sostenibles se sustenta en entender las preferencias y comportamiento de los consumidores, proporcionándoles productos que se ajusten mejor a sus necesidades.

El surtido de productos de una compañía está definido por aquellos artículos que dispone en una determinada tienda durante un cierto periodo de tiempo. De este modo, la selección del surtido y la introducción de nuevos productos, constituyen una de las decisiones más complejas a las que se enfrenta un retailer. Por lo general, dicha elección se efectúa de manera intuitiva, sin contar con información transaccional previa de los productos considerados para disponer al público. En determinadas categorías, como lo son artículos tecnológicos y bienes de carácter duradero, la gran cantidad de opciones disponible, así como la presencia de espacios de visualización limitados y restricciones presupuestarias contribuyen a la complejidad de la decisión. De este modo, dado un determinado espacio en tienda o un canal web, el surtido de productos involucra un *trade-off* entre tres elementos: el número de categorías a considerar por el retailer, el total de productos a disponer en cada una de ellas y, por último, la cantidad de inventario a tener por cada artículo.

A dicha dificultad, se suma la heterogeneidad de los consumidores, cuyo nivel de preferencia por los atributos de un producto es diferente entre segmentos, por lo que un diseño generalizado basado en la intuición o, errado respecto al proceso de toma de decisión de los individuos, no generará un stock de productos óptimo que se ajuste a las preferencias reales de las personas. De este modo, se infiere que la gestión de la variedad e introducción de nuevos productos tiene un efecto sobre la elección de los individuos y, por ende, sobre las utilidades que recibe un retailer.

La planificación del surtido e introducción de nuevos productos tiene como finalidad determinar la selección de artículos y atributos que maximiza las utilidades y el porcentaje de participación de mercado. Entre los factores que inciden en ello se encuentran: la cantidad de espacio disponible para exhibición, la incorporación de algunas marcas imprescindibles, renovaciones por temporada, cambio en las preferencias de los consumidores, entre otras.

Como una forma de organización, la mayoría de los retailers segmenta sus productos o *Stock Keeping Units (SKU)* en categorías, las que constituyen conjuntos de artículos que poseen características o atributos en común. Según el departamento de Agricultura de Estados Unidos, se estima que, durante 2016, tan sólo en la categoría de comida y bebestibles se introdujo un total de 21.435 nuevos productos, cifra que se incrementa año tras año y cuyo crecimiento se extiende a otras categorías.¹

Bajo este escenario, los consumidores se enfrentan a una variedad considerable de alternativas. Ello involucra un mayor esfuerzo de evaluación durante el proceso y decisión de compra. Existe evidencia que, al procesar un gran número de productos o características, los individuos tienden a emplear heurísticas o reglas de decisión no compensatorias, que les ayudan a considerar o descartar un determinado producto del total disponible, centrándose en aquellos atributos que califican como más relevantes. (Gigerenzer et al., 1999; Drolet & Luce 2004; Bennett & Blamey, 2001).

El uso de heurísticas para la elección de alternativas depende del contexto de la decisión. De esta forma, dichas reglas son más usadas bajo ciertas condiciones, que comprenden: la existencia de un gran número de productos y características a ser evaluadas, presión por el tiempo de búsqueda, encontrarse en etapas tempranas del proceso de decisión y cuando los consumidores poseen un grado de familiarización con la categoría a evaluar.

Los principales hallazgos en materia de planificación de surtido y diseño de nuevos productos tienen como supuesto el uso de reglas de decisión compensatorias, vale decir, bajo la premisa que un consumidor considera todas las alternativas y atributos a los que se enfrenta. No obstante, si los consumidores suelen utilizar reglas no compensatorias, el diseño de productos óptimos basado en reglas compensatorias puede derivar en resultados con menor grado de acierto respecto a las preferencias de las personas.

En consecuencia, al momento de escoger el surtido o crear los productos que se disponen en una tienda o sitio web, se debe tener en consideración que existen segmentos de consumidores heterogéneos, con distintas preferencias. Cada grupo tendrá un comportamiento en particular, que determinará la forma en que considera y escoge entre distintas opciones. Así, al establecer el surtido óptimo o diseñar nuevos productos, no sólo

¹ Servicio de investigación económica, Departamento de Agricultura de EE. UU. (www.ers.usda.gov)

se deben considerar los costos y beneficios que provee la venta futura de éstos, sino, además, el modelamiento y entendimiento de los hábitos de compra de los individuos.

1.1 Motivación

Las decisiones de diseño de productos constituyen un determinante en la decisión de compra de los consumidores y en las utilidades que recibe una compañía desarrolladora. Al momento de vincular dicho proceso con el comportamiento de los consumidores se enfrentan grandes desafíos, que comprenden el entendimiento de aspectos conductuales complejos, como lo son el efecto sustitución, la recolección y agregación de información y la heterogeneidad de los individuos.

Debido a una mayor variedad y disponibilidad de productos, en la actualidad, la industria del retail experimenta altos niveles de competencia, por lo que la adquisición de mayor poder de mercado depende en gran medida de la forma en que los distintos artículos que se comercializan son presentados, de manera que sean atractivos para el público. En este aspecto, resulta crucial que el diseño de productos considere las prácticas y hábitos que describen el comportamiento de compra y elección de los consumidores. Con ello se logra determinar los atributos y tipos de productos que reciben mayor atención y aceptación, presentándose una oportunidad para determinar conjuntos de productos óptimos que maximizan el bienestar de los clientes y las ganancias, destinando recursos para el desarrollo de los atributos o bienes con mayor relevancia.

De esta manera, las personas se enfrentan a una gran cantidad de decisiones, afectadas en gran medida por el contexto en que se realizan, como lo es el número y tipo de alternativas disponibles. Cuando los consumidores procesan una cantidad considerable de productos o características suelen recurrir a heurísticas o reglas de decisión no compensatorias, que ayudan a que el proceso de elección sea más simple.

De este modo, la optimización del diseño de productos se sustenta en el estudio del comportamiento de los individuos a través de distintos modelos de elección, basados en sus preferencias. Determinar la selección de características óptimas de un producto previo a su introducción al mercado, ya sea de un producto nuevo o un rediseño de producto, de acuerdo a las exigencias de los consumidores puede impactar de forma sustancial en las ventas y utilidades de una empresa, mientras que una inadecuada planificación puede derivar en pérdidas.

La literatura actual en materia de diseño de productos contempla preferencias compensatorias, asumiendo que un consumidor considera todas las alternativas y atributos disponibles al momento de tomar la decisión de compra. No obstante, si los consumidores escogen de manera no compensatoria, el diseño de productos basado en preferencias compensatorias puede resultar poco efectivo.

Este trabajo pretende evaluar el diseño de productos bajo reglas de decisión no compensatorias, como una forma de valorar si dicha distinción tiene un efecto significativo respecto al uso de modelos compensatorios. De esta forma, la contribución al estado del arte es brindar un enfoque no compensatorio al diseño de productos. En otras palabras, se buscará determinar la forma en que cambia la configuración del diseño de productos óptimo cuando se utilizan reglas no compensatorias, estimando las preferencias de los consumidores, en base a datos simulados y empíricos, para luego incorporarlas a los respectivos modelos de optimización.

Si bien el estudio de surtido de productos bajo preferencias no compensatorias es relevante para definir qué perfiles disponer en una página de productos o góndola, para efectos de este estudio, se centrará el análisis en el diseño de productos. De algún modo, el retailer construye el surtido en base a las opciones de productos presentes en el mercado, por lo que se trabajará en una de las primeras etapas de la cadena de suministro y se considerará el diseño de surtido como un símil al diseño de productos.

Se espera que los resultados generados sean de importancia para futuras investigaciones en materia de diseño y planificación de surtido de productos, en que se exploren efectos como la competencia entre retailers, costos de producción, sustitución entre productos y cambios de temporada bajo modelos de elección no compensatorios.

1.2 Objetivos

1.2.1 Objetivo general

Estudiar el diseño de productos cuando los consumidores utilizan reglas de decisión no compensatorias.

1.2.2 Objetivos específicos

- Estudiar reglas de decisión no compensatorias, específicamente de tipo lexicográfico, para la caracterización del comportamiento de los individuos a través de modelos de elección.
- Desarrollar una metodología para estimar el diseño de productos óptimos asumiendo preferencias no compensatorias.
- Determinar si la estimación basada en reglas de decisión no compensatorias tiene un desempeño significativamente distinto respecto al uso de preferencias compensatorias en el diseño de productos.
- Estudiar y establecer alcances y limitaciones del modelo, analizando su consistencia en distintos escenarios, tanto simulados, como reales.

1.3 Metodología

La oportunidad de investigación involucra el uso de herramientas estadísticas para el análisis y simulación de datos, así como modelos de elección y métodos de optimización. De esta forma, se ha definido la siguiente metodología de trabajo:

1.3.1 Revisión bibliográfica

Como parte inicial del trabajo se efectúa una revisión bibliográfica de las distintas investigaciones y avances en materia de preferencias del consumidor, diseño de productos y diseño de surtido de productos o *assortment planning*. Lo anterior, con la finalidad de comprender aspectos básicos vinculados al tema. Además, se realiza una recopilación literaria con la intención de determinar metodologías adecuadas para el cumplimiento de los objetivos planteados.

1.3.2 Estimación de preferencias no compensatorias

Se modelan preferencias no compensatorias, con especial énfasis en las de tipo lexicográfico a través de un algoritmo Greedy que permite obtener los órdenes de preferencia de los individuos sobre las alternativas y atributos disponibles.

1.3.3 Diseño del problema de optimización

Se desarrolla un modelo matemático que permite determinar un producto óptimo asumiendo preferencias tanto compensatorias, como no compensatorias, con la finalidad de distinguir si existen diferencias considerables entre ambos enfoques. Para ello se recurre a herramientas de programación dinámica.

1.3.4 Optimización del modelo

Se provee un diseño de productos óptimo bajo preferencias compensatorias y no compensatorias. Para ello se utilizan datos simulados, creados en función de una distribución de preferencias.

1.3.5 Validación

Se analiza la capacidad predictiva de los modelos generados, comparando su desempeño con un conjunto de datos reales pertenecientes a la industria del retail. En particular, se trabaja con información de una encuesta del tipo análisis conjunto con productos pertenecientes a la categoría de cámaras digitales.

1.4 Alcances

A continuación, se detallan los alcances que tendrá el desarrollo del proyecto, los cuales han sido establecidos a partir de la evaluación de la literatura e información disponible en materia de reglas de decisión no compensatorias y modelos de optimización de diseño de productos, junto la complejidad de considerar ciertos efectos cuya estimación se vuelve compleja dados los recursos y datos disponibles:

- Dentro de los tipos de reglas de decisión de carácter no compensatorio, se trabajará con aquellas de tipo lexicográfico.
- No se abordará el efecto de competencia entre distintos retailers al momento de diseñar nuevos productos.
- No se considerará la implementación de los resultados obtenidos en este trabajo de título.

1.5 Resultados esperados

A través del desarrollo de este trabajo de título se espera obtener una metodología que permita incorporar el efecto que tiene el uso de modelos no compensatorios en la optimización del diseño de productos.

En este aspecto, se espera obtener un marco comparativo en que se establezca bajo qué condiciones el uso de modelos no compensatorios debiese preferirse a la utilización de modelos compensatorios.

A continuación, se presenta el detalle de los resultados esperados:

- Descripción de la situación actual de los métodos de estimación de preferencias compensatorias y no compensatorias.
- Obtener una metodología que permita la identificación de los atributos más relevantes para cada individuo, representando sus preferencias a partir de modelos no compensatorios.

- Optimización del diseño de nuevos productos, incorporando preferencias no compensatorias.
- Comparación del uso un enfoque no compensatorio en el diseño de nuevos productos en contraste al enfoque compensatorio tradicional.

1.6 Estructura del trabajo de título

A continuación, se presenta la estructura del trabajo de título:

- **Capítulo 1. Introducción:** Establece el contexto del trabajo de título, describe las características generales y los pasos principales de la investigación, motivación, justificación, objetivos, alcances, además del valor que genera dentro del área de investigación.
- **Capítulo 2. Antecedentes generales:** Se muestra una revisión de la literatura relativa al tema de las reglas de decisión, diseño de productos y planificación de surtido de productos.
- **Capítulo 3. Estimación de preferencias no compensatorias:** Se representan las preferencias de los individuos a través de funciones de utilidad que cumplen con las características de la regla lexicográfica, correspondiente a una regla de decisión no compensatoria. A su vez, en base a los datos disponibles, se representan dichas preferencias empleando algoritmos que modelan el comportamiento de los individuos.
- **Capítulo 4. Estimación de preferencias compensatorias:** Se representan las preferencias de los individuos a través de la regla de decisión compensatoria, precisando la metodología de simulación y estimación.
- **Capítulo 5. Diseño de producto óptimo bajo preferencias simuladas:** A través de uso de programación dinámica y búsqueda exhaustiva, se efectúan estimaciones del diseño de productos óptimos utilizando un enfoque compensatorio y no-compensatorio, basado en la regla de decisión lexicográfica. Se presentan los principales resultados de la comparación y las conclusiones obtenidas.

- **Capítulo 6. Diseño de producto óptimo bajo preferencias reales:** A través de uso de programación dinámica y búsqueda exhaustiva, se efectúan estimaciones del diseño de productos óptimos utilizando un enfoque compensatorio y no-compensatorio sobre un set de datos reales. Se compara la estimación de ambos tipos de preferencias y las posibles diferencias en los perfiles de productos obtenidos con cada uno.
- **Capítulo 7. Conclusiones:** Conclusiones finales del trabajo realizado y futuras líneas de investigación.

Capítulo 2

Antecedentes generales

2.1 Estado del arte

2.1.1 Diseño de nuevos productos

Al tratarse del diseño de nuevos productos, las investigaciones en la materia se orientan a identificar y evaluar el desempeño que tendrán éstos previo a su introducción al mercado (Kohli y Krishnamurti ,1987).

El diseño ofrece la oportunidad de abordar las necesidades emergentes de los consumidores y responder a cambios tecnológicos dinámicos, así como a cambios sociales y la incertidumbre creada por los mercados globales.

Un problema tradicional en el diseño de nuevos productos es la integración de la información que provee el consumidor. La capacidad de atender las necesidades de los clientes se ve afectada debido a la mayor dificultad técnica de contar con muchos productos (Goffin, 1998; Nambisan, 2002) y, complicada por la amplia variedad de productos existente. Al mismo tiempo, la complejidad del diseño de nuevos productos aumenta con el número de componentes del producto, el nivel de novedad e interacción entre estos (Tidd, 1995).

De este modo, la dificultad de predecir el grado de aceptación de un producto en el mercado ha sido estudiada en profundidad. El enfoque tradicional para estimar las preferencias de los consumidores por un determinado producto ha puesto énfasis en el uso de técnicas basadas en atributos. Con ello, se busca determinar la configuración óptima de un producto a partir de un conjunto de posibles atributos y niveles de atributo, que se utilizan como punto de partida para caracterizar las alternativas a las que se enfrentan las personas.

Las principales técnicas basadas en atributos incluyen el análisis conjunto, el escalamiento multidimensional y el análisis de valor de múltiples atributos (Shi, Olafsson y Chen, 2001; Tomkovick y Dobie, 1995).

De esta forma, cuando se emplea análisis conjunto, las preferencias de los consumidores se representan mediante una función de utilidad, en que, dado un número de alternativas, un individuo escoge aquella que maximiza su bienestar.

Una forma de optimizar el diseño de nuevos productos es mediante heurísticas de programación dinámica. De este modo, el producto óptimo queda definido como aquel ítem que maximiza el porcentaje de elección, vale decir, el número de consumidores que lo prefieren (Kohli y Krishnamurti, 1987).

La literatura en torno al diseño óptimo de productos que utiliza técnicas basadas en atributos emplea preferencias compensatorias, asumiendo que un consumidor evalúa todos los atributos al momento de discriminar entre una alternativa u otra, siendo capaz de compensar la utilidad de aquellos que considera peores con una mayor valoración en otras características.

2.1.2 Assortment planning

Se define como surtido o *assortment* al set de productos que dispone una tienda en un determinado momento. De esta forma, el diseño de surtido o *assortment planning* busca maximizar las ventas y el porcentaje de participación de mercado sujeto a ciertas restricciones.

Los modelos de diseño de surtido de productos se construyen basándose en temas como: variedad de productos y diseño de línea de productos, asignación de espacio en estanterías, sistemas de inventario y percepción de variedad por parte del consumidor.

De igual forma, en la planificación del surtido de productos, parte importante es la definición de los modelos de demanda que describen el comportamiento del consumidor. Entre los principales modelos utilizados se encuentran: Modelo Logit Multinomial (MNL), Modelo de Demanda Exógena y Modelos de Elección de Localización. El logit multinomial es un modelo de elección discreto, que asume consumidores racionales que maximizan su utilidad al escoger entre alternativas. Por otro lado, los modelos de demanda exógena especifican directamente la demanda de cada producto y las acciones que toma un individuo si dicho producto no está disponible. Finalmente, el modelo de elección de localización es otro modelo basado en utilidad, pero que incorpora el efecto de la distancia de cada producto respecto a un estado ideal del consumidor.

Las investigaciones en surtido de productos y manejo de inventario han avanzado rápidamente en los últimos años, particularmente en modelar la sustitución entre productos aplicando teorías de elección del consumidor a nivel individual. De esta forma, el modelo Multinomial (MNL) ha sido comúnmente usado por investigadores.

Van Ryzin y Mahajan (1999) fueron los primeros en estudiar la planificación del surtido y las decisiones de inventario bajo el modelo MNL, determinando que el surtido óptimo consiste en los productos más populares del conjunto de alternativas potenciales a ofrecer.

La decisión de agregar un nuevo producto a un surtido no sólo debe considerar los costos y beneficios por venta del producto, sino, además, el beneficio indirecto que tiene un surtido más amplio en simplificar los costos de búsqueda del consumidor. Cachon et al. (2005) generalizan el proceso de elección de los individuos para incorporar los costos de búsqueda, demostrando que ignorar la búsqueda de los consumidores en la estimación de la demanda puede resultar en un surtido con menor variedad y beneficios esperados significativamente más bajos en comparación a la solución óptima.

Miller et al. (2010) consideran el problema de selección del surtido con clientes heterogéneos y prueban el impacto de diferentes modelos de elección del consumidor en el surtido óptimo. Ellos desarrollan un modelo de elección secuencial en el cual los clientes forman primero conjuntos de consideraciones y luego hacen elecciones de productos basadas en el modelo MNL. De acuerdo a su análisis, los consumidores simplifican la elección formando un set de consideración de productos que evaluarán de forma más detenida antes de tomar una decisión de compra. Saber qué productos forman el set de consideración puede llegar a explicar un 80% de la incertidumbre que podría explicar un MNL. Ello evidencia de que el uso de heurísticas puede ser más eficiente y llevar a mejores elecciones que razonamientos más complejos.

Recientemente, se han incorporado modelos de orden de preferencias para estimar el diseño de productos. De esta forma, Honhon et al. (2010) estudian el problema de planificación de inventario cuando los clientes se clasifican según su orden de preferencia de productos. Suponen que la demanda total del cliente es aleatoria, y el mercado está compuesto de proporciones fijas de diferentes tipos de clientes, basadas en el orden de preferencia.

En contraste con la literatura expuesta, el uso del Modelo Multinomial y las investigaciones efectuadas hasta el momento no consideran el uso de modelos de preferencia no compensatorios para la estimación del surtido de productos óptimo. De esta forma, el perfil de surtido y la predicción de participación de mercado de los productos que lo componen podría diferir substancialmente, dependiendo si se emplea un enfoque compensatorio o no compensatorio.

2.1.3 Reglas de decisión

Generalmente, los consumidores procesan un gran número de productos o características al momento de la decisión de compra. Para ello, en determinadas ocasiones, recurren a heurísticas para considerar o descartar un determinado producto del total de alternativas disponibles, a las que se denomina reglas no compensatorias.

Tradicionalmente, los principales modelos de decisión se basan en la idea de maximización de utilidad, como lo son los modelos de peso aditivo (*Weighted Additive Models*) derivados del análisis conjunto. El análisis conjunto se utiliza para evaluar cómo las distintas características de un producto contribuyen en las preferencias por dicho producto. De esta forma, en presencia de pocos atributos y opciones a considerar, la aproximación estándar (*Weighted Additive Models*) provee una exactitud predictiva más alta para tareas tanto de ranking, como de rating, mientras que el uso de un modelo no compensatorio puede tener un mejor desempeño pronosticando la elección de una gran cantidad de atributos. (Diekmann et al, 2009).

Es así como el estudio de reglas de decisión no compensatorias cobra gran relevancia para la industria. Sin embargo, no todas las decisiones que se efectúan implican el uso de heurísticas. En consecuencia, el uso de heurísticas depende del contexto de la decisión. De esta forma, las reglas heurísticas son más usadas cuando hay un gran número de productos y características a ser evaluadas, cuando existe mayor presión por tiempo, el consumidor se encuentra en etapas tempranas del proceso de elección y cuando los consumidores están más familiarizados con la categoría a evaluar (Hauser, et al, 2009).

Entre los tipos de reglas de decisión no compensatorias se distinguen: Disyuntivas, Conjuntivas, Subconjunto Conjuntivo, Disyunciones de Conjunciones, Lexicográficas y Eliminación por Aspectos.

Las reglas de decisión no compensatorias, aplicadas a elecciones de consideración, especialmente las de tipo lexicográfico, han tenido mayor atención por parte de la comunidad académica en los últimos años por su relevancia para la gestión en la industria. Según las investigaciones realizadas en la materia, en determinados escenarios, las reglas de decisión no compensatorias pueden predecir mejor que las meramente compensatorias (Diekmann et al, 2009) y las nuevas tecnologías están permitiendo medir su impacto con nuevos métodos de estimación.

De aquí nace el interés por aplicar dichas reglas al diseño de productos, con el fin de evaluar si existen diferencias substanciales respecto al uso de reglas compensatorias.

2.2 Marco conceptual

Parte del objetivo principal del trabajo de título consiste en la evaluación de las preferencias de los consumidores. Para ello se debe conocer la forma en que los individuos escogen dentro de un grupo de alternativas. A continuación, se describen los principales conceptos que guardan relación con esta materia:

2.2.1 Análisis conjunto

El análisis conjunto constituye una metodología que permite, por medio de una encuesta, analizar las preferencias de los consumidores frente a determinados productos. Cada perfil de producto se construye a partir de atributos, cada uno definido por un cierto número de niveles, utilizando un diseño factorial (Green & Rao, 1971).

Mediante información de las preferencias y empleando métodos de optimización, es posible encontrar una combinación de atributos adecuada que representa el producto con mayor aceptación.

El análisis de los datos obtenido resulta apropiado para conocer la importancia relativa de los atributos, la utilidad de ciertos perfiles de producto y contrastar diferencias entre individuos. Comercialmente, esta técnica es ampliamente utilizada en simulaciones de marketing para introducir nuevos productos al mercado, realizar segmentaciones de clientes y aplicar técnicas de pricing.

Dependiendo de los objetivos de la investigación, existen diferentes tipos de análisis conjunto. Uno de los más populares en el área de marketing corresponde al análisis conjunto basado en elecciones (Choice Based Conjoint - CBC), que corresponde a una herramienta que simula procesos de selección de opciones individuales (Haaijer and Wedel, 2007), en que los encuestados expresan sus preferencias escogiendo productos de un set, en lugar de efectuar un rating o un ranking de ellos.

Una de las ventajas de este tipo de análisis se fundamenta en que los datos recolectados simulan decisiones de compra, una característica más realista y simple de entender para los encuestados. Además, a diferencia de los Análisis Conjunto tradicionales, el CBC resulta más flexible en términos de diseño, pues abarca una amplia gama de elección, permitiendo la incorporación de productos alternativos, productos de la competencia, la opción de No Comprar, entre otras variantes (DeSarbo et al., 1995) y (Louviere and Woodworth, 1983).

2.2.2 Reglas de decisión

Se entienden como el conjunto de condiciones bajo las cuales se evalúan alternativas de productos o servicios. Cuando se realiza una elección existen distintos criterios empleados por el tomador de decisión. Este estudio se enfocará en modelos de elección caracterizados por preferencias no compensatoria (Grether & Wilde, 1984).

2.2.2.1 Reglas de decisión compensatorias

Los métodos de elección compensatorios son aquellos en que el consumidor elige considerando todos los atributos o beneficios de un producto o servicio, de modo que la ausencia de atributos de una alternativa se compensa con la presencia de otras características.

2.2.2.2 Reglas de decisión no compensatorias

Al igual que los métodos de elección compensatorios, los métodos de elección no compensatorios involucran un trade-off entre las utilidades que representa cada atributo de una alternativa. No obstante, la ausencia de uno no puede ser compensada por otros, puesto que en el proceso de decisión el individuo descarta alternativas, basado en criterios definidos previamente. Al existir igualdad de alternativas, el individuo continúa el proceso con el atributo que considere más relevante. Según el tipo de método utilizado existen distintas estrategias para evaluar y escoger alternativas. Entre las principales reglas de carácter no compensatorio se encuentran:

- a) **Reglas lexicográficas:** Los atributos se clasifican por importancia. Luego se examinan todas las alternativas en términos del atributo más importante y se escoge aquella con la utilidad más alta. Si más de una alternativa es seleccionada, se considera el segundo atributo más importante y se continúa hasta que sólo quedara una alternativa.
- b) **Reglas conjuntivas:** Se define un estándar mínimo para cada atributo. Se consideran todas las alternativas que cumplan con dicha condición para todos los atributos.
- c) **Reglas disyuntivas:** Se consideran todas las alternativas que cumplan con el estándar mínimo definido para alguno de los atributos.

Además, se encuentran las reglas de Subconjunto Conjuntivo, Disyunción de Conjunciones, Eliminación por Aspectos y reglas q-Compensatorias.

2.2.3 Métodos de estimación y caracterización de preferencias

Para estimar las reglas de decisión que utilizan los consumidores se han realizado variados estudios, cuya metodología consiste en aplicar una encuesta del tipo análisis conjunto. En ella, los encuestados deben mostrar sus preferencias por una determinada combinación de productos. Con los datos obtenidos y la información de la encuesta, se proponen modelos para predecir las elecciones, reglas de decisión y orden de importancia de atributos.

2.2.3.1 Logit Multinomial

Un producto puede representarse mediante un perfil de n atributos, cada uno de los cuales posee un número finito de niveles. De esta forma, se define L_k al conjunto de niveles del atributo k ($k = 1, \dots, n$). A su vez, se define x como el perfil de un producto, cuyo nivel para el atributo k queda representado por $x_k \in L_k$ ($k = 1, \dots, n$).

$$x = (x_1, \dots, x_n)$$

Si se cuenta con m consumidores, es posible definir D_i como el conjunto de productos que son considerados o han sido adquiridos por el consumidor i ($i = 1, \dots, m$). De esta forma, D_i representa el set de consideración del individuo i .

Cuando se introduce un nuevo producto y al mercado, el set de consideración del individuo i se puede caracterizar como:

$$C_i = D_i \cup y$$

Así, cada individuo asocia una utilidad $u_i(x)$ al perfil del producto x . Cuando se emplea un logit multinomial asumiendo compensación entre atributos, dicha función de utilidad puede representarse como:

$$u_i(x) = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \dots + \beta_n * x_n + \varepsilon_i$$

Por otro lado, se define una regla de elección probabilística como una función:

$$\pi_i(x) = \pi(u_i(x) | x \in C_i)$$

Por definición, $0 \leq \pi_i(x) \leq 1$, para todo $x \in C_i$ y $\sum_{x \in C_i} \pi_i(x) = 1$, para todo $i = 1, \dots, m$.

En consecuencia, la probabilidad de elección promedio asociada al perfil de producto y puede definirse como:

$$\bar{\pi}(y) = \frac{1}{m} \sum_{i=1}^m \pi_i(y)$$

Cuando el valor de m aumenta, $\bar{\pi}(y)$ se aproxima al valor de la probabilidad esperada de compra, o equivalentemente al porcentaje de participación de mercado del perfil de producto y .

Para una muestra fija de m consumidores, optimizar $\bar{\pi}(y)$ equivale a maximizar la suma de las probabilidades de elección individuales:

$$\text{Maximizar } \sum_{i=1}^m \pi_i(y)$$

$$\text{con } y = (y_1, \dots, y_n) \mid y_k \in L_k \quad \forall k = 1, \dots, n$$

De esta forma, al analizar datos empleando un modelo multinomial, el modelo de optimización de las preferencias considera las siguientes probabilidades de elección individuales:

$$\pi_i = \frac{\exp(u_i(y))}{\sum_{x \in C_i} \exp(u_i(x))}$$

2.2.3.2 Modelo lexicográfico

Para estimar preferencias lexicográficas, se ha probado que dada una cantidad n de atributos de un producto, una secuencia de pesos w_1, \dots, w_n representa un orden lexicográfico sobre una construcción binaria de atributos x_1, \dots, x_n , si $w_k > w_{k+1} + \dots + w_n$, para todo $k = 1, \dots, n - 1$. De esta forma, no existe una combinación del resto de los atributos, de menor prioridad, en que la utilidad generada sea mayor o igual a la que genera un atributo de mayor prioridad.

$$u_i(x) = w_1 * x_1 + w_2 * x_2 + \dots + w_n * x_n$$

En particular, si $w_k = \frac{1}{2^k}$, se tiene una representación de las preferencias lexicográficas de un individuo. (Kohli y Jedidi 2007).

Para obtener los órdenes de prioridad de la regla lexicográfica se puede recurrir al uso de un algoritmo Greedy, que será descrito a continuación.

2.2.4 Heurística Greedy

Esta heurística funciona por medio del algoritmo Greedy, el cual busca soluciones que son óptimas localmente, para así encontrar un óptimo global (Chvatal, 1979).

Un concepto relevante para el uso de esta heurística corresponde al hit rate, el cual es un indicador del porcentaje de veces que el valor predicho por un modelo, coincide con el valor real.

De esta forma, en el contexto de la heurística mencionada, el hit rate define el porcentaje de los casos en que las alternativas arrojadas por Greedy coinciden con las alternativas reales escogidas por los encuestados.

A continuación, se explican los pasos que sigue la Heurística Greedy:

1. Se definen las variables específicas del problema de ordenamiento de atributos, como el número de individuos, elección de alternativas, cantidad de atributos, niveles de atributo, número de preguntas y alternativas por pregunta.
2. Para cada individuo, se asume una función de utilidad lexicográfica, de modo que, en una primera iteración, cada atributo se considera como el más importante, asignándole el coeficiente w_1 y se calcula el hit rate por atributo.

3. Se selecciona el atributo que posee el máximo hit rate, vale decir, aquel en que la utilidad dada por orden de preferencias asignado genera un mayor número de elecciones que coinciden con las preferencias reales de los individuos, fijándose como el atributo más importante.
4. El atributo escogido es removido del conjunto inicial de atributos y se continúa explorando el segundo atributo más importante sobre los atributos restantes, a los que se asigna el coeficiente w_2 . Se escoge aquel atributo que, junto al atributo más importante, genera una utilidad que tiene un mayor hit rate respecto a las preferencias reales de los individuos.
5. El proceso se repite, calculando los *hit rates* de las elecciones hasta ordenar todos los atributos.

Este tipo de heurística resulta útil para predecir el orden de importancia de atributos, especialmente para individuos que utilizan reglas de decisión no compensatorias como la lexicográfica. Cabe mencionar que este método no es capaz de asegurar un óptimo global en términos de órdenes de importancia, ya que trabaja definiendo soluciones parciales en cada etapa, sin replantear el resultado obtenido en etapas anteriores. El próximo capítulo aborda la estimación de preferencias no compensatorias mediante el uso de esta heurística, detallando el algoritmo matemático subyacente y presentando su desempeño en distintos escenarios de simulación.

Capítulo 3

Estimación de preferencias no compensatorias

3.1 Representación de modelos de preferencia lexicográficos

3.1.1 Regla lexicográfica

De acuerdo a la teoría económica, un consumidor que se rige por un comportamiento racional debe ser consciente de todas las alternativas disponibles al momento de realizar una elección de compra, siendo capaz de efectuar una correcta evaluación de cada una de ellas para seleccionar la configuración óptima (Schiffman & Kanuk 2007). Sin embargo, la habilidad de los consumidores para efectuar dicha elección puede estar altamente condicionada, ya sea por información limitada y poco adecuada, por aspectos motivacionales, restricciones de tiempo o por valores y relaciones sociales (Simon 1997). De esta forma, se infiere la pertinencia que tiene el uso de métodos de simplificación para comparar alternativas a través de un número limitado de características.

Una forma de simplificar el proceso de elección es a través de reglas de decisión no compensatorias, mediante las cuales un consumidor utiliza aspectos mínimos de aceptación para calificar productos, tales como establecer rangos de precio o tamaño mínimo a considerar. De esta forma, si un atributo relevante no cuenta con el requisito de mínima aceptación, dicho producto no será considerado, a pesar de su valoración en otros aspectos.

Una de las reglas comúnmente estudiadas corresponde a la regla lexicográfica, por medio de la cual un consumidor realiza un ranking de los atributos de un producto según su orden de relevancia. Si un producto se considera mejor que el resto para el atributo más importante, este será escogido por el consumidor. Del mismo modo, si dos o más productos se consideran igualmente atractivos en el atributo más importante, se comparan las características del segundo atributo más importante. Este proceso se repite hasta que no existan empates (Engel, Blackwell and Miniard, 1995). Por ejemplo, un consumidor puede decidir que el tamaño es el atributo más importante de un televisor,

optando por tamaños superiores a 50 pulgadas, mientras que la resolución es el segundo atributo más importante. Si solo una de las opciones de televisores mide más de 50 pulgadas, el comprador escogerá dicho producto, no importando los televisores de otros tamaños. Si dos o más televisores miden más de 50 pulgadas, el consumidor escogerá aquél con la mejor resolución.

La regla lexicográfica ha sido vastamente estudiada en la literatura, existiendo gran evidencia respecto a su uso. Investigaciones pasadas señalan que las personas optan por este tipo de criterio incluso frente a decisiones importantes como la elección de tratamientos médicos para el cáncer, siendo conscientes que en dicho caso el uso de reglas de decisión compensatorias sería más apropiado (Kahn & Baron, 1995). Al mismo tiempo, los estudios realizados por Yee, Dahan, Hauser & Orlin (2007), muestran que aproximadamente dos tercios de los participantes de su experimento emplearon algún tipo de regla lexicográfica en la elección de smartphones. Por otro lado, la evidencia empírica mostrada por Campbell, Hutchinson & Scarpa (2006) en un estudio de preferencias sobre territorios rurales en Irlanda, señala que la utilización de una regla lexicográfica afecta la disposición a pagar de las personas, de modo que para algunos encuestados la elección de una alternativa se basa únicamente en la evaluación de los atributos que consideran más relevantes.

Si bien el uso de preferencias lexicográficas es un indicador de que algunos atributos presentes en el set de consideración no son relevantes a nivel individual, existen otros factores que pueden influenciar el uso de este tipo de reglas de decisión. Ente ellos se encuentran factores propios del diseño experimental, como lo son el nivel de complejidad (DeShazo and Fermo, 2002; Heiner, 1983; Swait and Adamowicz, 2001) o el tipo de atributos presentados (Blamey et al., 2001). Por otro lado, entre los factores externos se encuentran habilidades cognitivas, creencias y características sociodemográficas propias de cada individuo (Rosenberger et al., 2003).

Dado que la habilidad de las personas para realizar comparaciones entre alternativas se encuentra limitada por su capacidad cognitiva, para un consumidor que se enfrenta a problemas de elección complejos, que involucran la evaluación de múltiples alternativas y características, será más complicado valorar y comparar diferencias entre opciones. (Drolet & Luce 2004). En experimentos de elección discreta, la complejidad cognitiva que enfrentan los individuos está relacionada a la cantidad de alternativas presentes en el set de consideración, al total de atributos, la correlación de atributos entre alternativas y el número de repeticiones (Bennett and Blamey, 2001; Caussade et al., 2005).

De este modo, las personas tienden a emplear estrategias lexicográficas en escenarios donde existe correlación entre atributos o cuando consideran que un atributo tiene una importancia relativa más alta que el resto. (Luce et al., 2000; Blamey et al., 2002).

En consecuencia, las investigaciones pasadas evidencian la existencia de la regla lexicográfica, cuyo uso se ve influenciado, tanto por factores inherentes a las preferencias

de cada persona, así como a características propias del diseño experimental y el contexto de la elección.

3.1.2 Utilidad lexicográfica

Si bien el uso de la regla lexicográfica ha sido ampliamente caracterizado a nivel conductual, el trabajo respecto a su representación matemática e inferencia ha sido limitado y menos explorado.

Según lo declarado por Debreu (1954), no es posible construir una función de utilidad que represente preferencias lexicográficas sobre dos o más atributos con valores continuos. Sin embargo, un orden de preferencias lexicográfico sobre atributos discretos sí puede representarse mediante una función de utilidad.

De este modo, Martignon y Hoffrage (2002) prueban que una secuencia de coeficientes w_1, \dots, w_n simboliza un orden lexicográfico de atributos binarios x_1, \dots, x_n , si cumplen la siguiente condición:

$$w_k > w_{k+1} + \dots + w_n \quad \forall k = 1, \dots, n - 1.$$

De esta forma, no existe una combinación de los atributos de menor prioridad, en que la utilidad generada sea mayor o igual a la que genera un atributo de mayor prioridad. En particular, si $w_k = \frac{1}{k}$, se tiene una representación de las preferencias lexicográficas de un individuo.

Posteriormente, los trabajos de Kohli muestran que todos los sistemas numéricos son lexicográficos y pueden utilizarse en la representación de preferencias lexicográficas sobre atributos discretos. (Lexicographic systems, Kohli 1999).

Con ello, es posible caracterizar preferencias lexicográficas a través de funciones de utilidad que describen un orden de importancia para atributos discretos de un determinado perfil de producto.

3.1.3 Tipos de representaciones lexicográficas

Los modelos de preferencias lexicográficas pueden clasificarse en aquellos de tipo estándar, en que se evalúan alternativas atributo por atributo, o en los de tipo no estándar, que son variaciones que permiten mayor flexibilidad para representar la forma en que se evalúan las características de un producto. A continuación, se describe cada una de ellas:

3.1.3.1 Preferencia lexicográfica estándar

El modelo lexicográfico estándar asume que una persona evalúa las distintas alternativas de un producto atributo por atributo. De esta forma, si un perfil de producto cuenta con un total de m atributos ($m \geq 2$) y cada atributo k contiene n_k niveles ($n_k \geq 2, \forall k = 1, \dots, m$), su utilidad queda representada de la siguiente forma:

$$u_i(x) = \beta_1 * x_1 + \dots + \beta_m * x_m$$

Bajo esta aproximación, x_k representa el k -ésimo atributo más relevante, con $k = 1, \dots, m$. Con ello, una condición necesaria y suficiente para que el modelo lineal anterior represente preferencias lexicográficas es que:

$$\beta_k * \min \Delta x_k > \sum_{j=k+1}^m \beta_j * \max \Delta x_j$$

Lo anterior se interpreta como que el menor cambio en la utilidad para el k -ésimo atributo más importante debe ser mayor que cualquier posible cambio en la utilidad de los atributos que lo suceden en orden de relevancia. Entendiéndose como cambio la variación en la utilidad de un atributo ante variaciones de los niveles escogidos para ese atributo.

3.1.3.2 Preferencias lexicográficas no estándar

Como una alternativa al modelo de preferencias lexicográfico estándar, se encuentra el modelo lexicográfico binario, cuya principal diferencia radica en que se elimina el supuesto de que las personas evalúan alternativas atributo por atributo, de manera que utilizan un orden sobre los niveles de atributos, en lugar de un orden por atributos.

De esta manera, los consumidores pueden ir y volver de un atributo a otro a medida que utilizan preferencias lexicográficas, clasificando las alternativas entre aquellas que cuentan con el nivel preferido entre todos los atributos y aquellas que no.

Cabe destacar que el modelo lexicográfico estándar constituye un caso particular de este modelo, coincidiendo con el escenario en que los niveles de un mismo atributo son evaluados uno tras otro.

Un modo de caracterizar este modelo es a través de la siguiente representación:

Sea $N = \sum_{k=1}^m n_k$ la sumatoria del total de niveles entre los m atributos de un producto. Sea $l = 1, \dots, N$ un orden de los N niveles de atributos en orden decreciente de preferencia. Se asocia el parámetro x_l al nivel l , que toma el valor de 1 si dicho nivel está presente en la alternativa considerada y 0 si no.

De esta forma, cada alternativa puede ser representada por el vector (x_1, \dots, x_N) , en que $x_l = 1$ para m valores de l y $x_l = 0$ para los $N - m$ valores restantes.

Finalmente, la función de utilidad está representada por:

$$u_i(x_1, \dots, x_N) = \sum_{l=1}^N \beta_l * x_l$$

Para satisfacer la condición de orden lexicográfico, los coeficientes β_l deben cumplir que:

$$\beta_l > \sum_{p=l+1}^N \beta_p \quad \forall 1 \leq l \leq N$$

De este modo, de acuerdo a lo expuesto por Kohli & Jedidi (2007), se comprueba que valores de $\beta_l = \frac{1}{2^l}$ satisfacen la condición anterior, ya que:

$$\sum_{p=l+1}^N \frac{1}{2^p} < \sum_{p=l+1}^{\infty} \frac{1}{2^p} = \frac{1}{2^l} = \beta_l \quad \forall 1 \leq l \leq N - 1$$

Finalmente, entre los modelos lexicográficos no estándar se encuentra la aproximación lexicográfica de satisfacción, que modela el hecho de que ciertos niveles de atributos resultan indiferentes para el consumidor. Bajo este enfoque, las personas crean límites superiores o inferiores para ciertos atributos, excluyendo las alternativas que no cumplen con dichos requisitos de aceptación. Para el caso de atributos nominales, la indiferencia entre niveles de atributos puede simplemente reflejar ausencia de preferencia (Kohli & Jedidi, 2007). De esta forma, para un atributo con n niveles, pueden existir desde 1 a n clases de indiferencia, en que el último caso corresponde al escenario en que cada nivel es su propia clase de indiferencia y, por ende, coincide con el modelo lexicográfico estándar.

Para incorporar la condición de satisfacción a los modelos lexicográfico estándar y binario, sólo basta con asignar el mismo entero x_k a dos o más niveles que son equivalentes en preferencia para un mismo atributo.

En definitiva, el uso de un modelo lexicográfico estándar es apropiado cuando los atributos se pueden representar como variables numéricas discretas, en que existe un orden de importancia claro, con aceptación entre la población, de los niveles de un mismo atributo. Por otro lado, cuando se tiene la presencia de atributos de tipo nominal, en que las preferencias por los niveles de un mismo atributo no son claras, es más adecuado el uso de un modelo binario.

No obstante, es común que los procesos de elección requieran la evaluación tanto de atributos cualitativos, como cuantitativos, de modo que en los atributos cuantitativos se justifique el orden de atributos por sobre el orden del total de niveles de atributo.

$$u_i(x_1, \dots, x_N) = \sum_{l=1}^N \beta_l * x_l$$

Considerando los antecedentes expuestos, este trabajo de título se desarrolla a partir de un modelo lexicográfico binario estándar, según el cual los primeros n_1 términos de la función de utilidad presentada se asocian a los niveles del primer atributo en orden creciente de preferencia. Posteriormente, se asocia los siguientes n_2 términos a los niveles del segundo atributo ordenados de manera creciente y así sucesivamente con el resto de niveles de atributos de carácter cuantitativo. De existir atributos nominales u ordinales, sus niveles pueden ser tratados de manera independiente, como un atributo distinto, sin establecer un orden entre ellos.

Debido a que, en un sentido más amplio, el enfoque utilizado permite la incorporación de variables nominales, es posible restringir el modelamiento a variables cuantitativas, como una forma de simplificación.

De esta forma, este modelo produce el mismo orden de importancia de atributos que un modelo lexicográfico estándar, en que las alternativas son evaluadas un atributo a la vez, pero en que los niveles de atributo están representados de manera binaria.

Sea (x_1, \dots, x_N) un perfil de producto, con m atributos y N niveles de atributo ordenados de manera decreciente según nivel de importancia. Con ello, la función de utilidad lexicográfica puede representarse de la siguiente forma:

$$u_i(x_1, \dots, x_N) = \beta_{11} * x_{11} + \dots + \beta_{1n_1} * x_{1n_1} + \beta_{21} * x_{21} + \dots + \beta_{mn_m} * x_{mn_m}$$

Nuevamente, los coeficientes β_{kl_k} deben cumplir con la restricción lexicográfica entre atributos y entre niveles de un mismo atributo, de modo que:

$$\beta_{kl_k} > \sum_{h=k+1}^{n_m} \beta_{hl_h} \quad ; \quad \beta_{kl_k} > \sum_{j=l_k+1}^{n_k} \beta_{kj} \quad \forall k = 1, \dots, m$$

Cabe destacar, que este enfoque permite la representación de cualquier nivel de atributo, tanto nominal, como ordinal o discreto, en función de enteros con valor 0 y 1.

3.2 Metodología de inferencia de modelos de preferencia lexicográficos

Para inferir órdenes de preferencia lexicográficos se ha recurrido a un método de estimación basado en el descrito por Kohli & Jedidi (2007), el cual representa un algoritmo de aproximación polinomial a partir de datos de tipo análisis conjunto.

Sea $x = (x_1, \dots, x_N)$ un perfil de producto compuesto por m atributos, cada uno de los cuales posee n_k niveles de atributo, de modo que $N = \sum_{k=1}^m n_k$ representa el total de niveles de atributo. De esta forma, x_k toma el valor 1 si el nivel k forma parte de dicha alternativa y 0 en caso contrario.

Sea $X = (x, \dots, x^w)$ una matriz, que representa un set de consideración con w perfiles de producto o alternativas de elección. La utilidad que reporta la alternativa j al individuo i está dada por la expresión:

$$u_i(x^j) = \sum_{l=1}^N \beta_l * x_l^j$$

De este modo, se asume que la alternativa escogida para el set de consideración X será aquella que represente mayor utilidad.

Para inferir el orden de preferencias lexicográfico que un individuo i tiene sobre los N niveles de atributos de un producto, se busca maximizar el porcentaje de veces en que la utilidad generada por un determinado orden de niveles, coincide con la alternativa real escogida para el set de consideración X . En consecuencia, se escoge aquel orden con el mayor número de éxitos para el total instancias de consideración.

En este aspecto, encontrar el orden lexicográfico que maximiza el porcentaje de aciertos constituye un problema de optimización combinatorial, que puede resolverse eficientemente por enumeración, siempre y cuando el número de posibles órdenes de niveles de atributo sea pequeño.

En caso contrario, es posible recurrir al uso de una heurística Greedy, como la que se señala:

Iteración inicial:

Sea $\{1, \dots, N\}$ un vector que denota el set de niveles del total de atributos.. Este vector puede representarse como una combinación de vectores $S_0 = \{S_{01}, \dots, S_{0m}\}$, de modo que S_{0k} representa los n_k niveles del atributo k , dispuestos en orden creciente de preferencia.

Un perfil de producto es representado por un vector $x = (x_1, \dots, x_N)$, en que el nivel l perteneciente al atributo k está identificado por un valor $x_{kl} \in [0,1]$, de manera que para cada individuo i y para cada alternativa x^j del set de consideración X , se computa:

$$u_{i1k}(x^j) = \sum_{l=1}^{n_k} \frac{1}{2^l} * x_{kl}^j, \forall k \in S_0$$

Sea $\omega = (x^j, J)$, con $J = (x^1, \dots, x^w) \setminus x^j$, los perfiles de producto que conforman el set de consideración X , siendo w igual al total de alternativas disponibles. Sea Ω equivalente al total de instancias de consideración ω a las que se enfrenta un individuo. Para cada $\omega \in \Omega$ se calcula:

$$d_{i1k}(x^j, J) = \begin{cases} 1, & u_{i1k}(x^j) < u_{i1k}(J) \\ 0, & \text{en caso contrario} \end{cases} \quad \forall k \in S_0$$

Sea

$$Z_{i1k}(\Omega) = \sum_{\omega \in \Omega} d_{i1k}(x^j, J), \quad \forall k \in S_0$$

Se denota k_1 como aquel atributo para el cual el valor de $Z_{i1k}(\Omega)$ tiene el menor valor respecto al total de atributos $k \in S_0$. De este modo:

$$Z_{i1k_1}(\Omega) = \min\{Z_{i1k}(\Omega) \mid k \in S_0\}$$

De esta forma, se selecciona k_1 como el atributo de mayor importancia. En caso de existir empate en el valor de $Z_{i1k_1}(\Omega)$, se escoge un atributo de manera aleatoria y se asigna:

$$u_{i1} = \left(\frac{1}{2}, \dots, \frac{1}{2^l}, \dots, \frac{1}{2^{n_{k_1}}} \right)$$

y

$$S_1 = S_0 \setminus \{k_1\}$$

En consecuencia, S_1 representa el set de atributos restantes, una vez que se ha escogido a k_1 como el atributo de mayor relevancia.

Recursión:

Sea $N_t = \sum_{k=1}^{t-1} n_k$ la suma de niveles de los primeros $(t - 1)$ atributos para los que se ha definido un orden de importancia. Para cada alternativa x^j del set de consideración X , se computa:

$$u_{itk}(x^j) = u_{i(t-1)k}(x^j) + \sum_{l=N_t+1}^{N_t+n_k} \frac{1}{2^l} * x_{kl}^j, \quad \forall k \in S_{(t-1)}$$

En que:

$$S_{(t-1)} = S_0 \setminus \{k_1, \dots, k_{(t-1)}\}$$

constituye el set de atributos no seleccionados tras el paso $t - 1$, mientras que:

$$u_{i(t-1)k}(x^j) = \frac{1}{2} * x_{k_1}^j + \dots + \frac{1}{2^{N(t-1)+n(t-1)}} * x_{k_{(t-1)n(t-1)}}^j$$

De esta forma, para cada $\omega = (x^j, J) \in \Omega$ se calcula:

$$d_{itk}(x^j, J) = \begin{cases} 1, & u_{itk}(x^j) < u_{itk}(J) \\ 0, & \text{en caso contrario} \end{cases} \quad \forall k \in S_{(t-1)}$$

Sea

$$Z_{itk}(\Omega) = \sum_{\omega \in \Omega} d_{itk}(x^j, J), \quad \forall k \in S_{(t-1)}$$

Se denota k_t como aquel nivel para el cual el valor de $Z_{itk}(\Omega)$ tiene el menor valor respecto al total de atributos $k \in S_{(t-1)}$. De este modo:

$$Z_{itk_t}(\Omega) = \min\{Z_{itk}(\Omega) \mid k \in S_{(t-1)}\}$$

De esta forma, se selecciona k_t como el t -ésimo atributo de mayor importancia. En caso de existir empate en el valor de $Z_{itk_t}(\Omega)$, se escoge un atributo de manera aleatoria y se asigna:

$$u_{it} = \left(\frac{1}{2^{N_t+1}}, \dots, \frac{1}{2^l}, \dots, \frac{1}{2^{N_t+n_{k_t}}} \right)$$

y

$$S_t = S_{(t-1)} \setminus \{k_t\}$$

Iteración final:

El algoritmo termina una vez que se ha obtenido el orden de preferencia para todos los atributos y sus respectivos niveles de atributo.

Si bien el uso de una heurística Greedy asegura un óptimo local, no necesariamente garantiza un óptimo global. Pese a ello, en caso de existir un orden de atributos que reproduce perfectamente un rango ordenado de atributos dado, el algoritmo Greedy descrito anteriormente resulta efectivo al momento de encontrar dicho orden.

3.3 Inferencia de modelos de preferencia lexicográficos sobre datos simulados

Con el objetivo de evaluar el desempeño de la metodología de inferencia planteada se simulon órdenes de importancia bajo distintas condiciones. En ellas se varió el número de atributos, la cantidad de niveles de atributos, el total de alternativas y el número de preguntas o sets de consideración presentados a cada individuo.

La ventaja de la simulación radica en la posibilidad de evaluar distintos escenarios, en los que se conoce con certeza la regla de decisión utilizada por cada individuo para su posterior estimación. Ello permite contrastar la efectividad de algoritmo respecto a las preferencias originales de los individuos. A continuación se describe la metodología para la creación de sets de consideración y simulación de preferencias:

En primer lugar, se definen los parámetros de la simulación: número de individuos, cantidad de atributos, niveles por atributo, alternativas y preguntas. Posteriormente, utilizando un diseño factorial, se crean sets de consideración para cada individuo.

Una vez que se ha creado la matriz de elección, se genera un orden de preferencia aleatorio de atributos para cada individuo de la muestra, que define el grado de importancia que recibe cada atributo al momento de evaluar alternativas.

De este modo, a partir de la función de utilidad lexicográfica descrita en la sección anterior, a cada nivel de atributo l se asigna un coeficiente del estilo $\frac{1}{2^l}$, que denota el orden de importancia relativo respecto al resto de los niveles de atributo, teniendo en consideración que, dentro de un mismo atributo, los niveles poseen un orden específico definido.

Dado un set de consideración $X = (x, \dots, x^w)$, con w alternativas de elección x^j , se calcula el beneficio de cada alternativa en función de la utilidad lexicográfica que define el orden de importancia de cada nivel de atributo.

$$u_i(x^j) = \sum_{l=1}^N \frac{1}{2^l} * x_l^j$$

A partir de lo anterior, se selecciona de manera determinística aquella alternativa que reporta mayor utilidad y se designa como la alternativa escogida por el individuo i en el set de consideración X .

La siguiente tabla muestra parte de un set de datos simulados con 3 atributos, con 2 niveles cada uno, 4 preguntas y 3 alternativas por pregunta.

Individuo	Pregunta	Alternativa	Atributo 1		Atributo 2		Atributo 3		Alternativa escogida
			Nivel 1	Nivel 2	Nivel 1	Nivel 2	Nivel 1	Nivel 2	
1	1	1	1	0	1	0	1	0	2
1	1	2	0	1	0	1	1	0	2
1	1	3	0	1	1	0	0	1	2
1	2	1	0	1	1	0	0	1	1
1	2	2	1	0	0	1	0	1	1
1	2	3	0	1	0	1	1	0	1
1	3	1	1	0	0	1	0	1	2
1	3	2	1	0	1	0	1	0	2
1	3	3	1	0	1	0	0	1	2
1	4	1	0	1	0	1	1	0	1
1	4	2	0	1	1	0	0	1	1
1	4	3	1	0	1	0	1	0	1

Tabla 1: Set de datos de elecciones *simuladas*.

Cada alternativa del set de consideración es representada por una combinación de niveles de atributos binarios, siendo escogida aquella que otorga mayor utilidad según el orden de preferencia lexicográfico de cada individuo.

Para evaluar el desempeño del método de inferencia de órdenes lexicográficos propuesto, se midió su desempeño en base a dos criterios:

- El hit rate o tasa de acierto, correspondiente a una medida de ajuste, que evalúa el porcentaje de predicciones correctas para la elección de alternativas entre el orden de preferencias simulado y el estimado para cada individuo.
- La recuperación del verdadero orden de los atributos, entendido como el porcentaje de veces en que el orden de preferencia estimado de un atributo coincide con el orden simulado para cada individuo.

A continuación, se presentan los resultados de la estimación para distintos escenarios de atributos, niveles de atributo y cantidad de sets de consideración. Cabe mencionar que todos los escenarios consideran una muestra de 100 individuos.

3.3.1.1 Estimación variando el número de atributos.

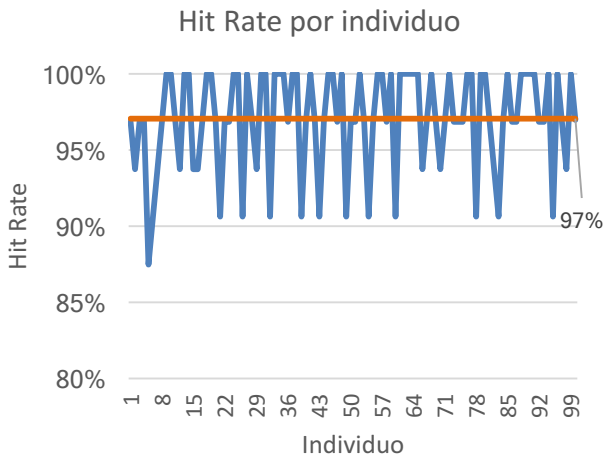


Gráfico 3: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 32 preguntas y 4 alternativas por pregunta.

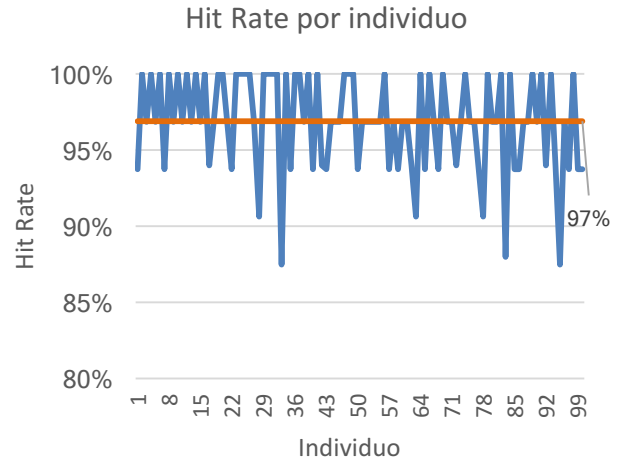


Gráfico 4: Hit rate por individuo para simulación con 8 atributos, 4 niveles por atributo, 32 preguntas y 4 alternativas por pregunta.

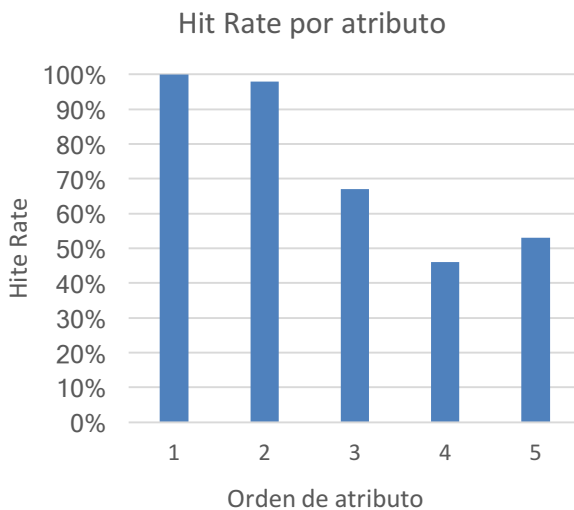


Gráfico 1: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 32 preguntas y 4 alternativas por pregunta.

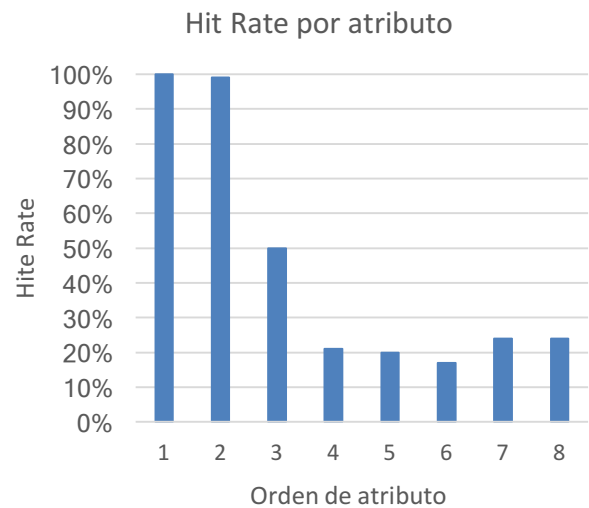


Gráfico 2: Tasa de recuperación del orden de atributos para simulación con 8 atributos, 4 niveles por atributo, 32 preguntas y 4 alternativas por pregunta.

Al variar el número de atributos de la simulación, manteniendo intacta la cantidad de niveles de atributos, preguntas y alternativas por pregunta, se aprecia que la efectividad del método de estimación decrece con el aumento de atributos. En este aspecto, el hit rate promedio por individuo al considerar 5 atributos es de un 97,1%, mientras que cuando se tienen 8 atributos, la tasa de acierto promedio es de un 96,9%. Con ello, se comprueba que el método de estimación utilizado resulta efectivo como predictor de elecciones no compensatorias ante variaciones del número de atributos.

Si bien este índice es, en ambos casos, superior al 90%, al observar la tasa de recuperación del orden de importancia de cada atributo, se aprecia que el método pierde potencia al momento de evaluar el grado de jerarquía de los atributos de menor preferencia, en especial cuándo se tiene una mayor cantidad de atributos. No obstante, si en el conjunto de elección no existen empates entre alternativas, basta utilizar un atributo para recuperar el orden lexicográfico y escoger la alternativa correcta.

3.3.2 Estimación variando el número de niveles de atributos.

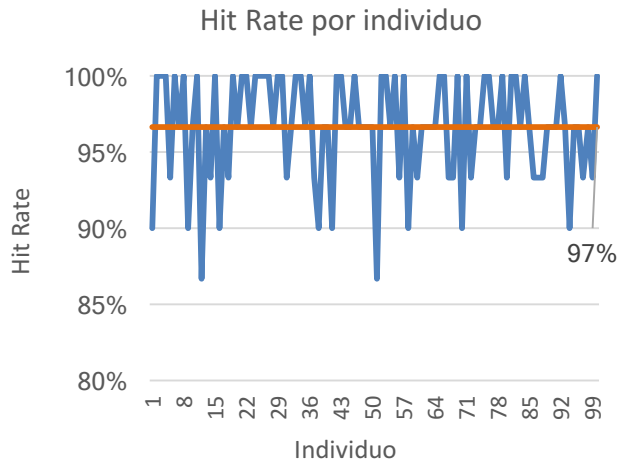


Gráfico 5: Hit rate por individuo para simulación con 5 atributos, 3 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.

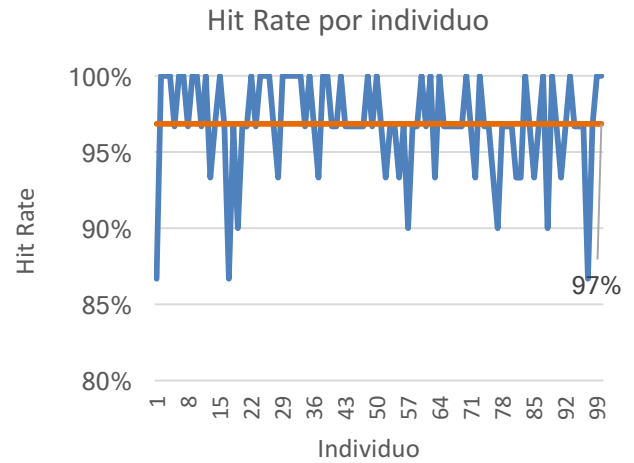


Gráfico 6: Hit rate por individuo para simulación con 5 atributos, 6 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.

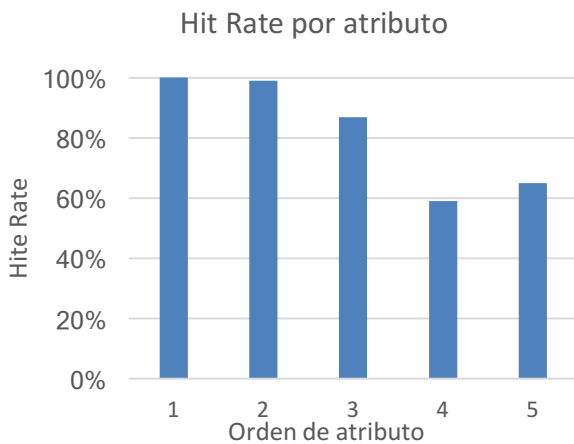


Gráfico 7: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 3 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.

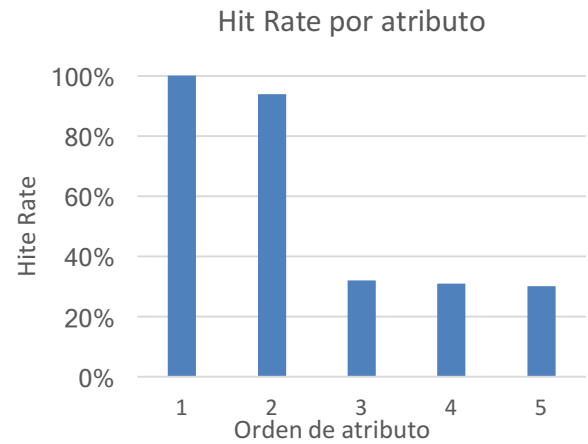


Gráfico 8: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 6 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.

Al variar la cantidad de niveles de atributos, no se aprecian grandes diferencias en el hit rate promedio por individuo. De este modo, tanto en presencia de 3, como 6 niveles por atributo, se obtienen tasas de acierto de un 97%. Por otro lado, se observan importantes diferencias en la tasa de recuperación del orden de importancia de cada atributo, de manera que con una menor cantidad de niveles el porcentaje de aciertos es mayor.

3.3.3 Estimación variando el número de preguntas.

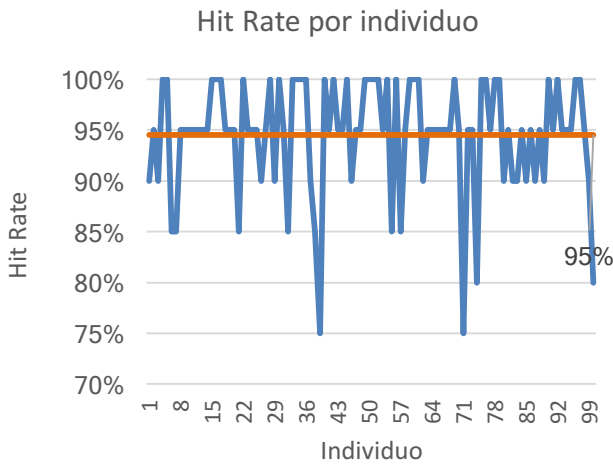


Gráfico 10: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 4 alternativas por pregunta.

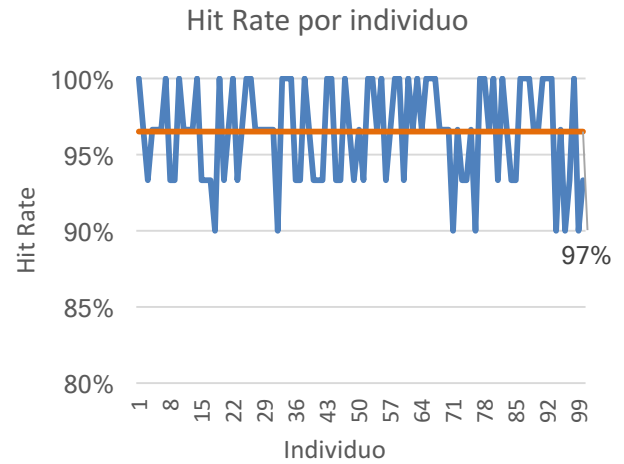


Gráfico 11: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.

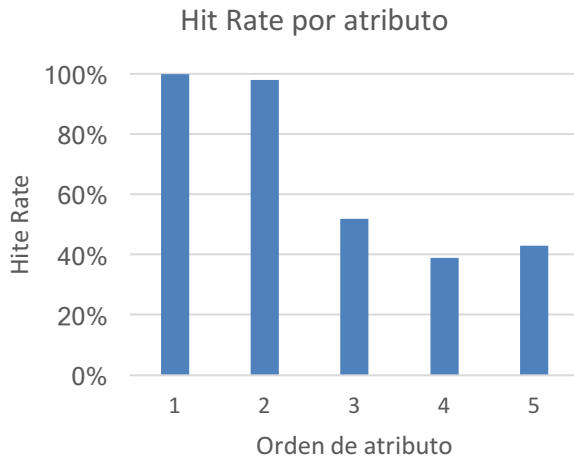


Gráfico 12: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 4 alternativas por pregunta.

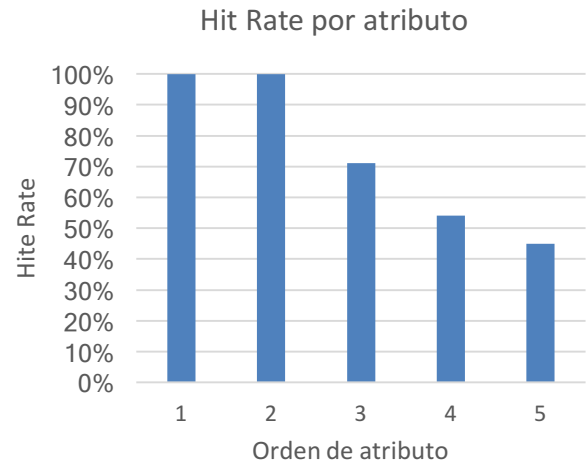


Gráfico 9: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 30 preguntas y 4 alternativas por pregunta.

A partir de los resultados anteriores, se observa que a medida que aumenta la cantidad de preguntas o sets de elección a los que se enfrenta un individuo, el porcentaje de aciertos respecto a las preferencias simuladas es mayor. A la vez, se cumple que la tasa de recuperación del orden de importancia de los atributos de menor preferencia mejora al considerar un mayor número de preguntas.

3.3.4 Estimación variando el número de alternativas por pregunta.

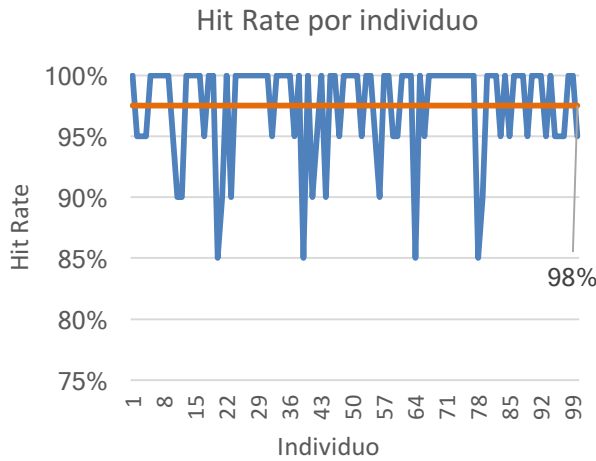


Gráfico 13: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 3 alternativas por pregunta.

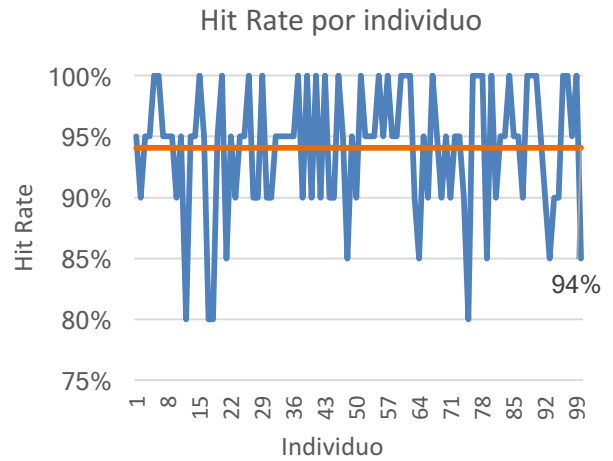


Gráfico 14: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 5 alternativas por pregunta.

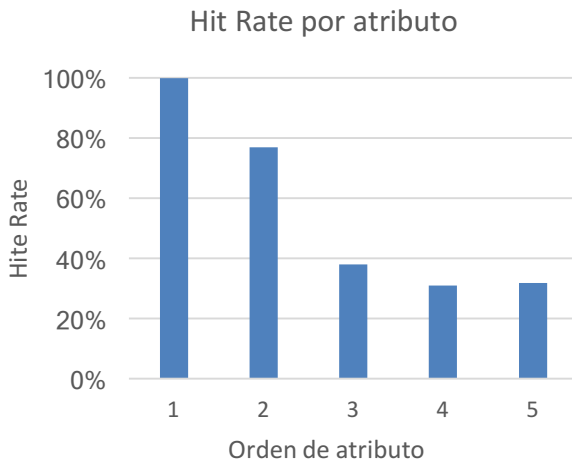


Gráfico 15: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 3 alternativas por pregunta.

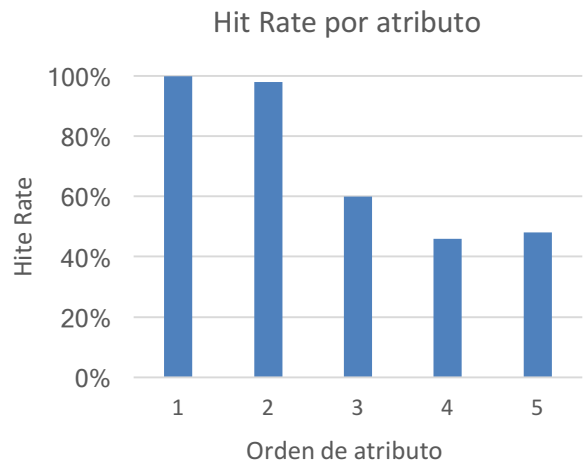


Gráfico 16: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas y 5 alternativas por pregunta.

Se aprecia que a medida que aumenta el número de alternativas, decrece el porcentaje de aciertos de la estimación de la elección. Ello puede explicarse porque al existir más perfiles de producto, la elección se distribuye en un rango mayor de posibilidades y, por ende, existe un mayor margen de error posible. Por otro lado, al tener mayor cantidad de alternativas mejora el hit rate por atributo, explicado por una mayor cantidad de instancias en que se presentan los distintos niveles de atributos con los que se puede construir un orden de importancia más robusto.

3.3.5 Estimación variando el coeficiente lexicográfico w

Finalmente, se realiza una comparación en la estimación de preferencias no compensatorias, variando el valor del coeficiente lexicográfico w .

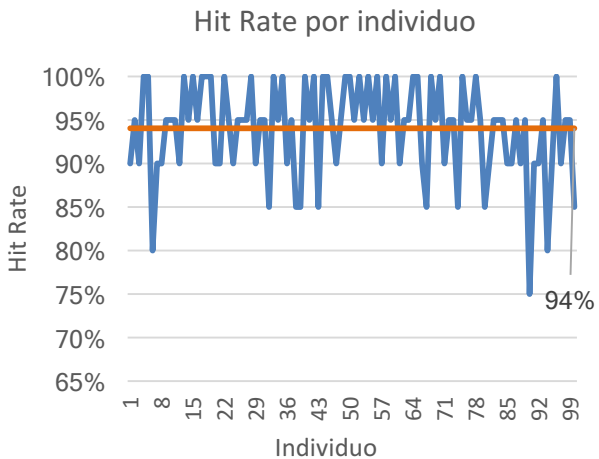


Gráfico 20: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y $w=1/3$.

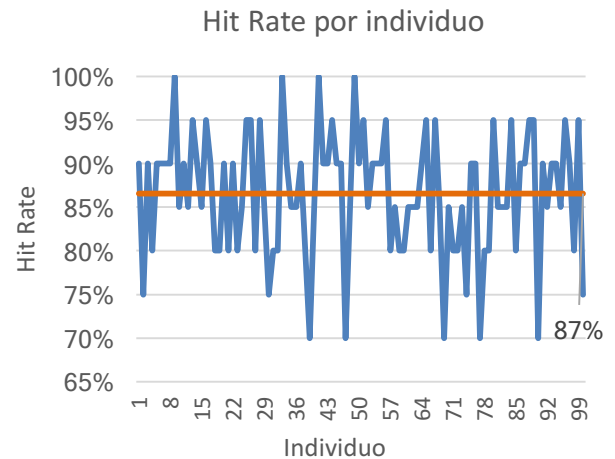


Gráfico 19: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y $w=1/10$.

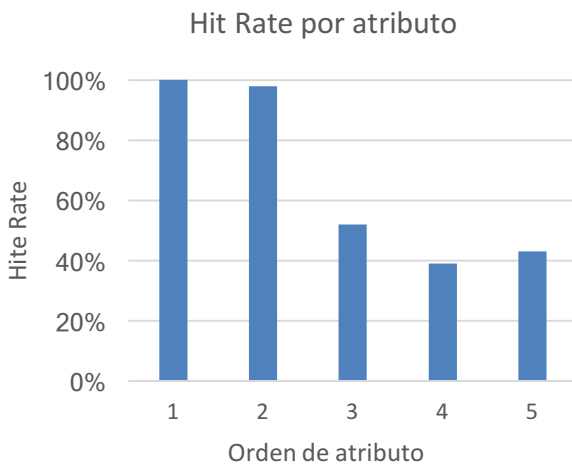


Gráfico 18: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y $w=1/3$.

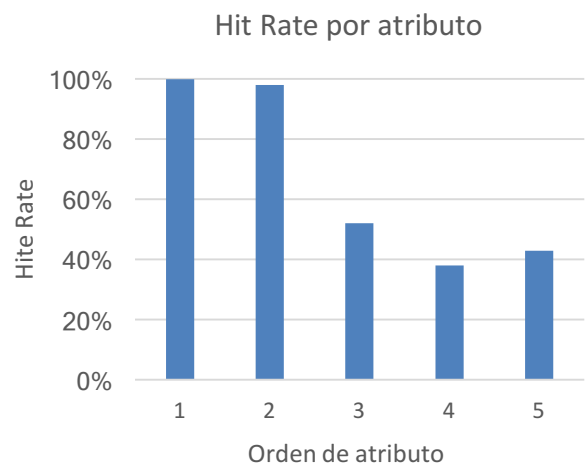


Gráfico 17: Tasa de recuperación del orden de atributos para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y $w=1/10$.

Al aumentar el denominador del coeficiente lexicográfico w , disminuye el hit rate promedio por individuo. Esto se explica dado que los pesos asociados a cada nivel de atributo disminuyen exponencialmente a una mayor tasa, por lo que terminan pareciéndose más entre sí, dificultando la capacidad para estimar correctamente la elección por individuo. Por otro lado, se obtiene el mismo hit rate por atributo, lo que indica que el método de estimación del orden de preferencia es robusto a este tipo de variaciones.

3.4 Conclusiones

A partir de los gráficos presentados, se aprecia que el uso de la heurística Greedy, como método de estimación de preferencias lexicográficas, tiene un desempeño satisfactorio al momento de predecir correctamente las elecciones de cada individuo en un escenario de simulación. En este aspecto, dependiendo de la cantidad de atributos, niveles, preguntas y alternativas consideradas, se logran tasas de acierto superiores al 90%.

Por otro lado, al evaluar su desempeño a través de la tasa de recuperación del verdadero orden de los atributos, se observa que a medida que crece el número de atributos, el porcentaje de predicciones correctas de los atributos de menor importancia decae considerablemente. Así, cuando se trata de predecir un orden lexicográfico, el método de estimación propuesto resulta más efectivo prediciendo el primer y segundo atributo de mayor importancia, que, en definitiva, constituyen aquellos atributos determinantes de la elección según la construcción matemática que define a la utilidad lexicográfica planteada.

Al mismo tiempo, este ratio se relaciona con la cantidad de niveles de cada atributo, de manera que, a mayor cantidad de niveles, el porcentaje de predicción es menor. Pese a que el número de atributos se mantiene constante en el análisis y, para efectos de la simulación, los niveles de un mismo atributo poseen un orden determinístico, estas variaciones se explican debido a que los pesos que condicionan la utilidad se distribuyen en los atributos y niveles de mayor importancia.

Finalmente, cuando se tiene una mayor cantidad de preguntas, el porcentaje de aciertos mejora gracias a que aumentan las instancias en las que un individuo manifiesta sus preferencias, otorgando un mejor espacio de entrenamiento para la predicción.

Los resultados discutidos anteriormente indican que la metodología empleada constituye un buen mecanismo para predecir elecciones derivadas de la regla lexicográfica en un entorno de simulación. No obstante, no se puede asegurar lo mismo en un escenario real, donde no es posible saber con antelación la regla empleada por cada individuo, que como se discutió, puede ser de carácter compensatorio o no compensatorio, dependiendo del contexto de la elección.

En los capítulos futuros se presentan alternativas de estimación de preferencias compensatorias, como una forma de mostrar el desempeño de metodologías tradicionales. A la vez, se buscará contrastar el grado de acierto de ambos enfoques cuando se utilizan indistintamente para la estimación de preferencias tanto compensatorias, como no compensatorias.

Capítulo 4

Estimación de preferencias compensatorias

4.1 Representación de modelos de preferencia compensatorios

Al momento de describir los procesos de elección y toma de decisiones, suelen emplearse modelos de preferencia compensatorios, los cuales asumen que una persona es capaz de evaluar todas las características de un producto, “compensando” la ausencia de un nivel deseado en un atributo por la presencia de otros niveles en el resto de atributos. De este modo, la utilidad de los peores atributos de una alternativa puede equilibrarse con la valoración en otros, haciendo que el perfil de producto sea atractivo en su conjunto.

Utilizando la descripción propuesta por Green y Rao (1971), un individuo evalúa alternativas maximizando la utilidad que le reporta cada perfil de producto, determinada por la suma ponderada de la utilidad de los atributos que lo conforman. De este modo, la utilidad que percibe el consumidor i ante un perfil de producto j puede especificarse a través de la expresión:

$$U_{ij} = u_{ij} + e_{ij}$$

$$U_{ij} = \beta_0 + \beta_{11} * X_{11} + \beta_{12} * X_{12} + \dots + \beta_{1n_1} * X_{1n_1} + \dots + \beta_{mn_m} * X_{mn_m} + e_{ij}$$

En que β_0 corresponde al intercepto o utilidad base, mientras que β_{kn_k} equivale al coeficiente asociado a la utilidad del nivel n_k del atributo k . A su vez, X_{kn_k} corresponde a una variable dummy que indica si el nivel n_k está presente en el perfil de producto j . Finalmente e_{ij} se asocia al error de la alternativa j .

4.2 Simulación de preferencias compensatorias

Con el objetivo de proveer un escenario de comparación a la estimación de preferencias no compensatorias, se simulan escenarios de elección de preferencias compensatorias que serán utilizados para analizar el desempeño de los métodos de estimación propuestos.

Siguiendo el procedimiento de simulación formulado por Arora & Huber (2001) y Toubia et al (2007b), la utilidad determinística de cada perfil de producto puede simularse en base a una distribución normal, en que las preferencias del set de niveles de cada atributo k siguen una media representada por un vector $\bar{\beta}$.

Para efectos de la simulación se trabajará con dos escenarios, que consideran una condición de ruido alta y otra baja para la elección de los consumidores. Al mismo tiempo, para incorporar el grado de heterogeneidad, se utilizará una matriz de covarianza $\Sigma = I$, con I igual a la matriz identidad. De este modo, en el caso de un diseño con 4 atributos y 4 niveles cada uno, el vector de medias $\bar{\beta}$ de la distribución normal para cada individuo queda representado por $(-\beta, -\frac{\beta}{3}, \frac{\beta}{3}, \beta)$. Así, para modelar los escenarios de alto y bajo ruido, el parámetro β puede adquirir valores 1 y 3 respectivamente.

Definidos los parámetros de la simulación, al igual que la simulación de preferencias no compensatorias, los pasos requeridos para su aplicación contemplan la generación de un diseño ortogonal que involucra un determinado número de preguntas o instancias de elección, en las que se presentan alternativas de productos con un número definido de atributos y niveles de atributo. Dichos sets de consideración son asignados a una cantidad específica de individuos, en la que cada uno posee preferencias derivadas de la distribución normal multivariada, con media $\bar{\beta}$ y varianza $\sigma_{\bar{\beta}}^2$, en que los términos de la covarianza son iguales a cero. Con ello, es posible construir, para cada individuo, el verdadero vector de preferencias sobre el total de niveles de atributo.

Definida la utilidad determinística de cada individuo, el término de error se incorpora en la probabilidad de elección mediante una distribución multinomial. De esta forma, un individuo escogerá la alternativa que reporte mayor beneficio, según las distribuciones que caracterizan a los coeficientes de la función de utilidad.

Cabe mencionar que simular preferencias distintas para cada individuo es otra forma de incorporar heterogeneidad a la muestra, proporcionando un escenario de simulación más realista.

4.3 Metodología de estimación de preferencias compensatorias

En un escenario compensatorio, en base a la utilidad que cada consumidor asigna a un perfil de producto dentro de un set de alternativas, es posible definir una función de elección probabilística para cada alternativa x del set de consideración X :

$$\pi_i(x) = \pi(u_i(x) | x \in X_i)$$

Como se explicó anteriormente en el capítulo 2, por definición, $0 \leq \pi_i(x) \leq 1$, para todo $x \in X_i$ y $\sum_{x \in X_i} \pi_i(x) = 1$, para todo $i = 1, \dots, I$.

Luego, la probabilidad de elección promedio asociada al perfil de producto x puede definirse como:

$$\bar{\pi}(x) = \frac{1}{I} \sum_{i=1}^I \pi_i(x)$$

Al aumentar la cantidad de individuos I , el valor de $\bar{\pi}(x)$ se aproxima al valor de la probabilidad esperada de compra. En consecuencia, cuando se tiene una muestra fija de consumidores, optimizar la participación de mercado esperada del perfil de producto x equivale a maximizar las probabilidades individuales de elección $\pi_i(x)$.

De esta forma, al analizar datos empleando un modelo logit multinomial, el modelo de optimización de las preferencias considera las siguientes probabilidades de elección individuales:

$$\pi_i = \frac{\exp(u_i(x))}{\sum_{x \in X_i} \exp(u_i(x))}$$

Para efectos de este trabajo de título, a través del modelo logit multinomial heterogéneo se estiman los coeficientes asociados a la función de utilidad U_{ij} empleando el principio de máxima verosimilitud. De este modo, la función de verosimilitud de la muestra se asocia a las probabilidades de elección individuales:

$$L = \prod_{ij} \pi_{ij}^{y_{ij}}$$

$$LL = \sum_{ij} y_{ij} * \ln (\pi_{ij})$$

En que y_{ij} constituye una variable discreta que indica si el individuo i escoge la alternativa j perteneciente al set de consideración X .

La estimación de estos coeficientes permite cuantificar la importancia relativa que tiene cada nivel de atributo para un individuo determinado. Con ello, como se detallará en los próximos capítulos, se puede utilizar programación dinámica para encontrar el perfil de producto óptimo.

4.4 Inferencia de modelos de preferencia compensatorios sobre datos simulados.

A continuación, se presentan distintos escenarios de simulación de preferencias compensatorias, en que se utiliza un modelo logit heterogéneo para estimar los coeficientes o part worths asociados a cada nivel de atributo. Cabe mencionar que todos los escenarios consideran una muestra de 100 individuos.

4.4.1 Estimación variando el número de atributos

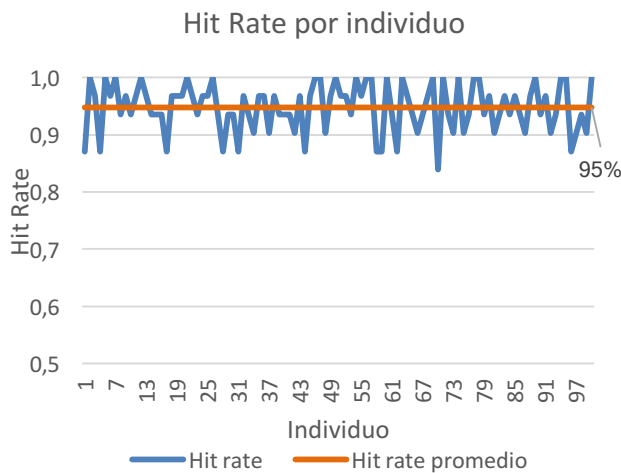


Gráfico 21: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 32 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta = 3$).

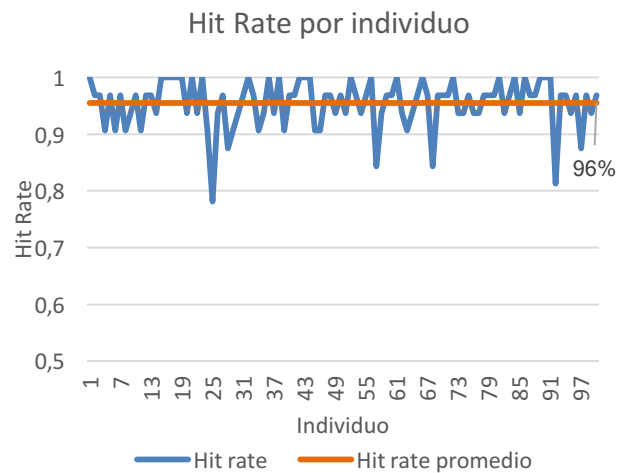


Gráfico 22: Hit rate por individuo para simulación con 8 atributos, 4 niveles por atributo, 32 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta = 3$).

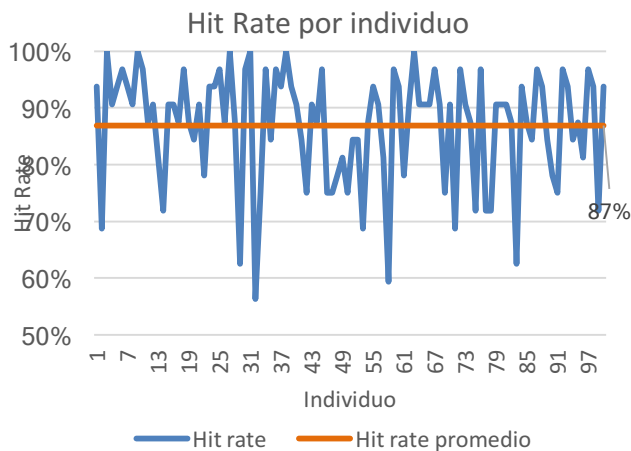


Gráfico 23: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 32 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta = 1$).

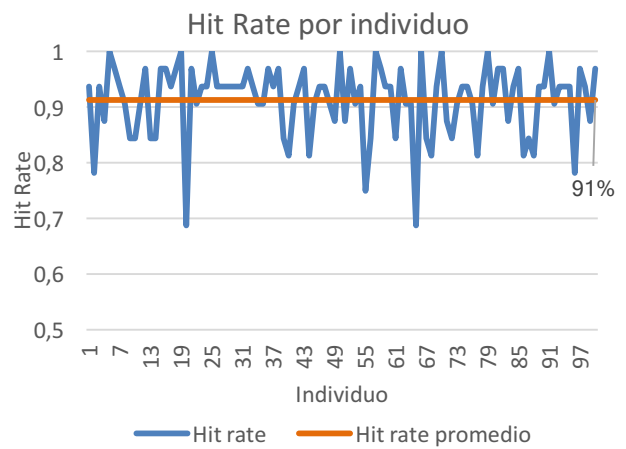


Gráfico 24: Hit rate por individuo para simulación con 8 atributos, 4 niveles por atributo, 32 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta = 1$).

A igual condición de ruido se aprecia que la tasa de acierto crece con un mayor número de atributos. Esta mejora en el ajuste se explica por un aumento en la cantidad de variables explicativas que se incorporan al modelo logit. Al aumentar el ruido, la diferencia entre las muestras con distinta cantidad de atributos se agudiza.

4.4.2 Estimación variando el número de niveles

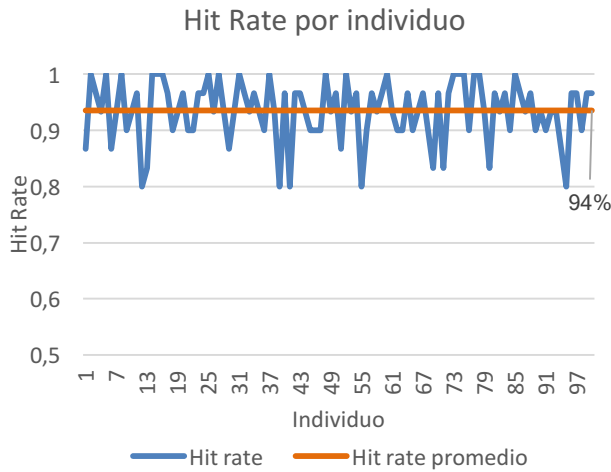


Gráfico 25: Hit rate por individuo para simulación con 5 atributos, 3 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta=3$).

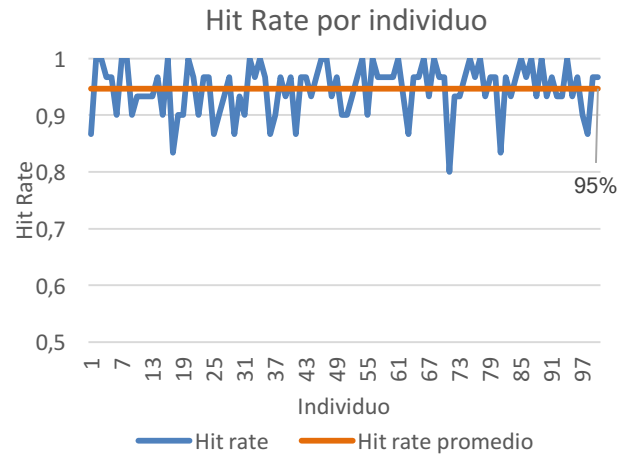


Gráfico 26: Hit rate por individuo para simulación con 5 atributos, 6 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta=3$).

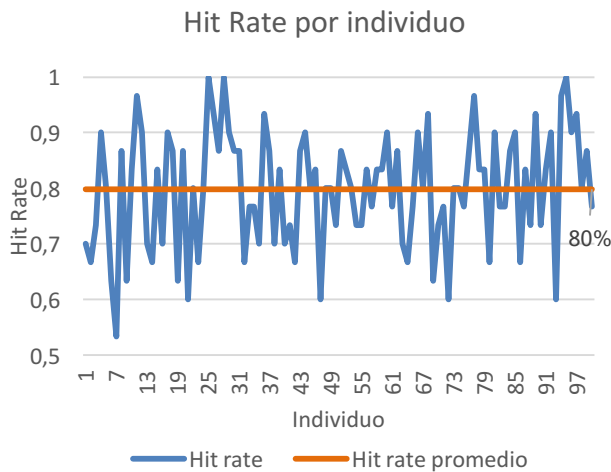


Gráfico 28: Hit rate por individuo para simulación con 5 atributos, 3 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta=1$).

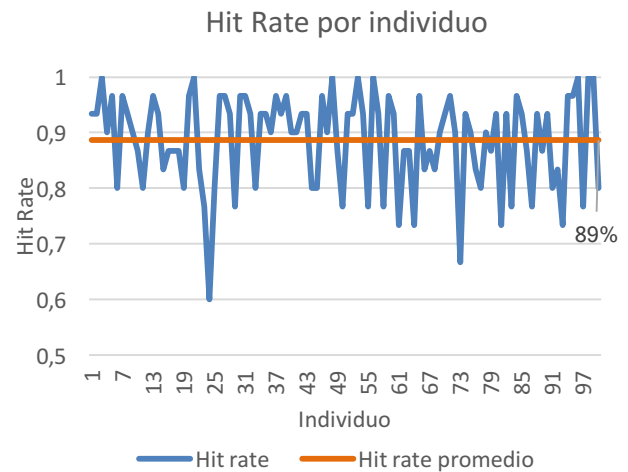


Gráfico 27: Hit rate por individuo para simulación con 5 atributos, 6 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta=1$).

En la condición de bajo ruido ($\beta=3$), al variar la cantidad de niveles de atributos, no se aprecian grandes diferencias en el hit rate promedio por individuo. De este modo, tanto en presencia de 3, como 6 niveles por atributo, se obtienen tasas de acierto similares. Por otro lado, en una condición de alto ruido ($\beta=1$), al considerar una mayor cantidad de niveles por atributo, se obtiene un mejor ajuste de las preferencias simuladas.

4.4.3 Estimación variando el número de preguntas

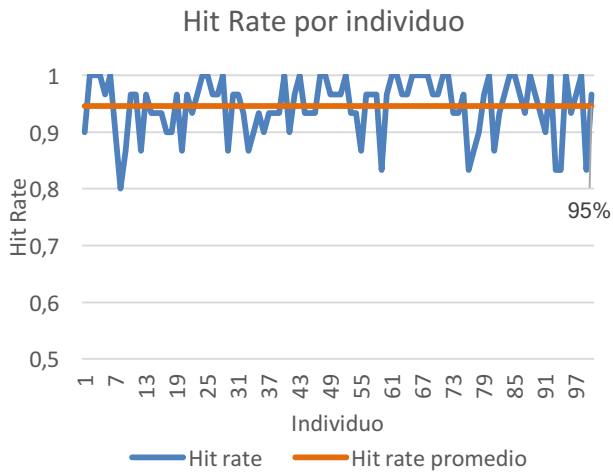


Gráfico 31: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta=3$).

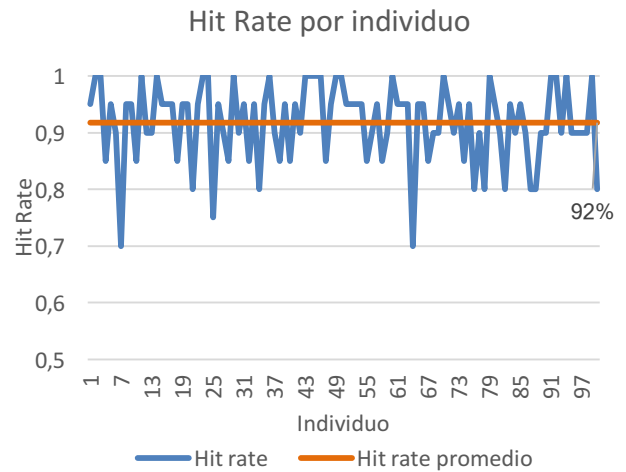


Gráfico 30: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y condición de bajo ruido ($\beta=3$).

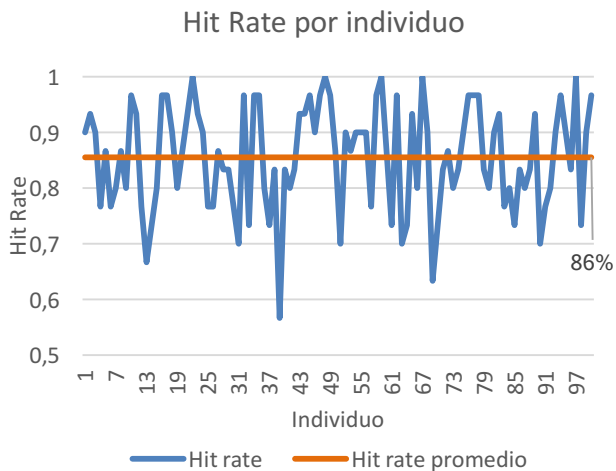


Gráfico 32: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 30 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta=1$).

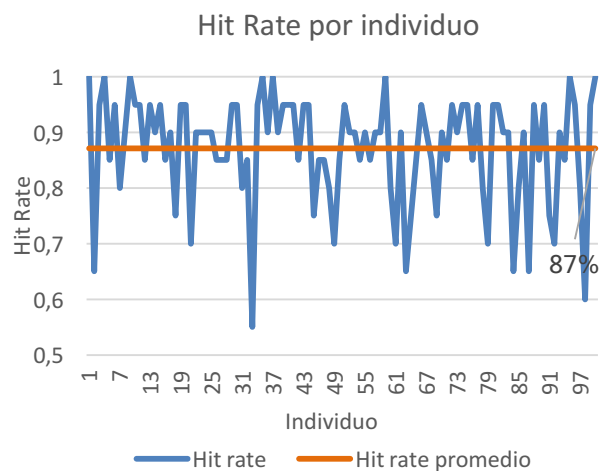


Gráfico 29: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 4 alternativas por pregunta y condición de alto ruido ($\beta=1$).

A partir de lo anterior se observa que, en una condición de bajo ruido, aumentar la cantidad de preguntas o sets de elección a los que se enfrenta un individuo mejora el porcentaje de aciertos respecto a las preferencias simuladas. No obstante, en una condición de alto ruido, no se distinguen diferencias significativas entre ambos escenarios.

4.4.4 Estimación variando el número de alternativas

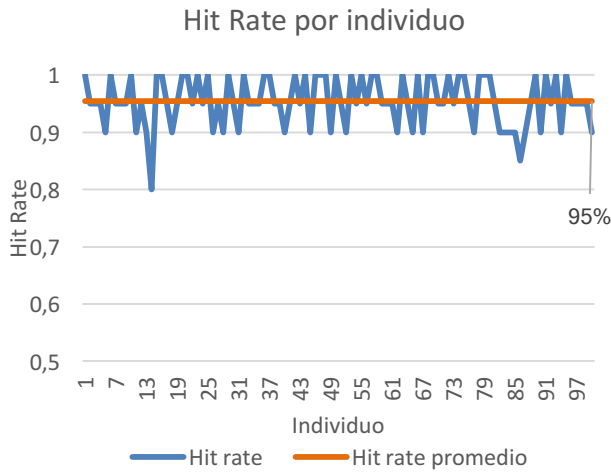


Gráfico 34: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 3 alternativas por pregunta y condición de bajo ruido ($\beta=3$).

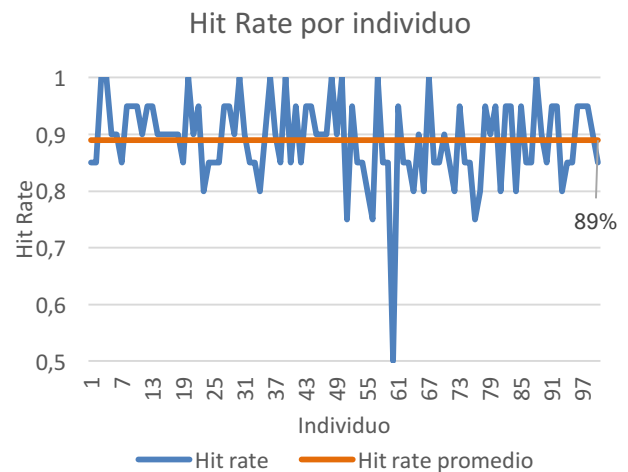


Gráfico 35: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 5 alternativas por pregunta y condición de bajo ruido ($\beta=3$).

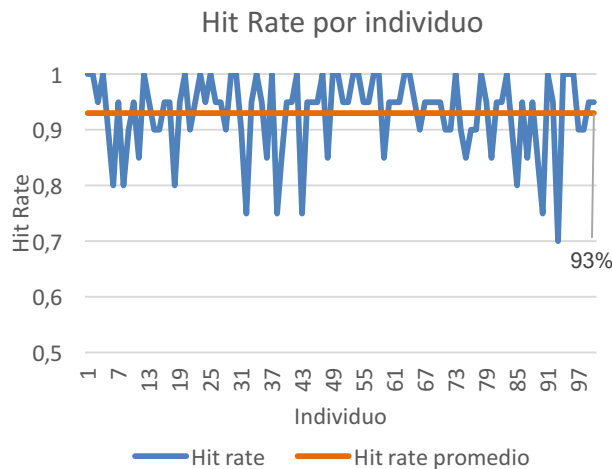


Gráfico 36: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 3 alternativas por pregunta y condición de alto ruido ($\beta=1$).

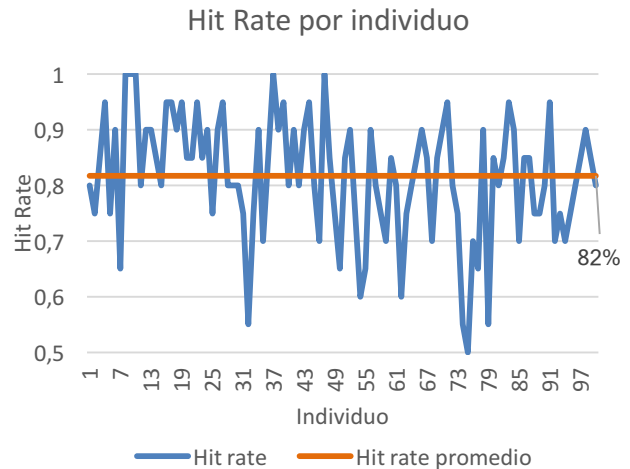


Gráfico 33: Hit rate por individuo para simulación con 5 atributos, 4 niveles por atributo, 20 preguntas, 5 alternativas por pregunta y condición de alto ruido ($\beta=1$).

Se observa que a medida que aumenta el número de alternativas, disminuye el porcentaje de aciertos de la estimación de la elección. Ello se explica porque al existir más perfiles de producto, la elección se distribuye en un rango mayor de posibilidades y, por ende, existe un mayor margen de error posible. Al mismo tiempo, al comparar entre ambas condiciones de ruido, se aprecia que esta diferencia se acentúa a menor β .

4.5 Conclusiones

El uso de un modelo de preferencias compensatorio constituye una metodología robusta para estimar la elección de los individuos, siendo ampliamente conocida y comúnmente utilizada por su simplicidad.

De esta forma, en un ambiente de simulación de preferencias compensatorias, se considera al modelo logit heterogéneo como una buena alternativa de comparación y estimación, dado que cuenta con una forma cerrada para la evaluación de la log-verosimilitud, siendo apropiado en un escenario en el que no se consideran variables sociodemográficas de los individuos y en el que interesa medir la capacidad de la estimación a partir de datos obtenidos de una encuesta de análisis conjunto.

De los resultados presentados, se infiere que el uso de un modelo compensatorio se ve afectado por variables propias del diseño experimental, como lo es la cantidad de atributos, el número de niveles, la proporción de preguntas y el tamaño de cada set de consideración. Es así, como una mayor cantidad de atributos y niveles genera un impacto positivo en el ajuste de la predicción, debido a que se adhieren más variables explicativas al modelo. Al mismo tiempo, una mayor cantidad de alternativas genera un impacto negativo en el porcentaje de aciertos de la elección, explicado por un mayor rango de opciones que se traduce en un mayor margen de error posible.

A partir de lo observado, se puede concluir que el grado de incidencia que tiene cada factor en la estimación será mayor en condiciones de alto ruido, a excepción de la cantidad de preguntas que, bajo este escenario, no influyen considerablemente en la predicción.

En definitiva, el modelo de simulación empleado refleja satisfactoriamente las decisiones de los consumidores, considerando situaciones con un alto y bajo índice de ruido, que varían la forma en que las variables del diseño experimental afectan la predicción. En este aspecto, el modelo logit heterogéneo tiene un buen desempeño prediciendo las decisiones de cada individuo, con tasas de acierto superiores al 80%. Sin embargo, al tratarse de un modelo a nivel individual, que utiliza las elecciones propias de cada persona para la estimación de sus preferencias, el ajuste de los resultados se encontraría limitado a cada muestra en particular. No obstante, se considera un método válido para contrastarlo con los resultados obtenidos de la estimación con un modelo lexicográfico, ya que ambos estiman preferencias a nivel individual y parte de los objetivos específicos de este trabajo es la comparación entre ambos modelos.

Capítulo 5

Diseño de producto óptimo bajo preferencias simuladas

El proceso de diseño de un producto involucra la definición de sus características, funcionalidad y propósito, sujeto a restricciones de tamaño, peso, forma, duración y apariencia. Para una compañía, su importancia radica en el grado de diferenciación que puede obtenerse respecto a otros productos de la competencia, en orden de atraer una mayor cantidad de consumidores. En este aspecto, el diseño se convierte en la manera de organizar los atributos y beneficios de un producto antes de su introducción al mercado y presentación al cliente.

El diseño de nuevos productos, así como la innovación y extensión de líneas de productos existentes, son un recurso fundamental para toda compañía que busca consolidar su permanencia y sostenibilidad en el tiempo. Esta relevancia se aprecia fundamentalmente en mercados en los que prima la tecnología, como lo es la venta de computadores y smartphones, no obstante, incluso en el área de servicios, el diseño de productos puede jugar un rol decisivo en la eficiencia y valor entregado al consumidor.

Al momento de introducir un nuevo producto al mercado, es fundamental definir y entender las preferencias de los consumidores. De este modo, contar con un diseño apropiado, que cumple las expectativas de los individuos, llevará a que el producto sea escogido por sobre el resto de alternativas existentes.

Debido a la creciente variedad de bienes y servicios disponible, el diseño óptimo de productos y surtido de productos se han convertido en un aspecto aún más importante para compañías desarrolladoras y retailers que buscan la construcción de ventajas competitivas a través del entendimiento del comportamiento de sus clientes.

La literatura actual en materia de diseño óptimo de productos contempla preferencias compensatorias, asumiendo que un consumidor considera todas las alternativas y atributos disponibles al momento de tomar la decisión de compra. No obstante, si los consumidores escogen de manera no compensatoria, el diseño y surtido de productos basados en preferencias compensatorias puede ser menos efectivo.

De este modo, si los consumidores escogen de manera compensatoria, una compañía debería concentrar sus esfuerzos en ofrecer productos que sean buenos en muchos atributos. En caso contrario, si los consumidores optan de manera no compensatoria convendría diseñar productos buenos sólo en los atributos más relevantes.

A continuación, se describen alternativas para evaluar el diseño óptimo de productos bajo reglas de decisión tanto compensatorias, como no compensatorias, realizando comparaciones entre ambos enfoques que permiten explorar si dicha distinción tiene un efecto significativo en el perfil de producto óptimo obtenido.

5.1 Metodologías de diseño de producto óptimo

Una forma simple de estimar el desempeño potencial de un nuevo producto es a través de part-worths derivados de las preferencias obtenidas en encuestas del estilo análisis conjunto. Cuando este objetivo se orienta a resolver un problema de participación de mercado, vale decir, maximizar el número de consumidores que comprarán el perfil de producto a evaluar, surgen distintas alternativas de cálculo.

La primera de ellas corresponde a una búsqueda exhaustiva, en que se evalúan todas las combinaciones posibles de atributos y niveles de atributo para escoger el perfil de producto con mayor aceptación. Sin embargo, este problema resulta difícil de resolver, debido a que a medida que aumenta el número de características, se requiere más tiempo de procesamiento.

Es así como, según lo descrito por Kohli y Krishnamurti (1989), este problema resulta de carácter NP-Complejo, es decir, relativamente difícil de abordar dentro de la familia de problemas que pueden ser resueltos por un computador. En otras palabras, cuando se tiene un problema de participación de mercado de perfiles de producto con m atributos y n niveles cada uno, seleccionar un perfil de producto óptimo requiere evaluar n^m perfiles de producto distintos, un desafío que crece exponencialmente con el valor de m .

Bajo este escenario, se recurre al uso de heurísticas que encuentran un perfil óptimo, ya sea global o local, que disminuyen los tiempos de procesamiento. Las metodologías más conocidas corresponden a la heurística de Programación Dinámica y la heurística Greedy. Ambas técnicas trabajan de manera similar empleando part-worths que permiten determinar el nivel óptimo de cada atributo para una muestra de individuos.

Al utilizar una heurística Greedy se calcula, para cada atributo k , las utilidades relativas promedio de todos individuos i según la expresión:

$$u_{kn} = \sum_{i=1}^I u_{ikn} - u_{ik\tilde{n}_k}$$

En que \tilde{n}_k corresponde al nivel del k -ésimo atributo para el perfil de producto base. De este modo, la heurística greedy selecciona el nivel n_k^* del atributo k que maximiza las utilidades relativas. En consecuencia, el nivel de cada atributo se escoge de manera independiente al resto de atributos.

Por otro lado, la heurística de Programación Dinámica trata cada atributo como la etapa de un proceso en el que los niveles simbolizan estados de cada fase. En cada iteración del algoritmo se define el mejor nivel para el atributo evaluado, sin considerar todas las interacciones entre atributos. De esta forma, en la primera iteración se determina el mejor nivel del primer atributo para cada uno de los niveles del segundo atributo, generando n_2 perfiles parciales de producto. Luego, en el siguiente paso, se generan n_3 perfiles parciales de producto (uno por cada nivel del atributo 3), encontrando la mejor combinación de niveles para el primer y segundo atributo de la etapa anterior. El proceso termina cuando se han evaluado todos los atributos, teniendo un conjunto de n_m perfiles parciales de producto a partir del cual se escoge el mejor.

Ambas heurísticas resultan similares en estructura, sin embargo se diferencian en que la heurística greedy realiza elecciones de manera iterativa, reduciendo el conjunto de opciones en uno más pequeño sin reconsiderar las decisiones efectuadas anteriormente, por lo cual no garantiza óptimos globales. Por otro lado, la programación dinámica es más exhaustiva, ya que en cada etapa toma en consideración las elecciones realizadas en la etapa anterior (Cormen, Leiserson, Rivest & Stein, 2009).

Dado lo anterior, el algoritmo Greedy puede resultar menos eficiente que la Programación Dinámica, siendo menos confiable en términos de la solución otorgada. Por esta razón, este trabajo de título emplea como principal heurística de optimización a esta última, cuyo mecanismo de cálculo se explica a continuación.

5.2 Diseño de producto óptimo con programación dinámica

El método de programación dinámica puede extenderse para el diseño óptimo de productos, evaluando la combinación de niveles de atributos que tiene un mayor grado de aceptación a partir de las preferencias manifestadas en una encuesta de análisis conjunto.

A continuación, se describe el algoritmo empleado, siguiendo la estructura propuesta por Kohli y Krishnamurti (1987):

Sea $C(k)$ una matriz que denota las preferencias individuales para cada nivel del atributo k , con $k \in \Omega = \{1, \dots, m\}$.

Sea $\Phi_k = \{1, \dots, J_k\}$ el set de niveles del atributo k .

De este modo, cada fila de $C(k)$ corresponde a un determinado individuo i , mientras que cada columna j equivale a un nivel del atributo k . Con ello, el elemento ij de $C(k)$ representa el coeficiente de preferencias relativo que el individuo i asocia al nivel j del atributo k . Sea $C_j(k)$ la j -ésima columna de $C(k)$.

Iteración inicial:

A cada columna de $C(1)$ se adiciona la columna $C_j(2)$, $j \in \Phi_2$ según la expresión que se describe a continuación:

$$S_j(2) = C(1) + [C_j(2)][1] \quad \forall j \in \Phi_2$$

En este caso, $[1]$ denota un vector fila con elementos equivalentes a la unidad. De este modo, se forman tantas matrices $S_j(2)$ como niveles compongan al atributo 2. Para cada individuo i y para cada nivel j del atributo 2, se escoge la columna de $S_j(2)$ que reporta mayor utilidad y se le asigna el valor 1, al resto de columnas se asigna el valor 0.

Por ejemplo, si se tienen dos individuos ($i = 1, 2$), en que el primer atributo cuenta con dos niveles ($j = 1, 2$) y el segundo atributo con J_2 niveles, se tiene que:

$$C_1(1) = (c_{11}(1) \ c_{21}(1))^t$$

$$C_2(1) = (c_{12}(1) \ c_{22}(1))^t$$

$$C_j(2) = (c_{1j}(2) \ c_{2j}(2))^t \ \forall j \in J_2$$

$$S_j(2) = \begin{bmatrix} c_{11}(1) & c_{12}(1) \\ c_{21}(1) & c_{22}(1) \end{bmatrix} + \begin{bmatrix} c_{1j}(2) \\ c_{2j}(2) \end{bmatrix} [1 \ 1]$$

$$S_j(2) = \begin{bmatrix} c_{11}(1) + c_{1j}(2) & c_{12}(1) + c_{1j}(2) \\ c_{21}(1) + c_{2j}(2) & c_{22}(1) + c_{2j}(2) \end{bmatrix}$$

Posteriormente, para cada nivel j del atributo 2 se selecciona la columna $s_j^*(2)$ de $S_j(2)$ cuya cantidad de valores iguales a 1 es mayor. En caso de existir empate, se escoge una columna de manera aleatoria.

Una vez seleccionadas las columnas $s_j^*(2)$ de todos los niveles Φ_k del atributo 2, se forma la matriz:

$$[S^*(2) = [s_1^*(2) \dots s_{j_2}^*(2)]]$$

Con ello, para cada individuo i , el valor de $s_{ij_1}^*(2)$ representa la suma de los coeficientes de preferencia relativos para el nivel j del atributo 2 y el nivel j_1 del atributo 1, en que j_1 es escogido de manera tal de maximizar el número de individuos para quienes la suma $c_{ij_1}(1) + c_{ij}(2)$ reporta mayor utilidad.

Recursión:

A cada columna de $S^*(k-1)$ se adiciona la columna $C_j(k)$, $j \in \Phi_2$ según la expresión que se describe a continuación:

$$S_j(k) = S^*(k-1) + [C_j(k)][1] \quad \forall j \in \Phi_k$$

En que [1] denota un vector fila con elementos equivalentes a la unidad. De este modo, se forman tantas matrices $S_j(k)$ como niveles compongan al atributo k . Para cada individuo i y para cada nivel j del atributo k , se escoge la columna de $S_j(k)$ que reporta mayor utilidad y se le asigna el valor 1, al resto de columnas se asigna el valor 0.

Posteriormente, para cada nivel j del atributo k se selecciona la columna $s_j^*(k)$ de $S_j(k)$ cuya cantidad de valores iguales a 1 es mayor. En caso de existir empate, se escoge una columna de manera aleatoria.

Una vez seleccionadas las columnas $s_j^*(k)$ de todos los niveles Φ_k del atributo k , se forma la matriz:

$$S^*(k) = [s_1^*(k) \dots s_{j_k}^*(k)]$$

Con ello, para cada individuo i , el valor de $s_{ij}^*(k)$ representa la suma de los coeficientes de preferencia relativos para el nivel j del atributo k y un nivel de cada uno de los atributos $1, \dots, (k-1)$. Dado un nivel j del atributo k , el nivel j_{k-1} del atributo $(k-1)$ es escogido de manera tal de maximizar el número de individuos para quienes la suma $s_{ij_{k-1}}(k-1) + c_{ij}(k)$ reporta mayor utilidad.

Iteración final:

Para cada individuo i y para cada nivel j del atributo m , se escoge la columna de $S^*(m)$ que reporta mayor utilidad y se le asigna el valor 1, al resto de columnas se asigna el valor 0.

Posteriormente, para cada nivel j del atributo m se selecciona la columna $s_j^*(m)$ de $S^*(m)$ cuya cantidad de valores iguales a 1 es mayor. En caso de existir empate, se escoge una columna de manera aleatoria.

En definitiva, el algoritmo trabaja creando perfiles de producto óptimo parciales, seleccionando el mejor nivel para el atributo evaluado en cada iteración, teniendo en consideración la combinación de niveles de atributo escogidos en las iteraciones previas.

5.3 Diseño de producto óptimo con búsqueda exhaustiva

Para tener una alternativa de comparación a la heurística de programación dinámica, se emplea un segundo enfoque correspondiente a búsqueda exhaustiva. Como se explicó, la dificultad de este método se asocia a los tiempos de ejecución, los cuales crecen exponencialmente con el número de atributos. Sin embargo, en un escenario de simulación en que el diseño experimental no contempla demasiados atributos y niveles de atributo, es posible evaluar la participación de mercado y utilidad que genera cada perfil de producto.

5.4 Comparación de producto óptimo entre modelos de preferencia compensatorios y no compensatorios

Para dar cumplimiento al objetivo principal de este trabajo de título, se utilizan los enfoques de optimización presentados a sets de datos simulados con preferencias tanto compensatorias, como no compensatorias.

Como primer punto de comparación se utiliza un diseño experimental dado por los siguientes factores:

- Número de atributos: 5
- Número de niveles de atributo: 4
- Cantidad de perfiles de producto en cada set de consideración: 4
- Condición de ruido para el modelo compensatorio: $\beta = 1$ y $\beta = 3$

Se escogió esta combinación de factores dado que constituye una configuración estándar frecuentemente usada en aplicaciones de análisis conjunto.

En ambos enfoques, para evaluar un potencial perfil de producto óptimo, éste se agrega como una alternativa adicional a los sets de elección y se obtiene un "índice de participación" basado en un criterio de máxima utilidad. Para ello, se calcula el porcentaje de veces en que se asigna la máxima utilidad al perfil del producto óptimo respecto al total de sets de elección presentados a los individuos. Por ejemplo, si se trata de 100 individuos y 20 preguntas para cada uno, se tienen 2.000 instancias de elección sobre las que se evalúa el desempeño de un perfil específico. Este índice es el similar a la cuota de mercado, correspondiente al número de individuos que asigna la máxima utilidad al producto óptimo, respecto al total de individuos.

Además, resulta importante mencionar que en este trabajo se tienen dos tipos de utilidades. En primer lugar, la utilidad que percibe un consumidor al momento de escoger un perfil de producto, basado en las preferencias por cada nivel de atributo. Esta utilidad se asume lineal, tanto para el caso compensatorio, como no compensatorio.

En segundo lugar, se encuentra la utilidad que recibe la compañía que diseña el producto. Para ello se definen precios y costos por nivel de atributo. Los costos están estructurados de tal forma que los peores niveles de cada atributo tienen asociados costos más bajos que los mejores niveles. Para incorporar el beneficio (precio) de cada atributo, se definió éste como el costo del nivel más caro. De esta forma, se asume que pasar de un nivel a otro cuesta lo mismo para todos los atributos. Por ende, independiente del atributo, el aumento del costo entre niveles es proporcional.

A continuación, se detalla la estructura de costos para el escenario de simulación planteado:

<i>Nivel Atributo</i>	Costo (\$USD)	Precio (\$USD)
1	20	80
2	40	80
3	60	80
4	80	80

Tabla 2: Estructura de costos para diseño de producto óptimo de datos simulados.

Si bien esta proporcionalidad entre costos de atributos es un gran supuesto, es una forma de modelar el hecho de que para la firma existe un trade-off entre ofrecer productos más atractivos, que tienen mejor aceptación, pero que son más costosos.

Este supuesto busca una representación más realista de un problema en el que no siempre se tienen los costos e ingresos derivados de la producción y comercialización de cada perfil de producto, pero que deben tenerse en cuenta como criterios de rentabilidad que son fundamentales para la firma.

A continuación, se presentan los resultados en términos de utilidad e índice de participación para el producto óptimo de cada modelo. Al mismo tiempo, se muestra el hit rate correspondiente a la estimación de preferencias. Cabe mencionar que las métricas están calculadas en base a las preferencias simuladas y no a las estimadas, aunque la obtención del producto óptimo se realiza utilizando los coeficientes de preferencia obtenidos con el método de estimación de preferencias respectivo.

5.4.1 Datos simulados de manera compensatoria

5.4.1.1 Simulación utilizando coeficiente de alto ruido ($\beta = 1$)

En primer lugar, se crea un set de datos simulados de manera compensatoria, los cuales se utilizarán para la obtención del producto óptimo asumiendo preferencias tanto compensatorias, como no compensatorias en la estimación de los coeficientes por individuo.

Los resultados del producto óptimo obtenido con programación dinámica en base a estimación asumiendo preferencias compensatorias son los siguientes:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	87%	Atributo 1-Nivel 1 Atributo 2-Nivel 1 Atributo 3-Nivel 3 Atributo 4-Nivel 1 Atributo 5-Nivel 1	0,25	\$6.487

Tabla 3: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=1$).

Se aprecia que la mayoría de los niveles de cada atributo corresponden a aquellos menos atractivos individualmente según la distribución de costos utilizada. Al mismo tiempo se obtiene un hit rate del 87% y un índice de participación igual al 25%.

Por otro lado, al utilizar el método de búsqueda exhaustiva asumiendo preferencias compensatorias se obtiene lo siguiente:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	87%	Atributo 1-Nivel 1 Atributo 2-Nivel 1 Atributo 3-Nivel 1 Atributo 4-Nivel 1 Atributo 5-Nivel 1	0,24	\$7.335

Tabla 4: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=1$).

Si bien se obtiene un índice de participación cercano al adquirido con la heurística de programación dinámica, existen diferencias en la utilidad y en uno de los niveles del producto óptimo obtenido.

Al mismo tiempo, se obtiene el producto óptimo mediante la estimación de preferencias no compensatorias para el set de datos simulados con preferencias compensatorias:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	41%	Atributo 1-Nivel 4 Atributo 2-Nivel 1 Atributo 3-Nivel 4 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,19	\$1.173

Tabla 5: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=1$).

Como es de esperar, al tratarse de datos simulados de forma compensatoria se obtiene una menor tasa de acierto respecto a la estimación con un método compensatorio. Ello, deriva en un menor índice de participación y, por ende, en una menor utilidad esperada.

Finalmente, se obtienen los mismos resultados, pero a través de búsqueda exhaustiva:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	41%	Atributo 1-Nivel 1 Atributo 2-Nivel 1 Atributo 3-Nivel 4 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,22	\$2.730

Tabla 6: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=1$).

Al comparar con la heurística de programación dinámica, se obtienen 3 puntos porcentuales más en el índice de participación de mercado, lo que repercute en una mayor utilidad esperada. No obstante, estas métricas siguen siendo menores a las obtenidas a través de la estimación de preferencias con un modelo compensatorio.

5.4.1.2 Simulación utilizando coeficiente de bajo ruido ($\beta = 3$)

Se presentan los resultados del producto óptimo obtenido con programación dinámica en base a estimación asumiendo preferencias compensatorias:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	92%	Atributo 1-Nivel 1 Atributo 2-Nivel 1 Atributo 3-Nivel 2 Atributo 4-Nivel 1 Atributo 5-Nivel 1	0,25	\$7.014

Tabla 7: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=3$).

En este caso, se obtiene un hit rate mayor al obtenido en la condición de alto ruido. Si bien se tiene el mismo índice de participación, la utilidad generada es mayor, dado que la configuración del producto óptimo considera un nivel más barato para el tercer atributo.

Al emplear el método de búsqueda exhaustiva asumiendo preferencias compensatorias se obtiene lo siguiente:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	92%	Atributo 1-Nivel 1 Atributo 2-Nivel 1 Atributo 3-Nivel 1 Atributo 4-Nivel 1 Atributo 5-Nivel 1	0,26	\$7.860

Tabla 8: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=3$).

Nuevamente, el producto óptimo difiere en un nivel de atributo, generando mayor índice de participación y utilidad esperada respecto a la optimización con heurística de programación dinámica.

Como punto de comparación se obtiene el producto óptimo mediante la estimación de preferencias no compensatorias para el set de datos simulados con preferencias compensatorias:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	42%	Atributo 1-Nivel 4 Atributo 2-Nivel 1 Atributo 3-Nivel 4 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,18	\$1.089

Tabla 9: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=3$).

Al igual que en la condición de alto ruido, al tratarse de datos simulados de forma compensatoria se obtiene una menor tasa de acierto respecto a la estimación con un método compensatorio. Ello, deriva en un menor índice de participación y, por ende, en una menor utilidad esperada.

Al efectuar la optimización mediante búsqueda exhaustiva se obtiene lo siguiente:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	42%	Atributo 1-Nivel 4 Atributo 2-Nivel 1 Atributo 3-Nivel 4 Atributo 4-Nivel 1 Atributo 5-Nivel 4	0,19	\$2.370

Tabla 10: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=3$).

Respecto a la optimización con búsqueda exhaustiva, se obtiene 1 punto porcentual más en el índice de participación de mercado, lo que repercute en una mayor utilidad esperada. A la vez, estas métricas siguen siendo menores a las obtenidas a través de la estimación de preferencias con un modelo compensatorio. No obstante, las diferencias entre ambos métodos de optimización son menores respecto a las evidenciadas en la condición de alto ruido.

5.4.2 Datos simulados de manera no compensatoria

Finalmente, se crea un set de datos simulados de manera no compensatoria, los cuales se utilizan para la obtención del producto óptimo asumiendo preferencias tanto compensatorias, como no compensatorias en la estimación de los coeficientes por individuo.

A continuación, se presentan los índices de acierto, configuración de producto óptimo, índice de participación y utilidad esperada dada la estimación de preferencias compensatorias.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	95%	Atributo 1-Nivel 2 Atributo 2-Nivel 4 Atributo 3-Nivel 3 Atributo 4-Nivel 4 Atributo 5-Nivel 1	0,32	\$3.840

Tabla 11: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera no compensatoria.

Para el caso de datos simulados con un orden de preferencias lexicográfico, el uso de un modelo de estimación compensatorio tiene un buen desempeño en términos de hit rate (95%), generando un índice de participación del 32%.

Por otro lado, al utilizar el método de búsqueda exhaustiva asumiendo preferencias compensatorias se obtiene lo siguiente:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	95%	Atributo 1-Nivel 2 Atributo 2-Nivel 4 Atributo 3-Nivel 3 Atributo 4-Nivel 4 Atributo 5-Nivel 1	0,32	\$3.840

Tabla 12: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera no compensatoria.

En este caso, los resultados obtenidos de la optimización con heurística de programación dinámica coinciden con los de búsqueda exhaustiva, evidenciando que la primera es capaz de encontrar el óptimo global.

Al evaluar el producto óptimo mediante la estimación de preferencias no compensatorias para el set de datos simulados con preferencias no compensatorias se obtiene:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	95%	Atributo 1-Nivel 1 Atributo 2-Nivel 4 Atributo 3-Nivel 3 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,52	\$4.144

Tabla 13: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera no compensatoria.

Se observa igual porcentaje de aciertos respecto a la estimación basada en un modelo compensatorio. A la vez, el índice de participación de mercado y la utilidad obtenida son mayores respecto a la estimación basada en un modelo compensatorio.

Finalmente, se obtiene el producto óptimo en base a búsqueda exhaustiva asumiendo preferencias no compensatorias:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	95%	Atributo 1-Nivel 1 Atributo 2-Nivel 4 Atributo 3-Nivel 2 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,46	\$4.698

Tabla 14: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera no compensatoria.

Existen diferencias en el índice de participación y en uno de los niveles del producto óptimo obtenido respecto a la programación dinámica, por lo que la utilidad total esperada también difiere.

5.4.3 Variación de la condición de heterogeneidad

A partir de los datos obtenidos, se espera que el grado de heterogeneidad de la muestra tenga un impacto en el tipo de producto escogido y, por ende, en la participación y utilidad esperada. Dicho esto, se realiza una simulación variando esta condición mediante la creación de segmentos de consumidores que internamente comparten el mismo orden de preferencia por cada atributo. Dichos segmentos se construyen dividiendo aleatoriamente la muestra de individuos en grupos que tienen igual distribución de preferencias por los atributos de un producto.

Por otro lado, a partir de lo evidenciado en los capítulos 3 y 4, las variaciones en la cantidad de atributos y niveles por atributo no tienen un mayor impacto en los índices de hit rate y recuperación del orden de atributos. Al mismo tiempo, si bien las variaciones en la cantidad de preguntas y alternativas por pregunta tienen un impacto en estas métricas, corresponden a variaciones propias del diseño experimental que no explicarían la diferencia entre los productos óptimos obtenidos entre un enfoque compensatorio y uno no compensatorio.

5.4.3.1 Datos simulados de manera compensatoria ($\beta = 1$)

A continuación, se presentan los índices de acierto, configuración de producto óptimo, índice de participación y utilidad esperada dada la estimación de preferencias compensatorias y no compensatorias con 2 y 10 segmentos de consumidores de una muestra simulada de manera compensatoria con un alto coeficiente de ruido:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	87%	Atributo 1-Nivel 3 Atributo 2-Nivel 2 Atributo 3-Nivel 1 Atributo 4-Nivel 2 Atributo 5-Nivel 3	0,83	\$14.931

Tabla 15: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=1$) y 2 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	37%	Atributo 1-Nivel 4 Atributo 2-Nivel 1 Atributo 3-Nivel 4 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,17	\$996

Tabla 16: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=1$) y 2 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	84%	Atributo 1-Nivel 3 Atributo 2-Nivel 1 Atributo 3-Nivel 1 Atributo 4-Nivel 2 Atributo 5-Nivel 1	0,32	\$7.680

Tabla 17: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=1$) y 10 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	42%	Atributo 1-Nivel 1 Atributo 2-Nivel 4 Atributo 3-Nivel 3 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,27	\$2.188

Tabla 18: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=1$) y 10 segmentos de individuos.

En el caso de la estimación con un modelo compensatorio, al aumentar la heterogeneidad de las preferencias de la muestra, disminuye el índice de aciertos. Respecto al producto óptimo obtenido, disminuye el índice de participación de mercado y con ello, la utilidad esperada. Por otro lado, al estimar preferencias compensatorias con un modelo no compensatorio, si aumenta la heterogeneidad de la muestra, aumenta el hit rate, el índice de participación y la utilidad esperada.

Al mismo tiempo, la diferencia entre la utilidad esperada de la estimación con un modelo compensatorio y no compensatorio disminuye a medida que los individuos son más heterogéneos.

5.4.3.2 Datos simulados de manera compensatoria ($\beta = 3$)

A continuación, se presentan los índices de acierto, configuración de producto óptimo, índice de participación y utilidad esperada dada la estimación de preferencias compensatorias y no compensatorias con 2 y 10 segmentos de consumidores de una muestra simulada de manera compensatoria con un bajo coeficiente de ruido:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	92%	Atributo 1-Nivel 1 Atributo 2-Nivel 1 Atributo 3-Nivel 1 Atributo 4-Nivel 3 Atributo 5-Nivel 1	0,61	\$16.003

Tabla 19: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=3$) y 2 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	36%	Atributo 1-Nivel 4 Atributo 2-Nivel 4 Atributo 3-Nivel 1 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,53	\$3.201

Tabla 20: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=3$) y 2 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	90%	Atributo 1-Nivel 1 Atributo 2-Nivel 1 Atributo 3-Nivel 4 Atributo 4-Nivel 1 Atributo 5-Nivel 1	0,37	\$8.820

Tabla 21: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera compensatoria ($\beta=3$) y 10 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	48%	Atributo 1-Nivel 4 Atributo 2-Nivel 1 Atributo 3-Nivel 4 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,54	\$3.240

Tabla 22: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera compensatoria ($\beta=3$) y 10 segmentos de individuos.

Al comparar con un escenario de alto ruido, en el escenario de bajo ruido se obtienen mejores tasas de hit rate, participación y utilidad esperada. En el caso de la estimación con un modelo compensatorio, al aumentar la heterogeneidad de las preferencias de la muestra, disminuyen el índice de aciertos, el índice de participación de mercado y la utilidad esperada. Por otro lado, al estimar preferencias compensatorias con un modelo no compensatorio, ocurre el efecto contrario al aumentar la heterogeneidad.

Al igual que en la condición de alto ruido, la diferencia entre la utilidad esperada de la estimación con un modelo compensatorio y no compensatorio disminuye a medida que los individuos son más heterogéneos. Esto dado que por un lado crece la utilidad obtenida de la estimación con un modelo no compensatorio y al mismo tiempo, disminuye el beneficio esperado de la estimación con un enfoque compensatorio.

5.4.3.3 Datos simulados de manera no compensatoria

Los siguientes resultados corresponden a los índices de acierto, configuración de producto óptimo, índice de participación y utilidad esperada dada la estimación de preferencias compensatorias y no compensatorias con 2 y 10 segmentos de consumidores de una muestra simulada de manera no compensatoria:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	94%	Atributo 1-Nivel 3 Atributo 2-Nivel 3 Atributo 3-Nivel 4 Atributo 4-Nivel 1 Atributo 5-Nivel 1	0,76	\$12.216

Tabla 23: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera no compensatoria y 2 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	88%	Atributo 1-Nivel 4 Atributo 2-Nivel 1 Atributo 3-Nivel 4 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,37	\$2.265

Tabla 24: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera no compensatoria y 2 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	95%	Atributo 1-Nivel 3 Atributo 2-Nivel 3 Atributo 3-Nivel 1 Atributo 4-Nivel 1 Atributo 5-Nivel 2	0,28	\$5.863

Tabla 25: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo compensatorio para datos simulados de manera no compensatoria y 10 segmentos de individuos.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación de Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	95%	Atributo 1-Nivel 1 Atributo 2-Nivel 4 Atributo 3-Nivel 4 Atributo 4-Nivel 4 Atributo 5-Nivel 4	0,46	\$2.806

Tabla 26: Producto óptimo obtenido con programación dinámica a partir de estimación de preferencias con modelo no compensatorio para datos simulados de manera no compensatoria y 10 segmentos de individuos.

A partir de los resultados se aprecia que, en el caso de la estimación con un modelo compensatorio, al aumentar la heterogeneidad de las preferencias de la muestra, disminuyen el índice de aciertos, el índice de participación de mercado y la utilidad esperada. Por otro lado, al estimar preferencias no compensatorias con un modelo no compensatorio, a medida que existe un mayor número de segmentos, mejora el índice de participación y la diferencia entre la utilidad esperada de la estimación respecto a un modelo compensatorio disminuye.

5.5 Conclusiones

Los resultados en términos de hit rate son consistentes con lo esperado de modo que, la estimación basada en un modelo compensatorio otorga mayor cantidad de aciertos respecto a la estimación basada en un modelo no compensatorio cuando se tienen datos simulados de manera compensatoria. Además, la utilidad obtenida de la estimación con un modelo compensatorio es mayor respecto a la obtenida con un modelo no compensatorio. A la vez, cuando se está en un escenario con bajo nivel de ruido ($\beta = 3$), el hit rate es mayor respecto a la condición de alto ruido ($\beta = 1$).

Cuando se trata de simulaciones no compensatorias, los resultados en términos de hit rate son consistentes con lo esperado, de modo que la estimación basada en un modelo no compensatorio otorga igual porcentaje de aciertos respecto a la estimación basada en un modelo compensatorio cuando existe máxima heterogeneidad, es decir, cuando no existen segmentos con igual orden de preferencia.

De alguna forma, cuando las preferencias son no compensatorias, la empresa está dispuesta a ceder utilidad al optar por mejores niveles para los atributos más importantes, y ganar en los atributos menos importantes. En el caso compensatorio pasa lo contrario, en que tiende a dar lo mismo el nivel de atributo escogido mientras genere mayor utilidad, independiente de su grado de importancia.

Al mismo tiempo, se observa que los índices de participación de mercado son mayores para el caso no compensatorio, lo que indica que en este caso existe un esfuerzo por otorgar mejores niveles para ganar utilidad a través de mayor aceptación, mientras que en el caso compensatorio se busca generar más utilidad seleccionando niveles con el menor costo, independiente del share entre individuos. En consecuencia, en el caso de un modelo no compensatorio importa más el factor demanda, mientras que en un modelo compensatorio es más relevante el factor precio.

Finalmente, cuando se varía el grado de heterogeneidad en las preferencias de la muestra, las diferencias entre el producto óptimo dado por las estimaciones de un modelo compensatorio y uno no compensatorio cambian. De este modo, cuando los datos son simulados de manera compensatoria o no compensatoria, al aumentar la diversidad en las preferencias, disminuye el índice de aciertos, el índice de participación de mercado y la utilidad esperada obtenidos con un modelo compensatorio. Por el contrario, estos índices aumentan cuando se utiliza un modelo no compensatorio. La explicación se da en que, al tener menos segmentos, es difícil generar un producto óptimo que cumpla con las condiciones lexicográficas de todos los individuos, por lo que al priorizar un atributo considerado el más importante para un segmento, se prescinde de otros atributos que pueden resultar más importantes para el resto de los consumidores.

Capítulo 6

Diseño de producto óptimo bajo preferencias reales

En la toma de decisiones interactúan un gran número de factores propios del individuo, como lo son sus gustos, necesidades, capacidad cognitiva y motivaciones. Ello, junto a factores externos, como el contexto de la elección, convierten a este proceso en uno de alta complejidad de entendimiento. En términos cuantitativos, la instancia de la elección constituye el ambiente en que es posible analizar a los consumidores a partir de sus acciones.

La simulación de preferencias permite evaluar el nivel de aceptación de un producto contrastando las preferencias estimadas versus las preferencias reales o simuladas. Sin embargo, cuando se realizan encuestas del tipo análisis conjunto no siempre se tiene certeza respecto a la regla de decisión utilizada por los individuos para escoger entre distintos perfiles de producto. He allí la relevancia de emplear métodos de estimación que caractericen de manera fidedigna las preferencias de cada consumidor.

De acuerdo a lo expuesto en capítulos anteriores, el tipo de regla utilizada, ya sea compensatoria o no compensatoria, tiene un impacto distintivo en las características del producto óptimo escogido. Estos resultados permiten deducir que dependiendo de la heterogeneidad en las preferencias de sus consumidores, una compañía debe ser capaz de diseñar productos adecuados a cada segmento, distinguiendo aquellos que presentan una tendencia de elección lexicográfica de los que siguen una elección compensatoria.

Para apreciar la efectividad de los métodos de estimación de preferencias y optimización de diseño de producto propuesto, se aplicarán a un set de datos de elección reales obtenidos de una encuesta del tipo análisis conjunto con perfiles de cámaras fotográficas.

6.1 Estructura y características del set de datos

El set de datos empleado considera perfiles de cámaras fotográficas digitales descritas por 5 atributos, cada uno de los cuales puede representarse por uno de 4 niveles de atributo. A continuación, se presenta la descripción de cada uno:

<i>Atributo</i>	<i>Niveles</i>			
<i>Precio</i>	US \$ 500	US \$ 400	US \$ 300	US \$ 200
<i>Batería</i>	150 Imágenes	300 Imágenes	450 Imágenes	600 Imágenes
<i>Resolución</i>	2 Megapíxeles	3 Megapíxeles	4 Megapíxeles	5 Megapíxeles
<i>Zoom Óptico</i>	2x	3x	4x	5x
<i>Tamaño</i>	SLR	Medio	Pocket	Compacto

Tabla 27: Atributos y niveles por atributo del set de datos empírico.

Cabe mencionar que los atributos Precio, Rendimiento de la Batería, Resolución y Zoom Óptico constituyen características discretas, en las que puede asumirse qué nivel es superior en utilidad a otro. Por otro lado, el tamaño representa una característica nominal, en que no es claro cuál nivel tiene mejor aceptación para un individuo. Para efectos de este trabajo de título, por simplicidad se asumirá que un menor tamaño se asocia a un mayor beneficio para el consumidor.

Este set de datos considera una muestra de 125 personas, quienes de manera individual contestaron una encuesta online compuesta por 20 preguntas de elección. En cada interrogante, tal como fue expuesto en el capítulo de simulación, cada encuestado debió escoger una de cuatro alternativas con distintos perfiles de producto asignadas de manera aleatoria. Los atributos del producto fueron presentados y descritos a los participantes antes de aplicar el cuestionario.

A continuación se muestran algunos estadísticos descriptivos para la muestra de individuos, como un forma de analizar de manera exploratoria la distribución de la preferencia por cada nivel de atributo. Posteriormente, se presentan los resultados de la estimación utilizando el modelo compensatorio y lexicográfico, para finalmente mostrar el producto óptimo obtenido en cada caso. Con ello, se espera dar respuesta a interrogantes como: ¿Qué nivel de heterogeneidad caracteriza a los individuos en términos de su capacidad para definir los atributos con mayor relevancia? ¿Qué diferencias describen al producto óptimo obtenido en el caso compensatorio respecto al no compensatorio? ¿Existe un *insight* relevante que explique estas diferencias?

6.2 Análisis descriptivo

Previo a cualquier estimación, se realizó un estudio exploratorio de los datos, incluyendo análisis gráficos y observacionales, dentro de estos últimos se corroboró que no existieran valores mal computados o errores de transcripción. Posteriormente, como primera instancia de acercamiento a los datos, se analiza la distribución en la elección de cada nivel de atributo. Como se aprecia en siguiente histograma, en su mayoría, los individuos optan por aquellas características que objetivamente reportan una mayor utilidad.

6.2.1 Distribución de la elección por nivel de precio

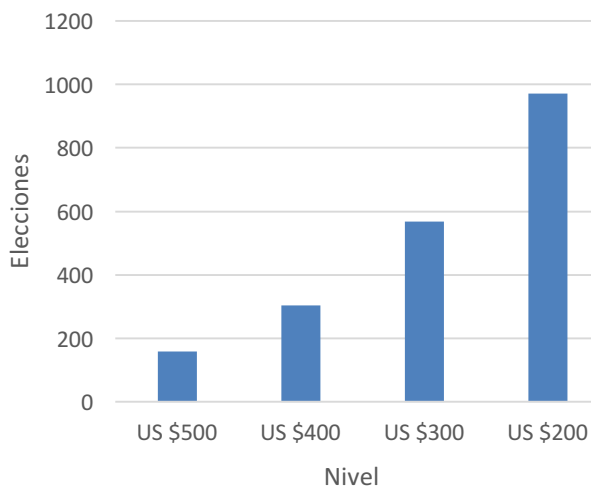


Gráfico 37: Distribución de la elección por nivel de precio.

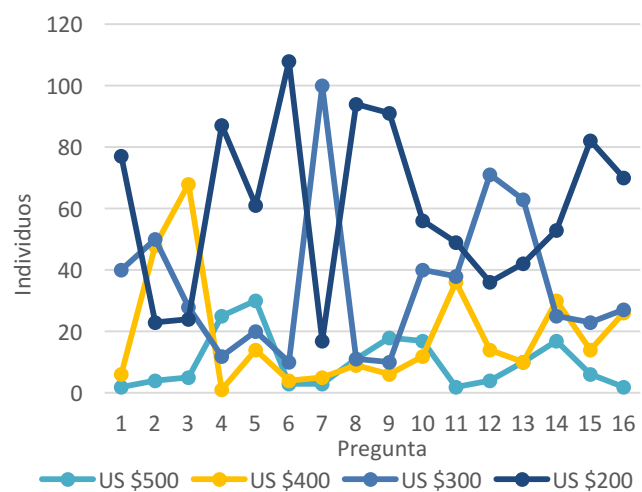


Gráfico 38: Distribución de la elección por nivel de precio y por pregunta.

6.2.2 Distribución de la elección por nivel de resolución

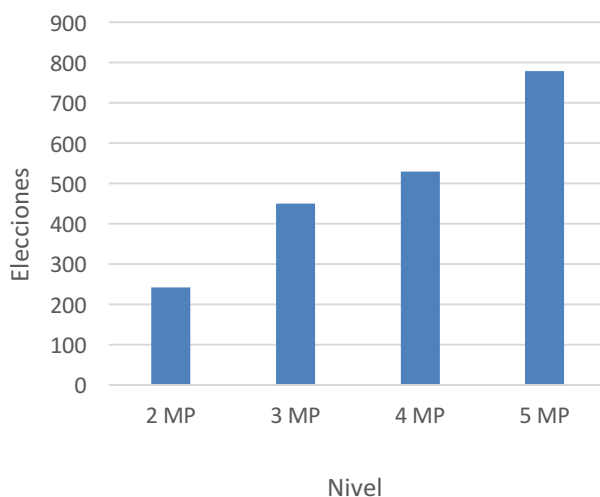


Gráfico 39: Distribución de la elección por nivel de resolución.

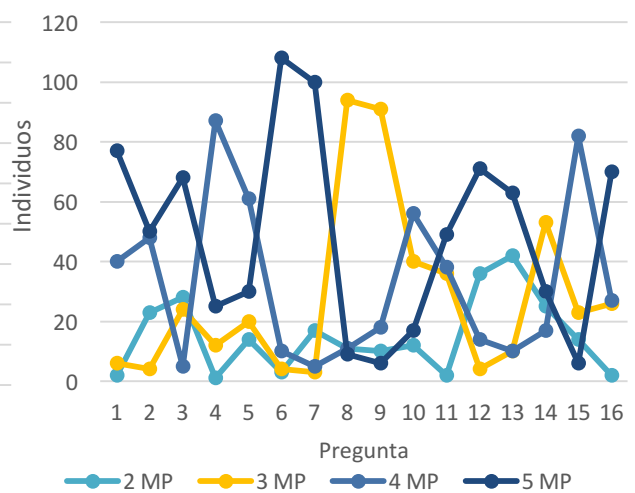


Gráfico 40: Distribución de la elección por nivel de resolución y por pregunta.

6.2.3 Distribución de la elección por nivel de batería

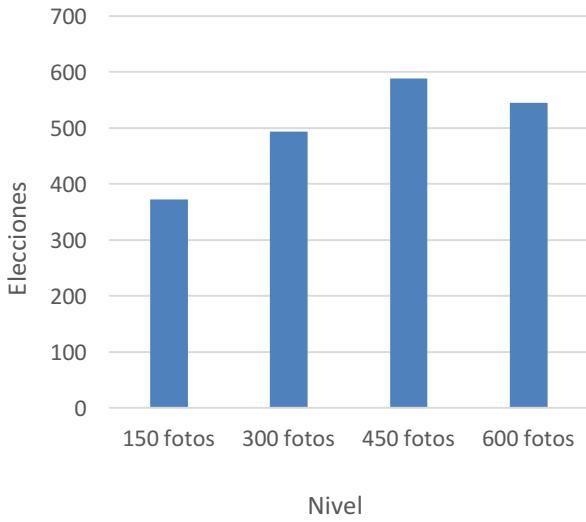


Gráfico 41: Distribución de la elección por nivel de batería.

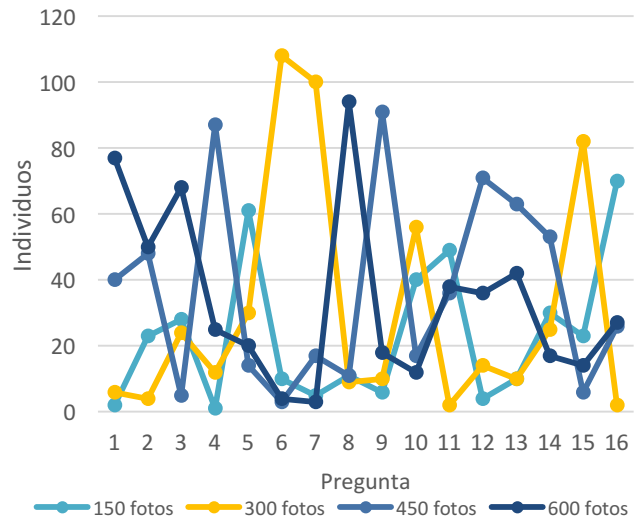


Gráfico 42: Distribución de la elección por nivel de batería y por pregunta.

6.2.4 Distribución de la elección por nivel de zoom

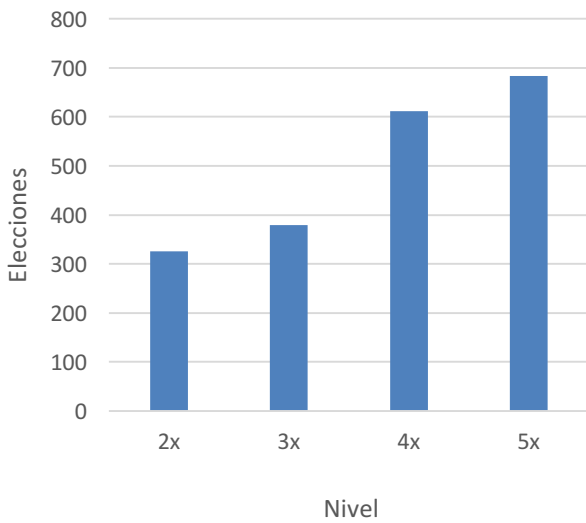


Gráfico 43: Distribución de la elección por nivel de zoom.

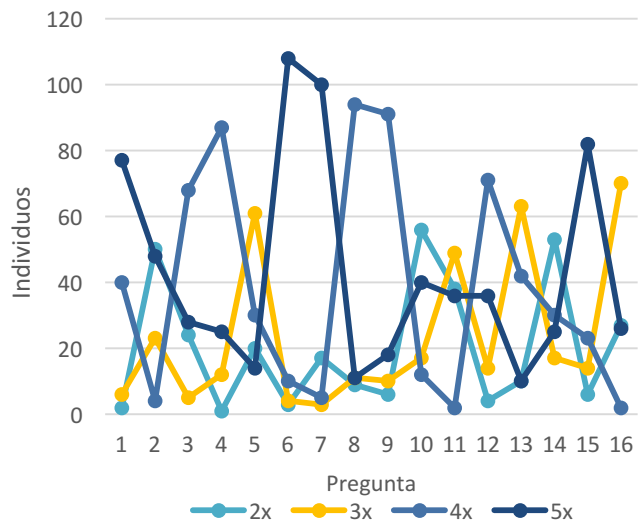


Gráfico 44: Distribución de la elección por nivel de zoom y por pregunta.

6.2.5 Distribución de la elección por nivel de tamaño

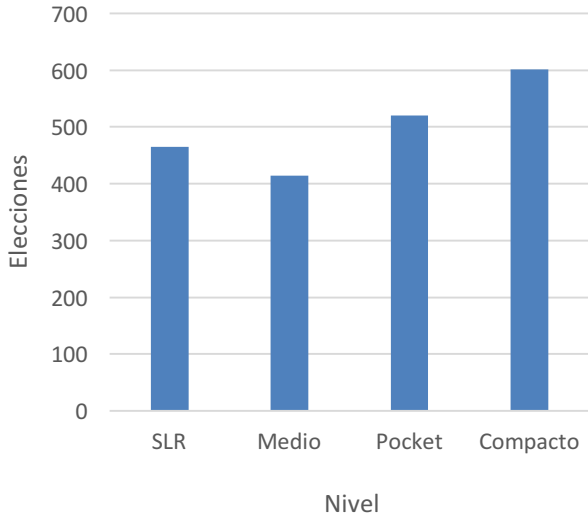


Gráfico 45: Distribución de la elección por nivel de tamaño.

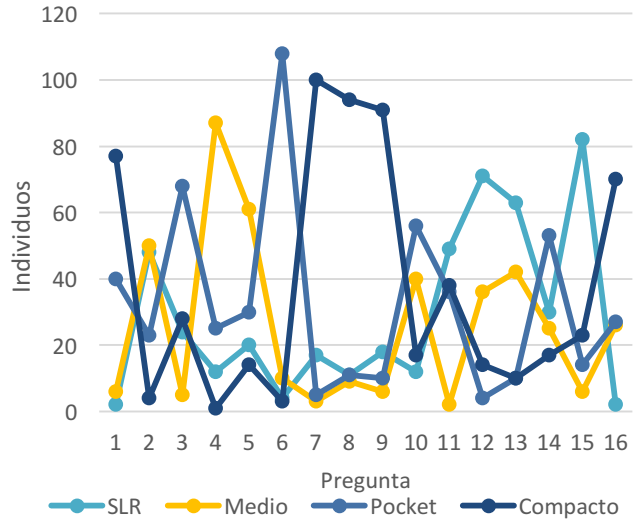


Gráfico 46: Distribución de la elección por nivel de tamaño y por pregunta.

A partir de los gráficos de línea, se observa que a lo largo de las instancias de consideración, las participaciones entre los niveles de un mismo atributo no tienen variaciones significativas, por lo que es difícil establecer algún patrón de tendencia. Para los atributos que son distintos al Precio, esta variabilidad establece un grado de dependencia con el resto de los atributos, de manera que la elección de un nivel estaría sujeta al resto de niveles presentes en el conjunto de alternativas. Por otro lado, al observar el comportamiento de la elección a nivel agregado, se observa que para el caso del Precio, Zoom, Resolución y Tamaño, existe una tendencia a escoger los mejores niveles de cada uno.

Para el caso del Tamaño, se observa menor variación en la elección entre niveles, reafirmando que característica resulta menos clara respecto a la preferencia de sus niveles. En el caso del precio se aprecia una tendencia a escoger valores más bajos, existiendo más variabilidad en la elección, lo que puede indicar que se trata de un atributo importante dentro de un orden lexicográfico.

Dado que el precio parece ser una de las variables determinantes en la elección de compra, se realiza un análisis gráfico evaluando las elecciones de cada atributo según el nivel de precio escogido:

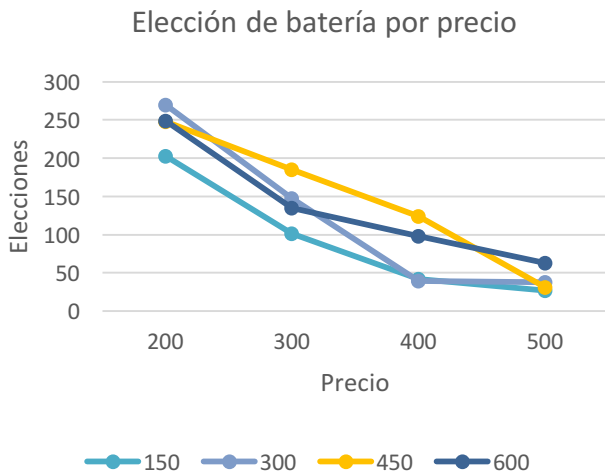


Gráfico 50: Elección de tipo de batería según nivel de precio.

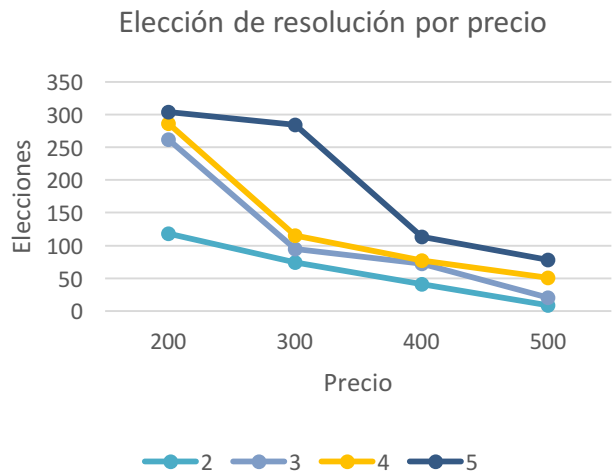


Gráfico 47: Elección de tipo de resolución según nivel de precio.

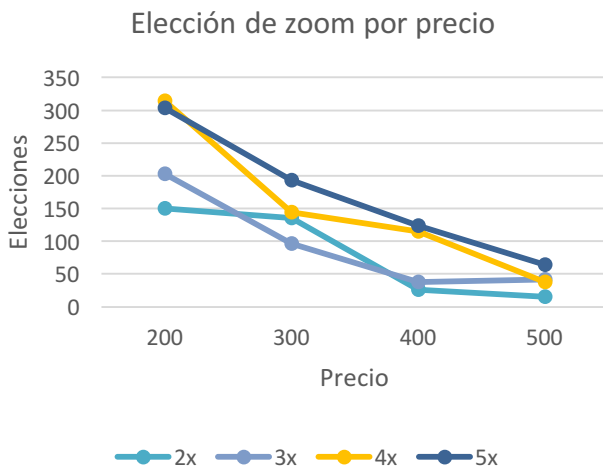


Gráfico 49: Elección de tipo de zoom según nivel de precio.

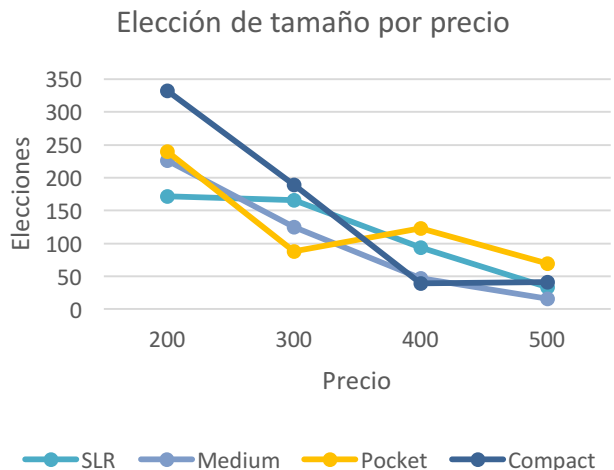


Gráfico 48: Elección de tipo de tamaño según nivel de precio.

Se observa que para un mismo atributo, independiente del nivel escogido, existe una correlación negativa respecto al precio. De este modo, a medida que el precio aumenta, la variación en la elección se reduce, manifestándose una menor preferencia a precios más elevados. En consecuencia, se observa que en general los consumidores son particularmente sensibles al precio, por lo que este debiese ser uno de los atributos más importantes al momento de evaluar distintos perfiles de producto.

Este primer análisis es útil para mostrar de manera preliminar aquellos niveles que tienen mayor preferencia entre los individuos, así como para identificar cambios en las preferencias de los niveles de un mismo atributo, dando una idea de qué atributos deberían tener un mayor orden de importancia.

6.3 Resultados

6.3.1 Estimación producto óptimo asumiendo preferencias compensatorias

Siguiendo la metodología de estimación de preferencias compensatorias del capítulo 4, se evalúan las preferencias del set de datos de cámaras a través de un modelo logit heterogéneo con estimación de preferencias individuales, en que los peores niveles de cada atributo son definidos como el nivel base para evitar colinealidad en la estimación. De este modo, se obtienen los parámetros o parth worths que definen las preferencias de un individuo por un determinado atributo. A través de este método se obtiene un hit rate del 98%, probando que la estimación mediante un modelo compensatorio presenta una alta tasa de precisión para esta muestra en particular.

Dado estos coeficientes y, siguiendo la metodología detallada en el capítulo 5, se estima el producto óptimo utilizando un enfoque de programación dinámica y búsqueda exhaustiva. De este modo, para evaluar un potencial perfil de producto óptimo, éste se agrega como una alternativa adicional a los sets de elección y se obtiene un "índice de participación" basado en un criterio de máxima utilidad. Para ello, se calcula el porcentaje de veces en que se asigna la máxima utilidad al perfil del producto óptimo respecto al total de sets de elección presentados a los individuos. Al mismo tiempo, se calcula la utilidad esperada en base a este índice y a los costos de cada nivel de atributo, los que se detallan en el Apéndice B:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	98%	Precio 500 Resolución 5 MP Batería 450 Zoom 4x Tamaño SLR	0,58	\$9.312

Tabla 28: Producto óptimo obtenido con programación dinámica a partir de estimación con modelo compensatorio para conjunto de datos reales asumiendo preferencias compensatorias.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	98%	Precio 500 Resolución 5 MP Batería 450 Zoom 4x Tamaño SLR	0,58	\$9.312

Tabla 29: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación con modelo compensatorio para conjunto de datos reales asumiendo preferencias compensatorias.

Para comparar con el desempeño de un enfoque no compensatorio, se evalúan los órdenes de preferencia del set de datos de cámaras a través de un modelo lexicográfico.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	28%	Precio 200 Resolución 2 MP Batería 600 Zoom 5x Tamaño Compact	0,04	\$240

Tabla 30: Producto óptimo obtenido con programación dinámica a partir de estimación con modelo no compensatorio para conjunto de datos reales asumiendo preferencias compensatorias.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	28%	Precio 200 Resolución 2 MP Batería 600 Zoom 5x Tamaño Medium	0,03	\$260

Tabla 31: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación con modelo no compensatorio para conjunto de datos reales asumiendo preferencias compensatorias.

A partir de estos resultados se observa que el uso de un modelo compensatorio genera buenos resultados en términos de hit rate, índice de participación de mercado y utilidad esperada cuando se asumen preferencias compensatorias. Esta deducción es consecuente con lo esperado, ya que se evalúa la configuración de producto óptimo en el mismo tipo de preferencias con las que fue obtenido.

Por otro lado, al evaluar el producto óptimo obtenido mediante un enfoque no compensatorio sobre preferencias compensatorias se genera una baja participación de mercado y con ello un pequeño beneficio esperado.

Cabe destacar que al igual que en los escenarios simulados del capítulo 5, no se aprecian grandes diferencias entre el método de programación dinámica y la búsqueda exhaustiva.

6.3.2 Estimación producto óptimo asumiendo preferencias no compensatorias

Considerando la metodología de estimación de órdenes de preferencia no compensatorios del capítulo 3, se evalúan las preferencias del set de datos de cámaras a través de un modelo lexicográfico.

De este modo, se obtienen los coeficientes que definen el orden de importancia de cada individuo por un cierto atributo. A través de este método se obtiene un hit rate del 28%. Pese a que este valor es bajo al obtenido con el enfoque compensatorio, sugiere que existe un grupo de individuos para el cual la estimación de preferencias a través de un modelo no compensatorio resulta útil.

Dado estos coeficientes, se estima el producto óptimo utilizando un enfoque de programación dinámica y búsqueda exhaustiva. Nuevamente, siguiendo la metodología mostrada en el capítulo 5, se calcula el índice de participación de mercado y la utilidad esperada de acuerdo a la estructura de costos definida en el Apéndice B:

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	28%	Precio 200 Resolución 2 MP Batería 600 Zoom 5x Tamaño Compact	0,60	\$3.612

Tabla 32: Producto óptimo obtenido con programación dinámica a partir de estimación con modelo no compensatorio para conjunto de datos reales asumiendo preferencias no compensatorias.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo No Compensatorio (Lexicográfico)</i>	28%	Precio 200 Resolución 2 MP Batería 600 Zoom 5x Tamaño Medium	0,56	\$3.900

Tabla 33: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación con modelo no compensatorio para conjunto de datos reales asumiendo preferencias no compensatorias.

Para comparar con el desempeño de un enfoque compensatorio, se evalúa la importancia relativa de cada atributo del set de datos de cámaras a través de un modelo logit.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Programación Dinámica a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	98%	Precio 500 Resolución 5 MP Batería 450 Zoom 4x Tamaño SLR	0,08	\$1.368

Tabla 34: Producto óptimo obtenido con programación dinámica a partir de estimación con modelo compensatorio para conjunto de datos reales asumiendo preferencias no compensatorias.

Método de estimación de producto óptimo	Hit Rate Estimación Preferencias	Producto Óptimo	Índice de Participación Mercado	Utilidad Esperada (\$US)
<i>Búsqueda Exhaustiva a partir de preferencias obtenidas con Modelo Compensatorio (Logit)</i>	98%	Precio 500 Resolución 5 MP Batería 450 Zoom 4x Tamaño SLR	0,08	\$1.368

Tabla 35: Producto óptimo obtenido con búsqueda exhaustiva a partir de estimación con modelo compensatorio para conjunto de datos reales asumiendo preferencias no compensatorias.

En base a los resultados obtenidos, se infiere que el uso de un modelo no compensatorio no genera buenos resultados en términos de hit rate, índice de participación de mercado ni utilidad esperada cuando se asumen preferencias no compensatorias para este set de datos en particular. Esto indica que las preferencias de los individuos de la muestra presentan cualidades de carácter más compensatorio. Por otro lado, al evaluar el producto óptimo obtenido mediante un enfoque compensatorio sobre preferencias no compensatorias se genera una menor participación de mercado y beneficio esperado respecto al primer escenario.

Es importante mencionar que al igual que en los escenarios simulados del capítulo 5, no se aprecian grandes diferencias entre el método de programación dinámica y la búsqueda exhaustiva.

Al evaluar los resultados de ambos enfoques, cuando se asumen preferencias compensatorias y no compensatorias, se cumple que las diferencias en la participación de mercado y la utilidad esperada obtenidas entre un modelo compensatorio y uno no compensatorio son menores cuando el producto óptimo es evaluado asumiendo preferencias no compensatorias.

Sea Δ^1 la diferencia en la utilidad al comparar el perfil de producto óptimo obtenido con un modelo compensatorio cuando se asumen preferencias compensatorias (X_C^C) frente al producto obtenido con un modelo no compensatorio cuando se asumen preferencias compensatorias (X_{NC}^C):

$$\Delta^1 = U_{X_C^C} - U_{X_{NC}^C}$$

Sea Δ^2 la diferencia en la utilidad al comparar el perfil de producto óptimo obtenido con un modelo compensatorio cuando se asumen preferencias no compensatorias (X_C^{NC}) frente al producto obtenido con un modelo no compensatorio cuando se asumen preferencias no compensatorias (X_{NC}^{NC}):

$$\Delta^2 = U_{X_{NC}^{NC}} - U_{X_C^{NC}}$$

Luego, a partir de los resultados obtenidos, se cumple que la variación en la utilidad Δ^1 al comparar X_C^C y X_{NC}^C es mayor a la variación Δ^2 resultante de comparar X_{NC}^{NC} y X_C^{NC} :

$$\Delta^1 \geq \Delta^2$$

6.4 Conclusiones

Al aplicar los métodos de estimación de preferencias y optimización de producto a un set de datos reales se evidencia que el uso de un modelo compensatorio resulta efectivo y robusto, alcanzando altas tasas de acierto, índice de participación de mercado y utilidad esperada. Cabe destacar, al igual que se detalló en el capítulo 4, que estos resultados se explican por el uso de un modelo logit a nivel individual, que utiliza las elecciones propias de cada persona para la estimación de sus preferencias.

Por otro lado, para este conjunto de datos en particular, la aplicación del modelo lexicográfico resulta ser menos efectiva, ya que no coincide con lo esperado en el análisis exploratorio, se obtiene un menor hit rate junto a bajos porcentajes de participación de mercado y por ende un menor beneficio esperado en comparación a un modelo compensatorio.

A pesar que la tasa de acierto obtenida con esta metodología no es alta, sugiere que existe un grupo de individuos que es más proclive al uso de una regla no compensatoria, pese a que el grupo en general muestre una tendencia compensatoria. Esto se sustenta en que existe un porcentaje de las elecciones que puede ser explicado por el uso de un modelo de estimación de preferencias no compensatorio.

Dicho esto, las diferencias en la participación de mercado y la utilidad esperada obtenidas entre un modelo compensatorio y uno no compensatorio son mayores cuando el producto óptimo es evaluado asumiendo preferencias compensatorias. Esto coincide con lo esperado, pues por construcción el modelo compensatorio tiende a ser más flexible, adecuándose mejor a las preferencias reflejadas por los datos, incluso en escenarios no compensatorios. Por el contrario, un modelo no compensatorio no tiene capacidad de adaptación en ambientes en que las preferencias son compensatorias, siendo más restrictivo en este aspecto.

En definitiva, cuando no se conoce la regla de decisión utilizada por los consumidores, convendrá aplicar ambos enfoques, compensatorio y no compensatorio de manera conjunta, analizando su desempeño desde una perspectiva individual y no sobre la muestra en general.

Capítulo 7

Conclusiones

7.1 Conclusiones generales

El diseño óptimo de productos se ha convertido en un aspecto relevante para retailers y desarrolladores, cuyas decisiones impactan en la elección de compra de los individuos y son determinantes en las utilidades que se perciben.

Estudiar el diseño de productos a partir de la elección de los consumidores por determinados perfiles de producto resulta útil e interesante de abordar, pues permite entender a los individuos desde una perspectiva conductual que toma en consideración variables que influyen en la decisión de compra. Estos factores contemplan aquellos inherentes a las preferencias de cada persona, así como características propias del contexto de la elección.

Si los consumidores se guiaran solamente por un comportamiento racional, deberían ser capaces de evaluar las características de todas las alternativas a las que se enfrentan cuando deben optar por un perfil de producto. En este supuesto se basa gran parte de la literatura actual en materia de diseño de productos, asumiendo modelos compensatorios para la estimación de preferencias. A partir de este escenario, en el trabajo de investigación desarrollado se evalúa el diseño óptimo de productos bajo reglas de decisión no compensatorias, estableciendo puntos de comparación que cuantifican las diferencias entre los perfiles obtenidos con ambos enfoques.

La investigación realizada se enmarca en un escenario experimental donde los individuos muestran sus preferencias por perfiles de producto en encuestas del tipo análisis conjunto. Este tipo de encuestas, ampliamente utilizado en la industria, resulta eficiente y práctico, dado que permite examinar simultáneamente muchos factores con una cantidad reducida de preguntas.

De esta forma, dado el uso de una regla lexicográfica, se puede capturar el grado de importancia que tiene cada atributo, siendo consistente con un escenario realista, en que es difícil evaluar todas las características de un producto de manera exhaustiva.

En un ambiente de simulación, cuando se conoce el tipo de regla de decisión utilizado por los individuos, ambos métodos de estimación, tanto compensatorio, como no compensatorio, son eficientes estimando preferencias individuales que recuperan el orden de importancia de los distintos niveles de atributo. De este modo, las buenas predicciones obtenidas en esta etapa guían el estudio a un contexto experimental donde la estrategia de decisión es desconocida.

En esta línea, al emplear los modelos de preferencia en el diseño óptimo de productos sobre un conjunto de datos simulados, se observa que en el caso de un modelo no compensatorio parece importar más el factor demanda, mientras que en un modelo compensatorio son más relevantes los costos de la firma. Esto se sustenta en los resultados expuestos en el capítulo 5 y 6, según los cuales el producto óptimo obtenido a partir de preferencias compensatorias presenta niveles de atributo más básicos, con un menor costo para la firma productora. Por otro lado, el producto óptimo derivado de preferencias no compensatorias se caracteriza por tener mejores niveles de atributo, que son más costosos, pero a la vez tienen mayor aceptación entre los individuos, haciendo que la utilidad esperada esté dada por un mayor porcentaje de participación. Esto se contrasta al primer caso, en la ganancia se relaciona a niveles de atributo con menores costos de producción. De este modo, cumplir con las exigencias de los consumidores cobra mayor relevancia cuando se diseña un producto en base a decisiones no compensatorias, dado que la utilidad está dada por un mayor porcentaje de elección.

De lo anterior, se deduce que la relevancia del diseño óptimo de productos está sujeta a variaciones en el contexto de la decisión. De esta forma, a medida que aumenta la diversidad en las preferencias, es decir, cuando se tiene una muestra más heterogénea, disminuye el índice de participación de mercado y la utilidad esperada obtenida con un modelo compensatorio. Por el contrario, cuando se utiliza un modelo no compensatorio, estas métricas tienden a ser crecientes a mayor grado de heterogeneidad. Ello indica que este tipo de modelo es más adecuado cuando se tiene una población heterogénea, resultando menos robusto a nivel general, en condiciones en que las distintas preferencias de los consumidores, en términos de órdenes de importancia, se encuentran distribuidas entre menos segmentos de individuos, con tendencias de elección distintas entre uno y otro.

Al comparar ambos enfoques con un set de datos reales, se observa que las diferencias en la participación de mercado y la utilidad esperada obtenidas entre un modelo compensatorio y uno no compensatorio son mayores cuando el producto óptimo es evaluado asumiendo preferencias compensatorias, reafirmando el hecho que el modelo compensatorio es robusto en múltiples escenarios. Pese a ello, el uso de ambos enfoques difiere en la configuración del producto óptimo obtenido, lo que conduce la investigación a la exploración de nuevas reglas de decisión, fuentes de datos y condiciones experimentales.

Así, se da cumplimiento al objetivo principal de este trabajo, por cuanto la metodología propuesta es capaz de estimar las preferencias de los consumidores, para luego utilizar dicha información en el diseño de productos usando modelos compensatorios y no compensatorios, permitiendo caracterizar los resultados bajo distintos escenarios, tanto simulados, como reales.

En conclusión, la estimación de preferencias basadas en información de elección de los consumidores es útil para el diseño de productos. No obstante, los resultados dependerán del tipo de modelo utilizado, los supuestos empleados y del grado de heterogeneidad de los individuos de la muestra. Dicho esto, una compañía que invierte en el diseño de productos y se involucra entendiendo las preferencias de sus consumidores puede obtener mayores retornos y ser escogida por sobre su competencia.

7.2 Discusión general y trabajos futuros

El desarrollo de esta investigación considera múltiples supuestos, los que han sido utilizados con el fin de proveer escenarios más flexibles y simplificados considerando el contexto de simulación y el tipo de datos disponibles para su validación.

Una de las primeras suposiciones tiene relación con asumir que los consumidores siempre escogen un perfil de producto. De este modo, se deja fuera la opción de no compra para estimar los modelos de elección en su forma original. Si bien, en un modelo compensatorio es posible modelar esta opción como una alternativa extra, a la que se puede asignar una utilidad base, dicha adaptación no es necesariamente replicable para el caso no compensatorio. Por otro lado, el objetivo de este trabajo es evaluar el diseño de productos basado en la importancia que recibe cada atributo. Por ende, en un escenario de simulación, la indiferencia por un set de perfiles de productos no es un indicador directo de la preferencia relativa de un atributo sobre otro. Sin embargo, se propone indagar en esta línea, explorando formas de incluir la opción de no compra para descubrir el impacto que tiene ello en el uso de reglas compensatorias y no compensatorias.

En segundo lugar, este trabajo se ha realizado aludiendo al diseño óptimo de productos, considerando que el diseño de surtido puede ser abordado de manera similar. De este modo, si bien es relevante estudiar la selección de características óptimas de un producto previo a su introducción al mercado, también lo es definir qué perfiles disponer en una góndola o página de productos, especialmente cuando se trata de un retailer, quien escoge el surtido considerando las opciones disponibles en el mercado. Así, resulta interesante evaluar esta relación entre diseño y surtido de productos en presencia de preferencias no compensatorias.

Por otro lado, los principales métodos de optimización utilizados corresponden a heurísticas de programación dinámica, que no siempre aseguran un óptimo global. Si bien se evalúan diferencias entre los productos óptimos del enfoque compensatorio y no compensatorio bajo una misma heurística y se realiza una comparación con búsqueda

exhaustiva, se propone el uso de un método de optimización alternativo, como por ejemplo el uso de algoritmos genéticos, ya que la optimización exhaustiva es costosa en términos de tiempo y recursos de procesamiento.

También, como futura línea de investigación se propone evaluar el diseño de producto óptimo cuando existe interacción entre atributos o niveles de un mismo atributo, ya que en un contexto más realista, es esperable que la presencia de cierto tipo de atributos determine el orden de importancia que pueden recibir otros. A la vez, en este estudio se asume un orden de importancia determinístico sobre el total de atributos, por lo que queda la posibilidad de evaluar un orden parcial de estos e incorporar probabilidades en la importancia de cada uno. Al mismo tiempo, se puede explorar la incorporación de otras variables que describen el comportamiento de elección de los consumidores y que pueden contribuir a una mayor predicción, como lo es la fijación visual, neuronal o factores sociodemográficos.

En este trabajo de título, se abordan las preferencias no compensatorias a través de la regla lexicográfica. No obstante, este es sólo uno del total de criterios de decisión conocidos. Es por ello, que se propone ampliar la investigación a otro tipo de reglas, como la conjuntiva o disyuntiva.

Como se explicó en el capítulo 5, en este trabajo existen dos funciones de utilidad distintas: La utilidad que percibe un consumidor al momento de escoger un perfil de producto, basado en sus preferencias por cada nivel de atributo, y la utilidad que recibe la compañía que diseña el producto. Para ello se simuló costos y precios para cada nivel de atributo. Lo anterior es una forma de modelar el trade-off presente entre ofrecer productos más atractivos que a la vez cuestan más. Esta conjetura busca una representación más realista de un problema en el que no siempre se tienen los costos e ingresos derivados de la producción y comercialización de cada perfil de producto, pero que deben tenerse en cuenta como criterios de rentabilidad fundamentales para la firma. Como trabajo futuro se propone utilizar información verdadera de costos para evaluar el diseño de productos bajo preferencias compensatorias y no compensatorias con un enfoque más realista desde el punto de vista de la firma.

Como se mencionó anteriormente, los individuos utilizan reglas de decisión compensatorias y no compensatorias dependiendo del contexto de la elección. Por ello, se sugiere que futuras investigaciones en materia de diseño de productos aborden ambos enfoques de manera conjunta, estudiando su interacción. En esta línea, se propone incursionar en la integración con un modelo Consider-Then-Choose, según el cual los consumidores eligen en dos etapas, primero seleccionando qué alternativas considerar y luego haciendo un ranking de ellas (Gaskin, Evgeniou, Bailiff & Hauser, 2007 ; Aouad, Farias & Levi, 2015). Con ello, se ampliaría la investigación desde un escenario ideal, en que se utiliza una regla u otra de manera excluyente, a otro más complejo en que ambas se vinculan y complementan en distintas etapas, permitiendo representar de manera más precisa la realidad en que están inmersas las preferencias de los consumidores.

Apéndice A

Estimación de orden de preferencias lexicográfico por nivel de atributo

Sea $x = (x_1, \dots, x_N)$ un perfil de producto compuesto por m atributos, cada uno de los cuales posee n_k niveles de atributo, de modo que $N = \sum_{k=1}^m n_k$ representa el total de niveles de atributo. De esta forma, x_k toma el valor 1 si el nivel k forma parte de dicha alternativa y 0 en caso contrario.

Sea $X = (x, \dots, x^w)$ una matriz con w perfiles de producto, que representa un set de consideración con w alternativas de elección. La utilidad que reporta la alternativa j al individuo i está dada por la expresión:

$$u_i(x^j) = \sum_{l=1}^N \beta_l * x_l^j$$

De este modo, se asume que la alternativa escogida para el set de consideración X será aquella que represente mayor utilidad.

Para inferir el orden de preferencias lexicográfico que un individuo i tiene sobre los N niveles de atributos de un producto, se busca maximizar el porcentaje de veces en que la utilidad generada por un determinado orden de niveles, coincide con la alternativa real escogida para el set de consideración X . En consecuencia, se escoge aquel orden con el mayor número de éxitos para el total instancias de consideración.

Para encontrar el orden lexicográfico que maximiza el porcentaje de aciertos es posible recurrir al uso de una heurística Greedy, como la que se señala:

Iteración inicial:

Sea $S_0 = \{1, \dots, N\}$ un vector que denota el set de niveles del total de atributos. Para cada atributo k , se disponen los n_k niveles en un determinado orden de preferencia. Un perfil de producto es representado por un vector $x = (x_1, \dots, x_N)$, en que el nivel l está identificado por un valor $x_l \in [0,1]$, de manera que para cada individuo i y para cada alternativa x^j del set de consideración X , se computa:

$$u_{i1l}(x^j) = \frac{1}{2} * x_l^j, \quad \forall l \in S_0$$

Sea $\omega = (x^j, J)$, con $J = (x^1, \dots, x^w) \setminus x^j$, los perfiles de producto que conforman el set de consideración X , siendo w igual al total de alternativas disponibles. Sea Ω equivalente al total de instancias de consideración ω a las que se enfrenta un individuo. Para cada $\omega \in \Omega$ se calcula:

$$d_{i1l}(x^j, J) = \begin{cases} 1, & u_{i1l}(x^j) < u_{i1l}(J) \\ 0, & \text{en caso contrario} \end{cases} \quad \forall l \in S_0$$

Sea

$$Z_{i1l}(\Omega) = \sum_{\omega \in \Omega} d_{i1l}(x^j, J), \quad \forall l \in S_0$$

Se denota l_1 como aquel nivel para el cual el valor de $Z_{i1l}(\Omega)$ tiene el menor valor respecto al total de niveles $l \in S_0$. De este modo:

$$Z_{i1l_1}(\Omega) = \min\{Z_{i1l}(\Omega) \mid l \in S_0\}$$

De esta forma, se selecciona l_1 como el nivel de mayor importancia. En caso de existir empate en el valor de $Z_{i1l_1}(\Omega)$, se escoge un nivel de manera aleatoria y se asigna:

$$u_{i1} = \frac{1}{2}$$

y

$$S_1 = S_0 \setminus \{l_1\}$$

En consecuencia, S_1 representa el set de niveles de atributo restantes, una vez que se ha escogido a l_1 como el nivel de mayor relevancia.

Recursión:

Para cada alternativa x^j del set de consideración X , se computa:

$$u_{itl}(x^j) = u_{i(t-1)l}(x^j) + \frac{1}{2^t} * x_l^j, \quad \forall l \in S_{(t-1)}$$

En que:

$$S_{(t-1)} = S_0 \setminus \{l_1, \dots, l_{(t-1)}\}$$

constituye el set de atributos no seleccionados tras el paso $t - 1$, mientras que:

$$u_{i(t-1)l}(x^j) = \frac{1}{2} * x_{l_1}^j + \dots + \frac{1}{2^{(t-1)}} * x_{l_{(t-1)}}^j$$

De esta forma, para cada $\omega = (x^j, J) \in \Omega$ se calcula:

$$d_{itl}(x^j, J) = \begin{cases} 1, & u_{itl}(x^j) < u_{itl}(J) \\ 0, & \text{en caso contrario} \end{cases} \quad \forall l \in S_{(t-1)}$$

Sea

$$Z_{itl}(\Omega) = \sum_{\omega \in \Omega} d_{itl}(x^j, J), \quad \forall l \in S_{(t-1)}$$

Se denota l_t como aquel nivel para el cual el valor de $Z_{itl}(\Omega)$ tiene el menor valor respecto al total de niveles $l \in S_{(t-1)}$. De este modo:

$$Z_{itl_t}(\Omega) = \min\{Z_{itl}(\Omega) | l \in S_{(t-1)}\}$$

De esta forma, se selecciona l_t como el t-ésimo nivel de mayor importancia. En caso de existir empate en el valor de $Z_{itl_t}(\Omega)$, se escoge un nivel de manera aleatoria y se asigna:

$$u_{it} = \frac{1}{2^t}$$

y

$$S_t = S_{(t-1)} \setminus \{l_t\}$$

Iteración final:

El algoritmo termina una vez que se ha obtenido el orden de preferencia para todos los niveles de atributo.

Apéndice B

Estructura de costos para conjunto de datos de elección de cámaras digitales

A continuación, se detalla la estructura de costos para el escenario con datos reales provenientes de una encuesta del tipo análisis conjunto de elección de cámaras digitales:

<i>Atributo</i>	Nivel de Atributo	Costo (\$USD)	Precio (\$USD)
<i>Precio</i>	US \$ 500	20	80
	US \$ 400	40	
	US \$ 300	60	
	US \$ 200	80	
<i>Batería</i>	150 Imágenes	20	80
	300 Imágenes	40	
	450 Imágenes	60	
	600 Imágenes	80	
<i>Resolución</i>	2 Megapíxeles	20	80
	3 Megapíxeles	40	
	4 Megapíxeles	60	
	5 Megapíxeles	80	
<i>Zoom Óptico</i>	2 x	20	80
	3 x	40	
	4 x	60	
	5 x	80	
<i>Tamaño</i>	SLR	20	80
	Medio	40	
	Pocket	60	
	Compacto	80	

Tabla 36: Estructura de costos para diseño de producto óptimo con datos reales de elección de cámaras digitales .

Esta distribución de costos es similar a la utilizada en el capítulo 5, en que la proporcionalidad existente es una forma de modelar el hecho de que para la firma existe un trade-off entre ofrecer productos más atractivos, que tienen mejor aceptación, pero que son más costosos. Por otro lado, el precio asociado al atributo Precio modela que precios más caros son más atractivos para la firma, a diferencia de los precios más baratos.

Bibliografía

- [1] Hauser, J., Ding, M., & Gaskin, S. P. (2009, May). Non-compensatory (and compensatory) models of consideration-set decisions. 2009 Sawtooth Software Conference Proceedings, Sequim WA.
- [2] Foerster, J. F. (1979). Mode choice decision process models: a comparison of compensatory and non-compensatory structures. *Transportation Research Part A: General*, 13(1), 17-28.
- [3] Gilbride, T. J., & Allenby, G. M. (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, 23(3), 391-406.
- [4] Cachon, G. P., Terwiesch, C., & Xu, Y. (2005). Retail assortment planning in the presence of consumer search. *Manufacturing & Service Operations Management*, 7(4), 330-346.
- [5] Dieckmann, A., Dippold, K., & Dietrich, H. (2009). Compensatory versus non-compensatory models for predicting consumer preferences. *Judgment and Decision Making*, 4(3), 200.
- [6] Gaskin, S., Evgeniou, T., Bailiff, D., & Hauser, J. (2007, October). Two-stage models: Identifying non-compensatory heuristics for the consideration set then adaptive polyhedral methods within the consideration set. In *Proceedings of the Sawtooth Software Conference* (Vol. 13, pp. 67-83).
- [7] Aouad, A., Farias, V. F., & Levi, R. (2015). Assortment Optimization Under Consider-Then-Choose Choice Models.
- [8] Mantrala, M. K., Levy, M., Kahn, B. E., Fox, E. J., Gaidarev, P., Dankworth, B., & Shah, D. (2009). Why is assortment planning so difficult for retailers? A framework and research agenda. *Journal of Retailing*, 85(1), 71-83.

[9] Intelligence Node. Product Assortment. Visitado el 10 de abril de 2017 en <http://www.intellicencenode.com/knowledge/product-assortment-planning-retail/>.

[10] Kohli, R., & Jedidi, K. (2007). Representation and inference of lexicographic preference models and their variants. *Marketing Science*, 26(3), 380-399.

[11] Maldonado, S., Montoya, R., & Weber, R. (2015). Advanced conjoint analysis using feature selection via support vector machines. *European Journal of Operational Research*, 241(2), 564-574.

[12] Urrutia, P. (2014) Estudio del uso de reglas de decisión lexicográficas utilizando herramientas de eye tracking.

[13] Arora N and Huber J (2001). Improving parameter estimates and model prediction by aggregate customization in choice experiments. *Journal of Consumer Research* 28(2):273–283.

[14] Wikipedia. NP-Hard. Visitado el 14 de septiembre de 2018 en <https://es.wikipedia.org/wiki/NP-hard>.

[15] Kohli, R., & Krishnamurti, R. (1989). Optimal product design using conjoint analysis: Computational complexity and algorithms. *European Journal of Operational Research*, 40(2), 186-195.

[16] Grether, David and Louis Wilde, An Analysis of Conjunctive Choice, *Journal of Consumer Research*, March 1984, 10, 373-385.

[17] Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). Introduction to algorithms. MIT press.

