



Universidade de Vigo

Trabajo Fin de Máster

Aplicación de técnicas de análisis conjunto para la determinación de la importancia de variables físicas en aprovechamientos forestales

David Iglesias González

Máster en Técnicas Estadísticas

Curso 2022-2023

Propuesta de Trabajo Fin de Máster

Título en galego: Aplicación de técnicas de análise conxunto para a determinación da importancia de variables físicas en aproveitamentos forestais
Título en español: Aplicación de técnicas de análisis conjunto para la determinación de la importancia de variables físicas en aprovechamientos forestales
English title: Application of conjoint analysis to determine the importance of physical variables in forest exploitation
Modalidad: Modalidad A
Autor: David Iglesias González, Universidad de Vigo
Directoras: María del Carmen Iglesias Pérez, Universidad de Vigo; Laura Alonso Martínez, Universidad de Vigo
Breve resumen del trabajo: En este trabajo se realizó un análisis conjunto basado en elecciones para explorar y cuantificar el sistema de valores de las principales empresas y asociaciones forestales de Galicia en la elección de una parcela para su explotación forestal a través de diferentes atributos o factores físicos del terreno.

Doña María del Carmen Iglesias Pérez, Catedrática de Escuela Universitaria de la Universidad de Vigo, doña Laura Alonso Martínez, Contratada FPU de la Universidad de Vigo, informan que el Trabajo Fin de Máster titulado

Aplicación de técnicas de análisis conjunto para la determinación de la importancia de variables físicas en aprovechamientos forestales

fue realizado bajo su dirección por don David Iglesias González para el Máster en Técnicas Estadísticas. Estimando que el trabajo está terminado, dan su conformidad para su presentación y defensa ante un tribunal.

En Vigo, a 14 de julio de 2023.

El/la director/a:
Don/doña María del Carmen Iglesias Pérez

El/la director/a:
Don/doña Laura Alonso Martínez

El/la autor/a:
Don/doña David Iglesias González

Declaración responsable. Para dar cumplimiento a la Ley 3/2022, de 24 de febrero, de convivencia universitaria, referente al plagio en el Trabajo Fin de Máster (Artículo 11, [Disposición 2978 del BOE núm. 48 de 2022](#)), **el/la autor/a declara** que el Trabajo Fin de Máster presentado es un documento original en el que se han tenido en cuenta las siguientes consideraciones relativas al uso de material de apoyo desarrollado por otros/as autores/as:

- Todas las fuentes usadas para la elaboración de este trabajo han sido citadas convenientemente (libros, artículos, apuntes de profesorado, páginas web, programas, . . .)
- Cualquier contenido copiado o traducido textualmente se ha puesto entre comillas, citando su procedencia.
- Se ha hecho constar explícitamente cuando un capítulo, sección, demostración, . . . sea una adaptación casi literal de alguna fuente existente.

Y, acepta que, si se demostrara lo contrario, se le apliquen las medidas disciplinarias que correspondan.

Agradecimientos

En primer lugar, me gustaría agradecer a M^a del Carmen Iglesias, Julia Armesto y Laura Alonso la oportunidad que me brindaron al ofrecerme este proyecto con el que he tenido mi primer acercamiento a la aplicación de técnicas estadísticas en situaciones reales, y me ha permitido comprobar la utilidad de los resultados obtenidos, haciéndome sentir que el trabajo ha merecido la pena. También quiero darles las gracias por la paciencia que han tenido conmigo y sobre todo por el tiempo y dedicación empleados.

También quiero agradecer a mi familia el haber estado conmigo desde el principio, apoyándome en todo. Así como a mis amigos y compañeros de trabajo que me animaron a seguir adelante.

Índice general

Resumen	xv
1. Planteamiento del problema	1
1.1. Introducción	1
1.2. Caso de estudio	4
1.3. Definición de las variables	6
2. Análisis conjunto	7
2.1. Presentación	7
2.2. Elementos y fases del análisis conjunto basado en elecciones	9
2.2.1. Diseño de la encuesta	10
2.2.2. Recogida de datos	17
2.2.3. Estimación de las utilidades parciales	18
2.3. Otros tipos de análisis conjunto	19
2.3.1. Análisis conjunto tradicional (AC)	19
2.3.2. Análisis conjunto adaptativo (ACA)	19
2.3.3. Análisis conjunto autoexplicativo	21
2.4. Selección de la técnica de estudio	22
3. Aplicación del análisis conjunto basado en elecciones	23
3.1. Descripción de los datos	23
3.1.1. Estadísticos descriptivos	24
3.1.2. Histogramas	25
3.2. Aplicación del ACBE	26
3.2.1. Diseño de la encuesta	26
3.2.2. Recogida de datos	34
3.2.3. Estimación de las utilidades parciales e interpretación	35
3.2.4. Conclusiones	40
A. Tablas de frecuencias	43
B. Encuesta de sondeo previa	45
C. Código empleado en SAS	47
D. Algoritmo del diseño	51
E. Encuesta	53
Bibliografía	61

Índice de figuras

2.1. Ejemplo de análisis conjunto adaptativo en dos partes	21
3.1. Histogramas de las variables cuantitativas	26
3.2. Resultados de la encuesta	35
3.3. Importancia relativa de los atributos	37
3.4. Importancia relativa de los niveles (los porcentajes negativos indican el porcentaje de una utilidad inferior a la media)	39

Índice de tablas

1.1. Valor añadido bruto y empleo de la cadena forestal-madera en el año 2018	2
1.2. Evolución de la superficie forestal arbolada en Galicia durante la 2ª mitad del siglo XX y principios del siglo XXI	5
2.1. Ejemplo de pregunta de un análisis conjunto basado en elecciones	10
2.2. Combinaciones de atributos y niveles ortogonales hasta 28 perfiles	12
2.3. Matrices ortogonales de cinco diseños diferentes	13
2.4. Cálculo de autovalores para una matriz de diseño	14
2.5. Formato de las respuestas	18
2.6. Ejemplo de un análisis conjunto tradicional. En I se establecen las preferencias entre pares de niveles para las variables marca y vida útil. En II se exponen tres posible perfiles completos para un diseño factorial	20
3.1. Disposición de las primeras diez filas de la matriz de datos	24
3.2. Estadísticos descriptivos de las variables cuantitativas	24
3.3. Tabla de frecuencias de las variables cualitativas <i>Frag</i> y <i>PteModa</i>	25
3.4. Atributos y niveles incluidos en el estudio final	27
3.5. Tamaños de diseños ortogonales	28
3.6. Perfiles del diseño lineal	29
3.7. Frecuencias de los niveles de los atributos y de sus combinaciones	30
3.8. Resultados del diseño de elección con menor Error D tras 100 iteraciones	31
3.9. Matriz de varianzas covarianzas del diseño de elección	32
3.10. Diseño de elección codificado y sin codificar	33
3.11. Resultados de la regresión logística multinomial	36
3.12. Cálculo de la preferencia por nivel del atributo área	38
3.13. 10 perfiles con mayor utilidad del total de 324	38
A.1. Tabla de frecuencias	43

Resumen

Resumen en español

Este trabajo ha tenido por objetivo explorar y cuantificar las preferencias de los profesionales de las principales empresas y asociaciones forestales de Galicia en la elección de una parcela para su explotación forestal, con el fin de desarrollar una herramienta de gestión que permita detectar, por un lado, terrenos con alto potencial forestal donde se consiga una mayor eficiencia en las explotaciones, y por otro, zonas con bajo potencial en las que sería conveniente realizar una reestructuración parcelaria para lograr mayor productividad. Para averiguar sus preferencias realizamos un análisis conjunto basado en elecciones tomando como atributos del modelo 5 variables físicas comunes a todos los terrenos forestales: la superficie del terreno, la forma de la parcela, la fragmentación colindante, la distancia a la vía más cercana y la pendiente del terreno. Mediante técnicas de diseño de experimentos elaboramos una encuesta y la distribuimos entre los profesionales del sector. Sobre las respuestas aplicamos una regresión logística multinomial y obtuvimos las utilidades parciales para cada una de las variables. Los resultados se presentaron detalladamente mediante tablas y gráficos de forma que puedan ser fácilmente interpretables.

English abstract

The objective of this work was to explore and quantify the preferences of professionals from the main forestry companies and associations in Galicia about choosing a plot for logging, in order to develop a management tool that allows detecting, on the one hand, lands with high forestry potential where greater efficiency is achieved, and on the other hand, areas with low potential where it would be convenient to carry out a land restructuration to achieve greater productivity. To find out their preferences, we carried out a choice-based conjoint analysis, taking 5 physical variables common to all forest lands as attributes of the model: the land surface, the shape of the plot, the adjacent fragmentation, the distance to the nearest road and the terrain slope. Using experimental design techniques, we prepared a survey and distributed it among professionals in the sector. On the responses we applied a multinomial logistic regression and obtained the partial utilities for each of the variables. The results were presented in detail using tables and graphs so that they can be easily interpreted.

Capítulo 1

Planteamiento del problema

1.1. Introducción

Galicia es la comunidad autónoma de España que más volumen de cortas de madera genera al año. Concretamente produce el 53,8 % del volumen de madera extraído en el país según el Anuario de Estadística Forestal (AEF) 2019, lo que se traduce en 9.676.134 m^3 de madera con corteza sobre el total de 17.986.742 m^3 extraídos en el territorio nacional. Podemos deducir, por tanto, que Galicia es una de la áreas de mayor peso a nivel forestal de España, pues las cortas de madera junto con el aprovechamiento para leña resultan ser, por lo general, las principales fuentes de producción de recursos de los montes y las que mayor valor de mercado generan. Existen otras motivaciones para la explotación de parcelas forestales distintas a la leña y la madera, como la extracción de corcho y resina o la recolección de castañas, piñones, trufas y demás hongos. Todas estas actividades tienen un peso considerable en el mercado, tanto nacional como internacional. Sin embargo, aún dentro de las actividades silvícolas, no resultan tan provechosas como las madereras, ya que este bien es demandado por numerosas industrias diferentes.

Centrándonos en la economía local, según datos de la Revisión Estadística das Contas Económicas de Galicia 2019, la cadena forestal-madera supuso en el año 2018 el 1,8 % del producto interior bruto (PIB) de la economía gallega, 1.144 millones de euros en términos monetarios, y 24.107 puestos de trabajo, es decir, un 2,1 % del empleo total en Galicia, como podemos observar en la [Tabla 1.1](#). De la cadena forestal-madera forman parte una serie de industrias con diferentes actividades principales, que se desagregan siguiendo la Clasificación Nacional de Actividades (CNAE-09) en cuatro sectores:

- 02 Silvicultura y explotación forestal. Incluye actividades de cultivo de madera, producción de leña y madera en rollo o en bruto, explotación de viveros forestales o la recolección de productos silvestres.
- 16 Industria de la madera y del corcho excepto muebles. Comprende la fabricación de productos madereros con los procesos de producción de serrar, cepillar, conformar, laminar y ensamblar.
- 17 Industria de papel. Incluye la fabricación de pasta de papel, papel y productos de papel transformado.
- 31 Fabricación de muebles. Esta división comprende la fabricación de muebles y productos afines de cualquier material, salvo piedra, hormigón y cerámica.

Podríamos decir que la base sobre la que se asientan las industrias 16, 17 y 31 es la rama número 02, pues es el sector que les suministra la materia prima sin la cual no podrían trabajar.

Tabla 1.1: Valor añadido bruto y empleo de la cadena forestal-madera en el año 2018

Ramas de actividad	Valor añadido bruto (millones)	% sobre el PIB	% sobre el sector	Puestos de trabajo	% sobre el total del empleo
02 Silvicultura y explotación forestal	417	0,7	36,5	9.346	0,8
16 Industria de la madera y el corcho, salvo muebles	324	0,5	28,3	7.982	0,7
17 Industria del papel	247	0,4	21,6	1.904	0,2
31 Fabricación de muebles	156	0,3	13,7	4.875	0,4
Cadena forestal-madera	1.144	1,8	100,0	24.107	2,1
Total economía	62.226	100,0		1.122.719	100,0

Fuente: Instituto Galego de Estatística. Contas económicas de Galicia. Revisión Estadística 2019

El reparto de la propiedad privada de los montes gallegos presenta un marcado carácter histórico minifundista, lo que supone la división de las parcelas en porciones muy numerosas y de pequeño tamaño que se van heredando dentro de las familias, de generación en generación, sin que lleguen a unificarse. En Galicia, según datos del catastro que introduciremos posteriormente, existen aproximadamente 7,3 millones de parcelas forestales, casi tres veces más que el total de la población gallega de 2,7 millones de personas en el año 2019. Estos números revelan la ineficiencia del bosque gallego, ya que explotar una parcela pequeña y aislada no es rentable. Lo conveniente para el progreso de la explotación forestal sería mejorar la ordenación de la propiedad, objetivo que no se ha cumplido desde finales del siglo pasado.

Dentro del sector de la silvicultura y explotación forestal en Galicia, existen grandes empresas que poseen parcelas forestales donde cultivan y extraen madera. Sin embargo, el método de obtención principal, tanto de grandes como de pequeñas empresas, se da a través de la compra de madera a particulares que están en propiedad de terrenos forestales que pueden ser explotados. El negocio comienza cuando algún propietario se pone en contacto con la empresa extractora de madera. Éste le comenta detalles acerca de las características de su parcela a la empresa, como la cantidad de árboles, superficie del terreno, ubicación, etc. En ese momento es decisión de la compañía si le interesa, o no, desplazarse hasta la parcela para talar su madera y posteriormente comprársela. Por ejemplo, en la página web de Ence (gran empresa dedicada a la extracción y transformación de madera en Galicia) detallan que la empresa gestiona en terreno gallego cerca de 10.500 hectáreas de cultivos forestales y que tienen relación con propietarios particulares, suministradores y comunidades de montes a los que compran madera anualmente por valor de 100 millones de euros.

Dentro de este contexto surge el interés del presente trabajo. Visto que muchos de los terrenos gallegos son improductivos y no rentables para la corta por parte de una empresa, sumado al interés del sector de la silvicultura por la compra de madera a particulares, pensamos que sería interesante estudiar las preferencias de estas empresas extractoras de madera a la hora de decidir un posible aprovechamiento forestal. Como indicamos anteriormente, las compañías madereras no acuden a todas

las peticiones de corta, puesto que si el aprovechamiento forestal va a proporcionar nulo o poco rédito, no sería conveniente. El decantarse por una petición de corta viene determinado, principalmente, por las características físicas del terreno. Hemos mencionado que el minifundismo influye en la superficie de la parcela, ya que a menor superficie disponible, menor cantidad de árboles a talar y, por lo tanto, menor cantidad de madera extraída. Pero, además de la superficie, existen también una serie de variables a tener en cuenta, presentes en todas las parcelas forestales, que influyen en la decisión de explotación. En este estudio hemos decidido que las más determinantes para las empresas silvícolas son: la pendiente del terreno, la distancia desde el terreno a la vía más cercana, la forma de la parcela y la fragmentación de las parcelas circundantes, que, junto a la superficie, sumarían cinco variables en total. En un apartado posterior justificaremos más en detalle la selección de estas variables.¹

El objetivo último del trabajo consiste en explorar y cuantificar el sistema de valores de los profesionales en la elección de una parcela para su explotación forestal, asumiendo que esa elección depende de diferentes atributos o factores físicos del terreno. Para llevar a cabo esta tarea de exploración y cuantificación del sistema de valores haremos uso del análisis conjunto, una serie de técnicas analíticas clásicamente empleadas en marketing para estudiar los productos más valorados por los consumidores. Existen numerosas técnicas de análisis conjunto, pero, en general, todas consisten en la elaboración de una encuesta formada por un número de preguntas determinado por el investigador. En cada pregunta (bloque) figuran diferentes versiones (perfiles) del producto a estudiar, donde se detallan una serie de características (atributos) presentes en el producto. Dichos perfiles contienen una descripción de los mismos atributos del producto, pero en diferentes cantidades (niveles). Los conceptos mencionados entre paréntesis en el texto anterior son los tecnicismos empleados en el ámbito del análisis conjunto y serán descritos con más detalle en el [Capítulo 2](#). El objetivo es que el encuestado escoja, ordene o puntúe, en función del tipo de análisis que realicemos, los perfiles para cada pregunta. En base a las respuestas obtenidas estimaremos las preferencias de las personas encuestadas, que serán profesionales del sector maderero.

Generalmente, lo que se busca con el análisis conjunto es conocer las características mejor valoradas por los consumidores de un producto, en definitiva, marketing. Dentro de este ámbito puede emplearse para múltiples tareas tales como decidir precios (en el artículo de Goldberg et al. 1984 se utiliza para poner precio a los productos de acogida en los hoteles), para diseñar un producto (en Dufhues et al. 2004 se empleó para diseñar mejores productos financieros para la población rural del norte de Vietnam) o para el posicionamiento del producto (en DeSarbo et al. 1995 los autores aplican el análisis conjunto basado en elecciones para determinar segmentos del mercado). No obstante, recientemente ha empezado a aplicarse esta metodología en diferentes disciplinas tales como economía ambiental (Baarsma 2003), economía de la salud (Skjoldborg and GyrdHansen 2003) o seguridad alimentaria (Finn and Louviere 1992).

La aplicación de técnicas de análisis conjunto sobre las variables físicas de las parcelas forestales nos permitirá generar un instrumento que cuantifique el grado de explotación forestal de un terreno, es decir, el valor que tiene dicho terreno a la hora de realizar un aprovechamiento forestal en él. Esto permitirá ampliar el conocimiento de los profesionales de la madera sobre las parcelas, favoreciendo la rentabilidad del aprovechamiento. Esta información tiene, además, múltiples aplicaciones que discutiremos una vez obtenidos los resultados, como elaborar un mapa con las zonas de mayor potencialidad forestal o proporcionar a los poderes públicos una herramienta objetiva para la ordenación del territorio.

El trabajo está estructurado en tres capítulos. En el presente [Capítulo 1](#) hemos introducido el proyecto indicando brevemente el interés del mismo, las técnicas que emplearemos y los resultados

¹Que conozcamos, no existe ningún estudio sistemático que acerque conocimiento a las compañías madereras sobre los terrenos más favorables a talar empleando variables físicas.

que esperamos conseguir. A continuación presentaremos el objeto de estudio: las parcelas forestales. Situaremos el marco teórico a nivel nacional y gallego para facilitar su comprensión y delimitar qué son y qué no son. También introduciremos las variables que, a priori, consideramos necesarias para la elaboración del trabajo. En el [Capítulo 2](#) exponemos los fundamentos metodológicos del trabajo, empezando por poner en contexto el análisis conjunto. Posteriormente describiremos sus diferentes utilidades, conceptos y terminología, así como sus principales tipos de aplicación. Cerraremos el capítulo escogiendo y justificando la metodología elegida. Por último, en el [Capítulo 3](#) aplicaremos las técnicas de análisis conjunto paso a paso, detallando y explicando los resultados obtenidos en el estudio de las preferencias de los profesionales del sector para la explotación de las parcelas forestales. Concluiremos el proyecto mostrando los diferentes resultados que hemos obtenido y su posible aplicación a la actualidad del sector silvícola.

1.2. Caso de estudio

Los datos correspondientes a las delimitaciones de las parcelas forestales se han descargado de la Sede Electrónica del Catastro. Luego se ha utilizado el Mapa Forestal Español de máxima actualidad para la comunidad de Galicia (MFE25) para discernir las parcelas forestales de las no forestales. El total de fincas consideradas como forestales es de 7.224.165 incluyendo en esta cifra las parcelas pertenecientes a montes vecinales en mano común.

Según se recoge en el artículo 2 de la Ley 7/2012, de 28 de junio, de montes de Galicia “se entiende por monte o terreno forestal todo terreno en el que vegetan especies forestales arbóreas, arbustivas, de matorral o herbáceas, sea espontáneamente o procedan de siembra o plantación, que cumplan o puedan cumplir funciones ambientales, protectoras, productoras, culturales, sociales o recreativas”.

En general para toda España existen dos clasificaciones diferentes en función de la titularidad: montes privados y públicos. El monte será de titularidad pública si pertenece al Estado, a las comunidades autónomas, a las entidades locales y a otras entidades de derecho público. Mientras que los montes privados pertenecen a personas físicas o jurídicas de derecho privado, ya sea individualmente o en régimen de copropiedad. Sin embargo, en el caso de Galicia y otras comunidades determinadas, existe el concepto de monte vecinal en mano común, mancomunidad de montes o simplemente monte vecinal, que se consideran montes privados, de naturaleza germánica, que pertenecen colectivamente y sin atribución de cuotas a las respectivas comunidades vecinales titulares, estando sujetos a las limitaciones de indivisibilidad, inalienabilidad, imprescriptibilidad e inembargabilidad según el artículo 13 de la ya citada Ley de montes de Galicia.

Conviene distinguir entre montes de titularidad pública y privada, pues a nosotros nos interesan solamente los de titularidad privada. Los montes de titularidad pública siguen un proceso de corta distinto, motivo por el que no se incluirán en el estudio. En el Anuario de Estadística Forestal de Galicia 2019, se indica que el volumen de aprovechamientos forestales en montes públicos sólo representa el 3 % del total, correspondiendo el resto a montes privados, lo cual deja la mayor parte del terreno susceptible de ser valorado según las utilidades que estudiaremos más adelante. Tampoco incluiremos los montes vecinales, ya que presentan características físicas muy diferentes de las parcelas privadas, lo que podría alterar las mediciones y los resultados del estudio. Por lo tanto, de las 7.224.165 parcelas forestales gallegas nos quedaremos con 6.689.822, las parcelas privadas.

Cabe destacar que, de los 2 millones de hectáreas forestales que posee Galicia, solamente están arboladas 1.454.298 hectáreas en total, las restantes 586.456 están desarboladas, lo cual deja a Galicia como la tercera Comunidad de España con mayor porcentaje de superficie forestal respecto a la autonómica, por detrás de Canarias, Asturias y empatada con Cantabria. Se considera Superficie Forestal Arbolada al terreno poblado con especies forestales arbóreas como manifestación vegetal dominante

y cuya fracción de cabida cubierta (FCC) es igual o superior al 10 %. Sin embargo, esta cifra no fue siempre tan positiva. Galicia presentó un importante aumento de la superficie forestal arbolada en la segunda mitad del siglo XX, como se demuestra en la [Tabla 1.2](#). No se poseen datos actualizados para años posteriores al 2011. El AEF publica todos los años una tabla pero la superficie arbolada de Galicia no varía porque la superficie forestal según el AEF se calcula en base al MFE25 y, como ya mencionamos, en Galicia no se actualiza desde diciembre de 2011. Observando la tabla veremos que la superficie forestal arbolada se triplicó a lo largo de 50 años, desde 1950 al 2000 gracias a repoblaciones forestales mediante el cultivo de especies de crecimiento rápido (Molina 1997) y regeneraciones naturales de tierras agrarias abandonadas. Desde el 2000 al 2011 aumentó ligeramente. Cabe mencionar estos datos, pues la actual voluntad de regeneración y expansión del bosque gallego derivará en un mayor número de parcelas en el futuro sobre las que aplicar los resultados de este trabajo.

Tabla 1.2: Evolución de la superficie forestal arbolada en Galicia durante la 2ª mitad del siglo XX y principios del siglo XXI

Período	Superficie (hectáreas)		
	Arbolada	Desarbolada	Forestal
1945 – 1950	538.271	1 552.597	2.090.868
1970 – 1975	1.129.361	863.443	1.992.804
1995 – 2000	1.405.452	634.123	2.039.575
Año			
2011 (IFN4)	1.424.094	606.586	2.030.681

Fuente: Primera revisión del plan forestal de Galicia e IFN4

Es importante señalar que para la realización del estudio no hemos tenido en cuenta la variable referida a las especies arbóreas que se ubican en cada una de las fincas. Actualmente ya se conoce que es una variable decisiva en función del tipo de madera que quiera vender la empresa. Según el AEF 2019 para Galicia, dentro de los 9.442.622 m^3 de madera con corteza de aprovechamientos declarados y solicitados en terrenos privados, 5,6 millones fueron extraídos de eucaliptos (*Eucalyptus spp.*) y 3,4 millones de especies de pinos (*Pinus pinaster*, *Pinus radiata* y *Pinus sylvestris*). El resto de la madera se divide entre otras especies de coníferas y frondosas.

Diferentes estudios socioeconómicos realizados en la comunidad (Crecente et al. 2007) señalan la existencia de múltiples perfiles de propietarios, cada uno con distintos objetivos y criterios de explotación forestal en parcelas de tamaños diversos, prevaleciendo, en mayor medida, fincas de tamaño reducido e ineficaces para una gestión eficiente. Este minifundismo deriva, en el caso de montes familiares o particulares, en una productividad unitaria poco elevada debida al desinterés general por poner a producir la finca, el desconocimiento de técnicas silvícolas apropiadas y la falta de capital necesario para la inversión inicial. Para lidiar con este problema endémico y estructural se han promovido reestructuraciones parcelarias desde la Xunta, como el Plan Forestal de Galicia de 1992, que pretendía mejorar la ordenación de la propiedad y reducir el minifundismo. A pesar de los esfuerzos, la estrategia se ha ido incumpliendo año tras año, invirtiendo más capital en apagar incendios que en reordenar los montes. Esta tendencia histórica de parcelas exageradamente pequeñas generará un gran cúmulo de terrenos con una superficie ínfima en comparación con el resto de parcelas, sobre todo con los montes

vecinales (Balboa López et al. 2013) que cuentan con superficies marcadamente grandes.

1.3. Definición de las variables

Como mencionamos en la introducción, el trabajo consta de cinco variables en base a las cuales estudiaremos las preferencias de los profesionales del sector silvícola. En el [Capítulo 3](#) incidiremos en cómo se han calculado y sus posibles valores, mientras que en esta sección nos limitaremos a definir las variables y justificar su inclusión en el estudio. El conjunto de datos consta de cinco variables físicas que describen las cualidades de cada terreno y otras dos que corresponden a las coordenadas centrales de la parcela y nos servirán más adelante para ubicarlas en el mapa.

1. *Area*. Corresponde a la superficie medida en hectáreas² para cada parcela. La importancia de esta variable es fácil de entender, porque la relación entre el área y los beneficios obtenidos es directa, a mayor área, mayor número de árboles a cortar, por lo tanto más madera. Suponemos que éste será un factor decisivo a la hora de ponderar una finca, ya que es la variable principal por la que preguntan los profesionales del sector. Pronosticamos que tendrá gran relevancia en el estudio.
2. *Forma*. Es un valor que indica la forma del terreno. Seguramente los profesionales de la madera valoren la forma que toma la parcela, ya que en función de ello será más o menos sencillo delimitar qué árboles pertenecen a la finca y cuáles no, además de facilitar el desplazamiento sobre ella si la parcela es más compacta (cuadrada o redonda) en lugar de rectangular y alargada o informe.
3. *Frag*, fragmentación. Representa si el área circundante a una determinada parcela está constituida por otras parcelas de pequeño o de gran tamaño. El interés de esta variable radica en conocer el entorno que rodea a cada parcela, por ejemplo, si llaman a la compañía para que corte madera en una parcela de área pequeña y dicho terreno está rodeado por otras fincas de tamaño pequeño, el trabajador podría considerar solicitar a los dueños de estas fincas adyacentes cortar y comprarles también a ellos su madera para compensar en mayor medida el desplazamiento a la zona. Sin embargo, cuando las fincas circundantes sean grandes, habrá menos posibilidades de que quieran venderles su madera, por lo que el trabajador, en última instancia, decidirá que el trabajo de talar y comprar tan poca madera no compensa el gasto en recursos por desplazarse a la zona.
4. *PteModa*, moda de la pendiente. Es la moda de la pendiente de la parcela, es decir, la pendiente más común del terreno delimitado por la parcela. Está expresada en porcentaje, por ejemplo, un valor del 20 % indica que por cada 100 metros de avance en horizontal se ascenderían 20 metros en vertical. El personal que trabaja cortando madera tiene en cuenta la pendiente del terreno, pues cuanto mayor sea ésta, más complejo resultará el trabajo sobre esa superficie a la hora de, por ejemplo, cortar con la motosierra, ya que es importante tener los pies fijos en el suelo y una parcela escarpada entorpece la operación. Además, podría imposibilitar el uso de maquinaria en la finca por falta de potencia para superar el grado de inclinación. En definitiva, conocer la pendiente es importante, ya que un relieve inclinado encarece y ralentiza el aprovechamiento.
5. *DistViasMe*, distancia media a vías. Es la distancia media (desde todos los puntos de una parcela) a la vía más cercana desde la que transportar la madera extraída, medida en metros. La implicación de esta variable es clara, ya que determina los costes de transporte una vez talada la madera, puesto que a mayor distancia, más se encarece el transporte de la madera por el terreno forestal, que como explicamos con la pendiente, puede llegar a complicar la tala.
6. *X*. Coordenada del punto central de la parcela sobre el eje x.
7. *Y*. Coordenada del punto central de la parcela sobre el eje y.

²La hectárea es una unidad de medida de superficie. 1 hectárea equivale a 10.000 m².

Capítulo 2

Análisis conjunto

2.1. Presentación

Dentro del mercado hay muchas decisiones interdependientes que se ven involucradas a la hora de crear una estrategia de marketing para una marca, ya sea de productos o servicios. Estas decisiones influyen no solo sobre las características del producto, sino también sobre su posicionamiento, comunicación, distribución y precio. Las decisiones deben tomarse en base a posibles reacciones de la competencia dentro de un mercado (muchas veces impredecible) en constante cambio. Si una empresa quiere conseguir el éxito, el proceso de toma de decisiones debe tener una comprensión clara de la mente de sus clientes y por qué eligen un producto u otro. Para ello, es imprescindible adoptar una visión global del problema, es decir, tomar en cuenta las diferentes alternativas de mercado y el conjunto de sus atributos, ya que, generalmente, los productos sobre los que los consumidores deciden son complejos, por lo tanto, no se toman decisiones en base a criterios unitarios, sino que se evalúan diversos atributos que en la mente de los clientes generan un estímulo determinado. Para conseguir esto debemos analizar el proceso de toma de decisiones en el mercado de manera multifactorial, lo que nos lleva a plantear la siguiente pregunta: ¿qué estrategias de decisión utiliza el consumidor para elegir un producto determinado?

Tratando de responder esta pregunta empezaron a surgir estudios en la década de los 60, dentro del campo de la investigación de mercados, que buscaban importar esta nueva visión del consumidor. En el año 1964 se inició una novedosa corriente con el artículo de Luce, un psicólogo matemático, y Tukey, un estadístico (Luce y Tukey 1964), que introdujeron el concepto de *conjoint measurements* o medidas conjuntas. Para poder comprender esta teoría primero debemos explicar tres conceptos básicos dentro de la terminología del análisis conjunto que son: los atributos, los niveles y los perfiles. Los atributos son variables que sirven para describir cierto producto, los niveles son los posibles valores que puede tomar un atributo y los perfiles son combinaciones concretas de los niveles de cada atributo para un producto. Tomando un teléfono móvil como ejemplo, podemos decir que dos de sus atributos son la marca y el precio, los niveles del atributo marca podrían ser {Apple, Xiaomi, Samsung} y del atributo precio {400€, 500€, 600€}, por último, ejemplos de perfiles podrían ser un móvil Apple de precio 400€ o un Samsung de precio 600€. Lo que pretende la teoría de medidas conjuntas es evaluar distintas combinaciones de niveles entre los atributos y con ello determinar una puntuación o medida para cada nivel o combinación de niveles. La puntuación puede ser categórica, ordinal o escalada. Siguiendo con el ejemplo del teléfono móvil, si le preguntamos a un individuo su opinión sobre un móvil de marca Apple y precio 400€ podría dar una respuesta categórica (e.g. barato para ser un Apple), ordinal (e.g. muy bueno, bueno, malo o muy malo) o escalada (e.g. 7/10 en escala de calidad precio). Si a ese mismo individuo se le pide que evalúe otra serie de perfiles, con los resultados podríamos estimar su función de utilidad para los teléfonos móviles. Podemos entender el concepto de función de utilidad

como una representación numérica de la satisfacción que obtienen los consumidores cuando disfrutan de una determinada cantidad de un producto o servicio.

Al tratar de implementar estos sistemas de medición conjunta en temas comerciales surgieron numerosos problemas prácticos, principalmente debidos a dificultades en los test¹ al tratar de comprobar si los datos recogidos satisfacían los axiomas del modelo. Estos procedimientos test requerían una cantidad ingente de datos y era complicado que se cumpliesen los axiomas incluso con un número bajo de encuestados. La metodología que surgió para solucionar este tipo de problemas se conoce como “análisis conjunto” (AC). En 1971 se publicó el primer artículo sobre análisis conjunto detallado y orientado hacia el consumidor en el ámbito del marketing (Green y Rao, 1971). Desde entonces, el AC es, probablemente, el enfoque más popular empleado para medir las preferencias de los clientes dentro del campo de la investigación de mercados. Está bastante relacionado con el campo de la *Information Integration Theory* (IIT)² y su sistema de mediciones funcionales.

Según Green y Srinivasan (1978) podemos definir el AC como una metodología de carácter descomposicional, en la que los encuestados valoran diferentes alternativas o perfiles a partir de los cuales se obtienen sus preferencias, con el fin de poder emprender, con mejores garantías, cualquier tipo de decisión con respecto al producto estudiado. Aunque no se considere análisis conjunto, la medición de preferencias también puede abordarse mediante un enfoque composicional (Fishbein 1967). La diferencia principal radica en que el análisis descomposicional es una aproximación *top-down*, es decir, partiendo de la evaluación final se “descompone” ésta en las utilidades relativas de los diferentes atributos, mientras que el análisis composicional es *bottom-up*, donde el primer paso consiste en preguntar a los sujetos cuál es la importancia que otorgan a cada atributo por separado.

Desde la incorporación de las técnicas de AC al mundo de la mercadotecnia, se ha ido desarrollando nueva metodología. A día de hoy, principalmente, existen cuatro métodos de análisis conjunto:

1. Análisis conjunto tradicional (conocido simplemente como AC).
2. Análisis conjunto basado en elecciones (ACBE).
3. Análisis conjunto adaptativo (ACA).
4. Análisis conjunto autoexplicativo.

Los tres primeros son técnicas descomposicionales, mientras que el AC autoexplicativo es composicional.

Como señalamos en la introducción, las técnicas de análisis conjunto pueden aplicarse en numerosos campos aparte del marketing gracias a su versatilidad. En Rao (2014) se identifican cinco aplicaciones fundamentales, no excluyentes entre sí, propias de los AC:

1. Es una técnica de medición que cuantifica las preferencias y determina los atributos más valorados por los compradores en ciertos productos o servicios;
2. Es una técnica analítica que sirve para predecir las reacciones de los consumidores ante nuevos productos o servicios;
3. Es una técnica de segmentación para la identificación de grupos de compradores con preferencias por intercambios similares;

¹Ver Corstjens y Gautschi (1983) para entrar en detalle sobre los métodos de test.

²Es una teoría propuesta por Norman H. Anderson (Anderson 1970) que describe y modela cómo las personas recopilan información de una serie de fuentes para luego desarrollar un juicio general.

4. Es una técnica de simulación que permite evaluar las reacciones del mercado competitivo ante nuevos productos o servicios; y
5. Es una técnica de optimización que busca maximizar la rentabilidad obtenida a través de productos o servicios.

A día de hoy, multitud de técnicas de análisis conjunto son cada vez más usadas en campos como las ciencias sociales o la salud. Por citar estudios recientes, en Ong (2022) emplean técnicas tradicionales de análisis conjunto usando el método de puntuación en una escala Likert del 1 al 7 para determinar las preferencias de los estudiantes de instituto de las Islas Filipinas acerca del sistema educativo online, que podría encuadrarse dentro de la primera, y más común, aplicación, diseñar un producto que vaya a ser valorado positivamente por sus futuros consumidores. En Igarashi (2022) emplean el análisis conjunto basado en elecciones para estudiar las preferencias de los ciudadanos ante una hipotética vacuna, lo que encajaría dentro de la aplicación 2, predecir la reacción de los habitantes japoneses ante un nuevo elemento (la vacuna). En Tariq (2018) se utiliza el análisis conjunto adaptativo para obtener la opinión de veterinarios sobre cuáles son los factores de riesgo que motivan la propagación de la gripe aviar en granjas avícolas de Pakistán. En Ramantswana (2020) emplean el análisis conjunto autoexplicativo para determinar las preferencias por la ubicación de la sede central de una empresa encuestando a CEOs de empresas que cotizan en la Bolsa de Valores de Johannesburgo.

2.2. Elementos y fases del análisis conjunto basado en elecciones

Todo análisis conjunto parte del interés por conocer en detalle las preferencias de los usuarios o consumidores de un determinado producto o servicio. Sin embargo, la metodología empleada varía en función del tipo de análisis conjunto. En este trabajo nos centraremos en el análisis conjunto basado en elecciones (ACBE), ya que fue la técnica escogida para llevar a cabo el caso práctico.

Una vez determinado el producto que queremos investigar, debemos valorar qué atributos y niveles emplear para el análisis. A partir de este punto es cuando empiezan a diferir las metodologías de los distintos tipos de AC. Concretamente, cuando tengamos atributos y niveles determinados, el ACBE consistirá en la elaboración de una encuesta conformada por un número de preguntas (también llamadas “bloques de alternativas” en el lenguaje del análisis conjunto) a establecer por el investigador, en las que se presentarán diferentes perfiles, es decir, combinaciones de los niveles de los atributos, buscando obtener la máxima información posible de los encuestados. El carácter de las preguntas y las posibles respuestas también son decididas por el investigador en función del producto a estudiar y del perfil de los consumidores. Basándose en la respuestas obtenidas se hará una estimación de la función de utilidad de los individuos. En el análisis conjunto es de suma importancia el concepto de utilidades parciales. Estas utilidades miden la influencia de cada uno de los niveles de los diferentes atributos en el proceso de toma de decisión por parte del encuestado. Conociendo la función de utilidad de los compradores de un producto pueden tomarse numerosas decisiones, por ejemplo, emplear el atributo mejor valorado de tu producto como estandarte en las campañas publicitarias, potenciar aquellos atributos peor valorados, cuantificar la mejora de los niveles de un producto para comprobar su rentabilidad, etc.

Antes de entrar en la teoría vamos a mostrar un ejemplo de pregunta de un ACBE recogido en Green y Srinivasan (1978) que nos ayudará a entender la estructura del cuestionario y su finalidad. Se plantea un mercado de neumáticos donde queremos estudiar los hábitos de compra de los consumidores, tomando como atributos la vida útil de los neumáticos (en km), la marca, el color lateral de los neumáticos y el precio.

En la [Tabla 2.1](#) se plantea la pregunta “¿Cuál escogería de las siguientes opciones?” El encuestado debe elegir uno de los tres perfiles que, como vemos, presentan niveles diferentes en algunos atributos e iguales en otros. Un ACBE completo consiste en un conjunto de cuestiones como esta, en este caso el

Tabla 2.1: Ejemplo de pregunta de un análisis conjunto basado en elecciones

¿Cuál escogería entre las siguientes opciones?

Atributos	Niveles de los atributos		
	Perfil 1	Perfil 2	Perfil 3
Marca	Sears	Firestone	Goodrich
Vida útil	50.000	40.000	50.000
Color	Blanco	Gris	Rojo
Precio	55	50	55

bloque (pregunta) contiene 3 alternativas o perfiles. Sin embargo, es en el diseño previo de la encuesta donde reside la verdadera dificultad del análisis conjunto, debemos decidir el número de preguntas, el número de alternativas para cada una y qué perfiles mostrar.

Tras darle contexto a la aparición del análisis conjunto y entendida su finalidad, podemos proceder a explicar cómo se elabora. Hemos visto que es posible aplicar estas técnicas analíticas a productos de todo tipo y con diversos fines, aún así, los pasos a seguir para su creación son siempre los mismos. Todo AC puede elaborarse en 3 pasos: (1) diseño de la encuesta, (2) recogida de los datos y (3) estimación de las utilidades parciales. Como hemos dicho, las tres etapas son comunes a todos los tipos de análisis conjunto, pero las metodologías varían. En este estudio nos centraremos en explicar los pasos siguiendo la metodología de un ACBE. En la [Sección 2.3](#) hablaremos sobre los otros tres tipos de análisis mencionados y en la [Sección 2.4](#) justificaremos la elección de la técnica escogida.

2.2.1. Diseño de la encuesta

Es la primera parte del análisis y la más importante, pues en función de las decisiones que se tomen los resultados pueden variar ampliamente. Esta etapa se divide en cuatro puntos: definición del problema, identificación de atributos, determinación de niveles y, por último, determinación de perfiles y estructura de la encuesta.

Definición del problema

Naturalmente, cualquier análisis conjunto empieza por definir el problema que se nos plantea y qué se pretende con los resultados, porque no siempre existirá un objetivo claro. Por ejemplo, imaginemos que una empresa quiere estudiar el diseño de un nuevo producto, con la salvedad de que dicha empresa ya tiene a la venta un artículo en esa línea de productos. En este caso, el problema principal para el investigador no es sólo determinar los mejores atributos del nuevo producto, sino que también debe tener en cuenta el número de ventas que este nuevo producto sustraerá del anterior. En definitiva, los investigadores deben asegurar, previamente, que el estudio proporcionará los resultados deseados.

Identificación de atributos

Es fundamental escoger correctamente los atributos, ya que la función de utilidad que obtendremos nos permitirá conocer las preferencias de los consumidores sobre esas características en concreto y no otras. Los atributos pueden ser de dos tipos: cualitativos y cuantitativos. Siguiendo con el ejemplo

previo de los teléfonos móviles, la marca sería un atributo cualitativo, ya que expresa una cualidad (Apple, Xiaomi, Samsung) o característica observable del objeto, mientras que el precio sería un atributo cuantitativo, pues no puede asociarse a una cualidad, sino que es un factor cuantificable que toma valores numéricos discretos (400€, 500 €, 600€) o en intervalo [400€; 500€), [500€; 600€) o [600€; 700€). No existe un criterio objetivo que permita seleccionar los atributos más favorables para el estudio, deben escogerse en función del problema que se plantee.

Determinación de niveles

Una vez determinados los atributos, deben decidirse sus niveles. Antes de continuar, es importante razonar que cuantos más atributos y niveles haya, mayor será la carga de información a la que deberemos someter a los encuestados. Entonces, o bien sacrificamos información a cambio de hacer un test más corto para los encuestados, o bien hacemos un test largo que nos permita recabar mucha información, pero sobrecargue a los individuos, con el riesgo de que nos den respuestas sin pensar para terminar la encuesta cuanto antes. Con esto queremos dejar claro que la elección de niveles y atributos debe sopesarse con cuidado.

Al igual que pasa con los atributos, no existe un criterio objetivo para determinar el número de niveles óptimo ni sus valores. En Wittink et al. (1989) los investigadores encontraron una regularidad empírica con respecto al efecto producido por diferentes números de niveles entre atributos. La conclusión principal es que los atributos con mayor número de niveles reciben, sistemáticamente, mayor importancia que aquellos con menos niveles. Por eso se considera recomendable diseñar estudios con prácticamente el mismo número de niveles para todos los atributos. En los artículos que hemos consultado para elaborar este estudio y que hemos ido referenciando, lo más frecuente fue elegir una cantidad de entre 2 y 5 niveles, común a todos los atributos.

Otra decisión a tomar serán los valores para cada nivel. En un atributo cualitativo la elección es relativamente más sencilla, por ejemplo, al estudiar la salida de un nuevo producto al mercado se escogerán aquellos valores que se plantean aplicar en el artículo, por ejemplo, para el atributo color de un teléfono móvil se plantean los siguientes valores {rojo, verde, negro}. Aquel nivel que obtenga una utilidad relativa mayor por parte de los encuestados será el elegido en el producto final. Sin embargo, cuando se trata de atributos cuantitativos, la decisión se complica, puesto que las posibilidades son infinitas. La metodología común consiste en escoger aquellos niveles comunes en el mercado y quizá añadir un nivel innovador propuesto por la compañía para comprobar su aceptación. También pueden basarse en los valores elegidos por estudios previos o realizar una pequeña encuesta, antes del AC, que permita identificar los valores sugeridos o esperados por los consumidores.

Determinación de perfiles y estructura de la encuesta

Habiendo decidido los atributos del producto y sus posibles niveles, podemos proceder a la creación de la encuesta así como los perfiles que presentaremos a los consumidores en cada bloque, o pregunta, para su evaluación. Recordemos que un perfil es una combinación concreta de niveles, pero ¿cómo se crean esas combinaciones? A diferencia de los atributos y niveles, sí existen criterios objetivos para la evaluar la eficiencia de la encuesta y la determinación de perfiles. Para establecer la estructura que seguirá la encuesta se recurre al diseño experimental, esto son, técnicas estadísticas que permiten identificar y cuantificar las causas de un efecto dentro de un estudio experimental. El diseño experimental de un análisis conjunto basado en elecciones se representa mediante una matriz, donde los atributos constituyen las columnas y las filas los perfiles del producto. Las filas se dividen en una serie de bloques (preguntas de la encuesta), cada uno formado por el mismo número de alternativas.

El primer paso del diseño consiste en seleccionar el número de perfiles de la encuesta, para lo cual será necesario conocer el total de perfiles posibles. Este número se calcula multiplicando los niveles de

todos los atributos. Volviendo sobre el ejemplo de Green y Srinivasan (1978) relativo a las preferencias de compra de neumáticos, teníamos cuatro atributos: *marca* y *color* mostraban tres niveles diferentes; *vida útil* y *precio* solo dos. Por lo tanto, el total de perfiles es $3 \times 3 \times 2 \times 2 = 36$. La combinación de perfiles posible para cuatro atributos con ese número de niveles es 36. Resulta una cantidad pequeña, ya que podríamos elaborar una encuesta con doce preguntas (tres perfiles por pregunta), una cantidad fácil de manejar por los encuestados. Cuando una encuesta consta del total de perfiles se conoce como diseño factorial completo. En la práctica no suele darse este caso, lo normal es tener que seleccionar una pequeña parte del total de perfiles, generando un diseño factorial fraccionado. Si al caso anterior añadimos una quinta variable cualquiera de cuatro niveles, el número de combinaciones total sería 144. Si quisiésemos elaborar una encuesta con todos los perfiles y tres alternativas por pregunta habría 48 preguntas, una cantidad demasiado alta de información a la que someter a los encuestados. En estos casos se opta por reducir el número de preguntas para no saturar a los individuos a costa de la pérdida de información.

Ahora, la pregunta que cabe hacerse es ¿cuántos perfiles utilizar? Para responder debemos introducir dos medidas fundamentales del diseño de experimentos que son la ortogonalidad y el balance. Diremos que un diseño está balanceado cuando cada nivel se presenta el mismo número de veces dentro de cada atributo en el total de preguntas, mientras que un diseño es ortogonal cuando cada pareja de niveles se presenta el mismo número de veces dentro de cada par de atributos en el total de preguntas. Si la matriz del diseño cumple ambas características se llamará matriz ortogonal y se dice que es 100 % eficiente y óptima. Estas matrices cumplen con propiedades muy deseables. La primera sería la parsimonia, esto son, modelos simples con la capacidad de explicar los datos mediante el número mínimo de parámetros. Segundo, permiten estimar todos los efectos principales del análisis conjunto. Por último, presentan una gran capacidad predictora. Sin embargo, son pocos y específicos los diseños capaces de cumplir ambas propiedades, pues dependerá del número de atributos y de niveles. Un diseño factorial completo siempre será 100 % eficiente y óptimo, pero es difícil que de conseguir. En la [Tabla 2.2](#) se muestran las combinaciones de atributos y niveles con posibilidad de matriz de diseño ortogonal. El número de la izquierda indica la cantidad de perfiles, mientras que los de la derecha representan el número de niveles elevado a la cantidad de atributos con esos niveles.

Tabla 2.2: Combinaciones de atributos y niveles ortogonales hasta 28 perfiles

4	2^3	12	2^{11}	16	2^{15}	18	$2^1 3^7$	21	$3^1 7^1$	24	2^{23}	25	5^6
6	$2^1 3^1$	12	$2^4 3^1$	16	$2^{12} 4^1$	18	$2^1 9^1$	22	$2^1 11^1$	24	$2^{20} 4^1$	26	$2^1 13^1$
8	2^7	12	$2^2 6^1$	16	$2^9 4^2$	18	$3^6 6^1$			24	$2^{16} 3^1$	27	3^{13}
8	$2^4 4^1$	12	$3^1 4^1$	16	$2^8 8^1$	20	2^{19}			24	$2^{14} 6^1$	27	$3^9 9^1$
9	3^4	14	$2^1 7^1$	16	$2^6 4^3$	20	$2^8 5^1$			24	$2^{13} 3^1 4^1$	28	2^{27}
10	$2^1 5^1$	15	$3^1 5^1$	16	$2^3 4^4$	20	$2^2 10^1$			24	$2^{12} 12^1$	28	$2^{12} 7^1$
				16	4^5	20	$4^1 5^1$			24	$2^{11} 4^1 6^1$	28	$2^2 14^1$
										24	$3^1 8^1$	28	$4^1 7^1$

Por ejemplo, el primer diseño en la lista (2^3) tiene tres atributos de dos niveles en cuatro perfiles y el segundo diseño ($2^1 3^1$) tiene un atributo de dos niveles y un atributo de tres niveles en seis perfiles. En la [Tabla 2.3](#) viene representada la matriz de los cinco primeros diseños. Cada fila representa un perfil y cada columna un atributo. Estos diseños están balanceados, puesto que cada nivel de cada atributo figura el mismo número de veces dentro de ese atributo. En la [Tabla 2.3](#) diseño $2^4 4^1$ el primer atributo (primera columna) presenta cuatro veces el valor 1 y cuatro veces el valor 2. Lo mismo ocurre en los siguientes tres atributos. Por último, el atributo de cuatro niveles presenta cada uno de ellos dos veces.

La ortogonalidad es un poco más complicada de ver. Primero, debemos dividir el diseño en parejas

Tabla 2.3: Matrices ortogonales de cinco diseños diferentes

2^3			$2^1 3^1$		2^7							$2^4 4^1$					3^4			
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	2	1	2	2	1	2	1	2	1	2	1	2	1	2	2	1	2	3	2
1	2	2	1	3	1	2	2	1	1	2	2	1	1	2	2	3	1	3	2	3
2	2	1	2	1	2	2	1	1	2	2	1	1	2	2	1	4	2	2	2	1
			2	2	1	1	1	2	2	2	2	2	2	2	2	1	2	3	1	2
			2	3	2	1	2	2	1	2	1	2	1	2	1	2	2	1	3	3
					1	2	2	2	2	1	1	2	2	1	1	3	3	3	3	1
					2	2	1	2	1	1	2	2	1	1	2	4	3	1	2	2
																	3	2	1	3

de atributos. El diseño 2^7 tiene 21 pares de atributos. Si identificamos cada atributo (columna) del 1 al 7 (en orden de izquierda a derecha) las 21 combinaciones serían: 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 2-3, 2-4, 2-5, 2-6, 2-7, 3-4, 3-5, 3-6, 3-7, 4-5, 4-6, 4-7, 5-6, 5-7 y 6-7. Ahora, dentro de cada uno de los 21 pares de atributos, hay cuatro posibles combinaciones de niveles: nivel 1—nivel 1, nivel 1—nivel 2, nivel 2—nivel 1 y nivel 2—nivel 2. Si nos fijamos en las dos primeras columnas veremos que estos pares de niveles se repiten todos dos veces. Lo mismo ocurre para las restantes 20 parejas de atributos. El diseño es por tanto ortogonal y, como ya vimos, balanceado, lo que significa que es 100 % eficiente y óptimo (cuando apliquemos el ACBE en el [Capítulo 3](#) se entenderán con mayor claridad ambas propiedades). La ortogonalidad implica que los niveles de los diferentes atributos se escogen independientemente unos de otros, por lo tanto, el efecto provocado por los niveles de cada atributo (utilidad parcial) es independiente del resto de efectos. A continuación presentaremos otro concepto fundamental en el diseño de encuestas, la eficiencia.

Para entender el concepto de diseño eficiente debemos presentar los autovalores o valores propios. Según Peña y Prieto (2003) un autovalor es una medida básica del tamaño de una matriz, que no se ve alterada si se produce un cambio de coordenadas equivalente a una rotación de los ejes. Las medidas globales de tamaño de una matriz, como el determinante o la traza, son una función de los autovalores, por tanto, también son invariantes ante transformaciones que preservan los autovalores. Los vectores propios representan las direcciones características de la matriz y no son invariantes. Llamaremos vectores propios de una matriz cuadrada de orden n a aquellos vectores cuya dirección no se modifica al transformarlos mediante la matriz. Por tanto u es un vector propio de la matriz A si verifica que $Au = \lambda u$, donde λ es un escalar, que se denomina valor propio de la matriz. Se supone $u \neq 0$, ya que si no es trivialmente cierta. Para calcular el vector propio podemos escribir la ecuación anterior como $(A - \lambda I)u = 0$ dando lugar a un sistema homogéneo de ecuaciones que tendrá solución no nula si, y solo si, la matriz del sistema, $(A - \lambda I)$, es singular. Por tanto, este sistema tiene solución no nula si se verifica

$$|A - \lambda I| = 0. \quad (2.1)$$

En la [Tabla 2.4](#) se calculan los valores propios partiendo de un diseño. Pensemos en el modelo de análisis conjunto como en un modelo lineal de forma $\hat{y} = X\hat{\beta}$ donde y contiene las valoraciones de los encuestados, X es la matriz de diseño codificada y $\hat{\beta}$ es el vector de utilidades parciales que queremos estimar. Entonces, volviendo sobre la [Tabla 2.4](#), la primera columna de la matriz codificada X representa el intercepto del modelo, donde el valor 2 se ha reemplazado por -1 . La traza de una matriz es la suma de los elementos diagonales de la matriz, la cual, por propiedades de los autovalores, es 0,75 tanto para $(X'X)^{-1}$ como para la matriz de autovalores de la inversa Λ , mientras que el determinante de $(X'X)^{-1}$ coincide con el producto de los autovalores y es 0,0009766.

Tabla 2.4: Cálculo de autovalores para una matriz de diseño

Diseño	X						$X'X$				$(X'X)^{-1}$				Λ			
1 1 1	1	1	1	1	1	1	6	0	-2	0	0,1875	0	0,0625	0	1/4	0	0	0
1 2 2	1	1	-1	-1	1	1	0	6	0	-2	0	0,1875	0	0,0625	0	1/4	0	0
1 2 2	1	1	-1	-1	1	1	-2	0	6	0	0,0625	0	0,1875	0	0	0	1/8	0
2 1 2	1	-1	1	-1	1	1	0	-2	0	6	0	0,0625	0	0,1875	0	0	0	1/8
2 2 1	1	-1	-1	1	1	1												
2 2 1	1	-1	-1	1	1	1												

Dada una cantidad de autovalores, generalmente queremos crear un solo número que resuma los valores del conjunto. La forma más obvia de hacerlo es mediante la media aritmética, un estadístico muy popular y útil, pero no es la única forma de promediar números. La menos conocida media geométrica se halla multiplicando entre sí p números y calculando la raíz p -ésima del producto. Para entenderlo mejor, consideremos el siguiente ejemplo. Digamos que nuestras inversiones se incrementaron un 7, 5 y 12% en un período de tres años. La media aritmética $(7 + 5 + 12)/3 = 8\%$ no representa el incremento medio que habrían tenido nuestros ingresos si hubiesen aumentado la misma cantidad año a año. Para encontrar esa cantidad necesitamos la media geométrica $(1,07 \times 1,05 \times 1,12)^{1/3} = 1,0796$. El incremento medio fue de 7,96%.

Presentados estos conceptos, estamos en condiciones de hablar sobre eficiencia. Las mediciones de eficiencia típicas de una matriz de diseño X de orden $N_D \times p$ se basan en la matriz de información $X'X$. Por ejemplo, el diseño de la [Tabla 2.4](#) corresponde a un problema de 3 atributos con 2 niveles cada uno en el que se consideran 6 perfiles. De modo que $N_D = n^0$ perfiles = 6 y $p = 3 + 1$ (n^0 atributos más intercepto). La matriz de varianzas covarianzas del vector $\hat{\beta}$ de parámetros estimados por mínimos cuadrados es proporcional a $(X'X)^{-1}$, concretamente es igual a $\sigma^2(X'X)^{-1}$, pero la varianza (σ^2) es una constante desconocida, y al ser constante puede ser ignorada o asumir que es igual a uno. Los elementos diagonales de $(X'X)^{-1}$ son las varianzas de los parámetros estimados. Un diseño eficiente tendrá una matriz de varianzas covarianzas “pequeña” medida a través de los autovalores de $(X'X)^{-1}$. El proceso de minimizar los autovalores o varianzas solo depende de la selección de perfiles de X y no del parámetro desconocido σ^2 .

Las principales medidas de eficiencia son dos, y se basan en cuantificar el “tamaño” de la matriz promediando los autovalores de las varianzas. La eficiencia A es una función de la media aritmética de los autovalores y se calcula como la traza de $(X'X)^{-1}/p$. Es quizá la medida más obvia de eficiencia, en el sentido de que a medida que las varianzas se hacen más pequeñas, la media aritmética de las varianzas también disminuye, por lo tanto aumenta la eficiencia. La segunda medida se conoce como eficiencia D, y se calcula como la media geométrica de los autovalores, es decir, $|(X'X)^{-1}|^{1/p}$. Esta medida es más fácil y rápida de optimizar computacionalmente que la eficiencia A. Existe una tercera medida llamada eficiencia G que se basa en σ_M , la máxima desviación típica de la predicción. Los tres criterios son funciones convexas de los autovalores de $(X'X)^{-1}$, por lo que están bastante correlados. Si el diseño está balanceado y es ortogonal los tres criterios tendrán eficiencia máxima. Por otro lado, cuanto más eficiente sea un diseño, más balanceado y ortogonal será. Un diseño será ortogonal y balanceado cuando se cumple que la diagonal de $(X'X)^{-1}$ es igual a $\frac{1}{N_D}I$ para la matriz codificada X . El diseño será ortogonal cuando la submatriz de $(X'X)^{-1}$, excluyendo la fila y columna del intercepto, sea diagonal. El diseño estará balanceado cuando todos los elementos fuera de la diagonal en la fila y columna del intercepto sean 0. La matriz X que escojamos se determinará la eficiencia del diseño. Idealmente queremos escoger aquella matriz X que sea ortogonal y balanceada o al menos esté muy cerca de serlo, es decir, queremos maximizar la eficiencia.

Las medidas de eficiencia se pueden escalar de 0 a 100 de la siguiente manera

$$\begin{aligned} \text{Eficiencia A relativa} &= 100 \times \frac{1}{N_D \operatorname{tr}((X'X)^{-1})/p} \\ \text{Eficiencia D relativa} &= 100 \times \frac{1}{N_D |(X'X)^{-1}|^{1/p}} \\ \text{Eficiencia G relativa} &= 100 \times \frac{\sqrt{p/N_D}}{\sigma_M}. \end{aligned} \quad (2.2)$$

Estas eficiencias miden la bondad de ajuste de un diseño en relación a un hipotético diseño ortogonal que podría no existir, así que no son útiles como medidas absolutas. En su lugar, deben emplearse relativamente para comparar la eficiencia de dos diseños diferentes, pero en la misma situación. Cuando la eficiencia D es 0 significa que uno o más parámetros no pueden ser estimados. Cuando es 100, el diseño será balanceado y ortogonal. Los valores intermedios implican que se pueden estimar todos los parámetros, pero con una precisión inferior a la óptima.

La metodología que emplearemos en este trabajo consistirá en construir un modelo lineal general en base al cual generar el modelo de elección. En Kuhfeld (2010) se indica que este razonamiento lo han seguido investigadores del campo del marketing durante años con buenos resultados. Asumen que un diseño lineal eficiente conllevará un buen diseño multinomial logístico, que será el que empleemos para estimar nuestro modelo de elección. Al construir diseños con este procedimiento conseguiremos modelos con máxima eficiencia y a su vez la independencia de los atributos (garantizada por la ortogonalidad del modelo). Este modelo multinomial logístico se basa en la premisa de que un individuo siempre escogerá aquella alternativa que le reporte mayor utilidad. Volviendo sobre este concepto, habíamos dicho que la utilidad para un consumidor representa numéricamente su nivel de satisfacción cuando disfruta de un determinado producto. Cuando a un individuo se le plantea una pregunta como la de la [Tabla 2.1](#), entra en juego su función de utilidad. Para explicar el proceso de elección de los individuos entre el conjunto de alternativas (perfiles) disponibles en cada bloque, se emplea la teoría de utilidades aleatorias, *random utility models* (RUMs) desarrollada en McFadden (1974). Esta teoría asume que la función de utilidad se divide en dos componentes, uno determinístico y uno aleatorio. El componente determinístico representa los atributos observables y se expresa mediante variables explicativas en el modelo, mientras que el componente aleatorio recoge todos los atributos y factores no observables del proceso de elección (gustos personales del individuo particular), que son las que dan nombre a la teoría. La utilidad es, por tanto, aleatoria, por lo que no existe un modelo que pueda predecir con total certeza la elección de un individuo.

En Rao (2014) se considera el problema de construir un modelo que determine la utilidad percibida por un individuo al escoger un perfil particular entre un conjunto de alternativas de un bloque S de n alternativas. Siendo \tilde{u}_k la utilidad aleatoria obtenida por un individuo para el perfil k ($\in S$), podría descomponerse de la siguiente manera

$$\tilde{u}_k = v_k + \epsilon_k, \quad (2.3)$$

donde v_k es el componente determinístico y ϵ_k es el componente aleatorio.

Basándonos en la premisa de que el componente determinístico de la utilidad se deriva de los atributos presentes en los perfiles, podemos escribir v_k como

$$v_k = \sum_{j \in T} b_{jk} x_{jk}, \quad (2.4)$$

donde T es el número de atributos incluidos en el modelo, x_{jk} es el nivel para el k -ésimo perfil del j -ésimo atributo y b_{jk} es peso asociado al nivel de x_{jk} .

La función es lineal en parámetros (coeficientes b). Además, es general y puede acoger funciones no lineales de las variables x . Para la modelización del componente aleatorio utilizaremos el modelo multinomial, que se basa en la asunción de que los componentes aleatorios (ϵ_k) , asociados con los perfiles de un bloque, son independientes e idénticamente distribuidos (*iid*) según una distribución de Gumbel. La función de densidad para esta distribución es

$$P(\epsilon_k \leq \epsilon) = e^{-e^{-\epsilon}}, \quad -\infty < \epsilon < \infty. \quad (2.5)$$

La forma de esta distribución fija la media en 0,575 y la varianza en 1,622 ó $\frac{\pi}{6}$, siendo ambas medidas adimensionales.

En base a lo visto hasta ahora podemos desarrollar una función que describa la probabilidad de que un individuo elija un perfil concreto. Si escoge el perfil k , implica que la utilidad aleatoria asociada a k es la más alta (entre las alternativas disponibles en el bloque). Denotaremos con y_j el perfil escogido, por lo tanto y_j tomará el valor 1 si es escogido el perfil j -ésimo y 0 en caso contrario. Entonces

$$\begin{aligned} P[y_j = 1] &= P[\tilde{u}_j \geq \tilde{u}_m; m \in S] \\ &= P[v_j + \epsilon_j \geq v_m + \epsilon_m; m \in S] \\ &= P[\epsilon_j - \epsilon_m \geq v_m - v_j; m \in S]. \end{aligned} \quad (2.6)$$

ó

$$P[y_j = 1] = P[\epsilon_m \leq v_j - v_m + \epsilon_j; m \in S].$$

La expresión 2.6 puede ser reescrita en función de las asunciones hechas sobre la distribución del componente aleatorio, de forma que la solución al problema es:

$$P[y_j = 1] = \frac{e^{v_j}}{\sum_{m \in S} e^{v_m}} = \frac{1}{1 + \sum e^{v_m - v_j}} \quad (2.7)$$

La Ecuación 2.7 representa el modelo de elección logístico multinomial. La probabilidad de escoger un perfil determinado depende de la utilidad determinística de los perfiles a valorar. Además, la probabilidad de elegir cualquier perfil es proporcional al exponente de su utilidad determinística.

Escrito en función de los atributos, la probabilidad de elección para el perfil k -ésimo (denotado como P_j) es:

$$P_j = \frac{1}{1 + \sum_{m \in S} e^{v_m - v_j}}; \quad \text{ó} \quad (2.8)$$

$$P_j = \frac{1}{1 + \sum_{m \in S} e^{\sum_{k \in T} b_{jm}(x_{jm} - x_{kj})}} \quad (2.9)$$

Los parámetros desconocidos en 2.9 son los coeficientes b . Para comprender el método de estimación de los coeficientes b que utilizaremos en este trabajo replantearemos el problema anterior con matrices. Partimos de la utilidad RUM descompuesta en el componente determinístico más el componente aleatorio

$$u = x_i \beta + e, \quad (2.10)$$

donde x_i es un vector de atributos que caracterizan a la alternativa i , β es un vector columna de K pesos asociados a dichos atributos, y e es la utilidad aleatoria. Suponemos que hay N bloques

de preguntas, C_n , indexados por $n = 1, 2, \dots, N$, donde cada bloque está constituido por los perfiles $C_n = \{x_{1n}, \dots, x_{J_n n}\}$. Pasando directamente al equivalente matricial de la [Ecuación 2.7](#) tenemos

$$P_{in}(X_n, \beta) = \frac{e^{x_{in}\beta}}{\sum_{j=1}^{J_n} e^{x_{jn}\beta}}, \quad (2.11)$$

donde X_n es una matriz que consta de J_n vectores fila, cada uno caracterizando al perfil x_{jn} . La unión vertical de las X_n matrices forma la matriz de diseño X . El objetivo consiste en estimar el parámetro β que maximice la verosimilitud de los datos. Bajo condiciones generales, el estimador de máxima verosimilitud es consistente y asintóticamente normal con la matriz de covarianzas

$$\Sigma = (Z'PZ)^{-1} = \left[\sum_{n=1}^N \sum_{j=1}^{J_n} z'_{jn} P_{jn} z_{jn} \right]^{-1} \quad (2.12)$$

$$\text{donde } z_{jn} = x_{jn} - \sum_{i=1}^{J_n} x_{in} P_{in}.$$

La [Ecuación 2.11](#) revela propiedades importantes sobre los modelos de elección (no lineales). Mientras que los modelos lineales están centrados con respecto al total de los perfiles, los no lineales están centrados en base a cada bloque de alternativas. Esto demuestra que en los diseños de elección, tanto la selección de perfiles como la asignación de los mismos a cada bloque, afecta a la matriz de covarianzas. Es más, en modelos lineales, la matriz de covarianzas no depende del valor del vector de parámetros, pero en los modelos de elección las probabilidades P_{jn} son funciones de β , y por lo tanto de la matriz de covarianzas. Lo que suele hacerse en los diseños del modelo de elección es asumir inicialmente que $\beta = 0$ porque simplifica el diseño del problema. Una vez se obtengan las respuestas y se estimen las utilidades parciales se demostrarán con los pertinentes test que la hipótesis inicial era falsa para cada uno de los niveles de los atributos, o no.

El objetivo de la matriz de diseño será, entonces, definir un conjunto de bloques, dado el valor anticipado de β , que minimicen el “tamaño” de la matriz de covarianzas, Σ definida en la [Ecuación 2.11](#). Como medida del tamaño del error emplearemos la eficiencia D, por tanto, trataremos de minimizar

$$\text{Error D} = |\Sigma|^{1/K}, \quad (2.13)$$

siendo K el número de parámetros.

2.2.2. Recogida de datos

En esta sección se procede a la recopilación de datos. Para ello, lo primero será elaborar la presentación de la encuesta obtenida en el apartado anterior de la forma más intuitiva y visual posible para los individuos encuestados. Lo primero será ofrecer al individuo una breve presentación del carácter y objetivo de la encuesta para, a continuación, dar paso al cuestionario. Esta etapa del análisis variará bastante en función de la técnica escogida, pues, como veremos más adelante, las encuestas siguen un formato distinto. En el caso de los ACBE, cada pregunta suele ir numerada y precedida por la cuestión pertinente, como en la [Tabla 2.1](#), y quizá acompañada por una imagen que represente a cada opción.

La encuesta podrá realizarse en persona u online. Una ventaja de realizar el cuestionario presencialmente es que conlleva mayor compromiso por parte de los participantes, con lo que se conseguirán resultados más fieles y detallados. Sin embargo, este método implica mayores costes en desplazamiento y tiempo que enviar un formulario online, sobre todo en el caso de no poder reunir a los participantes en un mismo lugar, teniendo que visitarlos uno a uno.

2.2.3. Estimación de las utilidades parciales

Los datos recogidos en la encuesta deben organizarse de una determinada manera para poder aplicar la regresión logística multinomial. El formato que seguirán se muestra en la [Tabla 2.5](#), en la que se presenta un número de filas equivalente a la cantidad de individuos que han participado en la encuesta multiplicado por el número de perfiles presentes en la encuesta. Lo que significa que, para cada encuestado, habrá una fila por cada perfil de la encuesta, lo haya escogido dentro de la pregunta correspondiente o no. Las primeras tres columnas de la tabla identifican el individuo, el bloque o pregunta y el perfil del que se trata. La columna y señala si, dentro de una pregunta, ese perfil ha sido elegido por el individuo. De ser así se señala con un 1, en caso contrario con un 0, por lo tanto habrá un único perfil con valor 1 para cada pregunta respondida por un encuestado. Los subíndices de Y_{111} indican por orden: el individuo, el bloque y el perfil. Siendo s el número total de perfiles del bloque por pregunta. La última columna indica los niveles de los atributos de cada perfil, donde p es el número total de niveles para los atributos codificados. Que un atributo esté codificado implica que se ha creado una variable ficticia para todos sus niveles salvo para uno, que será el nivel de referencia. Al ser variables ficticias tomarán el valor 1 si el perfil presenta ese nivel y 0 si no presenta ese nivel. Los subíndices x_{11} señalan por orden: el perfil y el nivel al que hace referencia.

Tabla 2.5: Formato de las respuestas

Individuo	Bloque	Perfil	Valor y para la alternativa ($y = 1$ o 0)	Variables X X_1, \dots, X_p
1	1	1	Y_{111}	$x_{11} \dots x_{1p}$
		2	Y_{112}	$x_{21} \dots x_{2p}$
		3	Y_{113}	$x_{31} \dots x_{3p}$

	1	s	Y_{11s}	$x_{s1} \dots x_{sp}$
1	2	1	Y_{111}	$x_{11} \dots x_{1p}$
		2	Y_{112}	$x_{21} \dots x_{2p}$
		3	Y_{113}	$x_{31} \dots x_{3p}$

	2	s	Y_{11s}	$x_{s1} \dots x_{sp}$

El método de estimación que mejor se ajusta para calibrar un modelo logístico a nivel individual es el de máxima verosimilitud. Asumamos que han respondido a la encuesta N individuos, que la elección de perfiles de la i -ésima persona se denota como (y_{i1}, \dots, y_{is}) , que S_i es el conjunto de perfiles respondido por la i -ésima persona y que cada y es igual a 1 o 0 dependiendo de si ha elegido ese perfil o no. El principio de estimación consiste en determinar los valores de los parámetros del modelo que maximicen la probabilidad (verosimilitud) de los datos observados. Para el i -ésimo individuo y el bloque de perfiles $S_i = \{1, 2, \dots, s\}$, la verosimilitud de las elecciones observadas es

$$L_i = \prod_{m=1}^S P_{im}^{y_{im}}. \quad (2.14)$$

La verosimilitud total sería $L = \prod_{i=1}^N L_i$. L es una función de los parámetros β desconocidos $L = L(\beta_1, \dots, \beta_T)$, siendo T el número de atributos.

2.3. Otros tipos de análisis conjunto

Haremos una breve presentación del resto de tipos de análisis conjunto para mostrar sus principales motivaciones y justificar nuestra elección del análisis conjunto basado en elecciones.

2.3.1. Análisis conjunto tradicional (AC)

El análisis conjunto tradicional tiene dos maneras de recoger las preferencias de los consumidores. La primera, y menos usual, se produce mediante la comparación de todas las posibles combinaciones de pares de niveles de los atributos del producto. La tarea del encuestado es ordenar estos pares de niveles para cada combinación de atributos. Mientras que la segunda emplea perfiles completos, en lugar de hacerlo de dos en dos como en el sistema anterior. Estos perfiles están formados por una serie de atributos que se consideran decisivos para el consumidor a la hora de distinguir entre un producto u otro. En función del número de atributos y de sus niveles existirá un mayor o menor número de combinaciones. Si al encuestado se le pregunta por todos los perfiles estaremos ante un diseño factorial completo. Sin embargo, no es lo más común, ya que suele trabajarse con demasiadas combinaciones como para someter al individuo a tantas preguntas. Generalmente se utiliza un diseño factorial fraccionado, es decir, no se plantean el total de las combinaciones, sólo una parte.

Pongamos el ejemplo recogido en Green y Srinivasan (1978). Supongamos que queremos estudiar los hábitos de compra de los consumidores en el mercado de los neumáticos. En él se toman como atributos importantes la vida útil de los neumáticos, la marca, el color lateral de los neumáticos y el precio. Si planteásemos un análisis conjunto por pares de atributos deberíamos comparar uno por uno los pares: vida útil-marca, vida útil-color, vida útil-precio, marca-color, marca-precio y color-precio. En la [Tabla 2.2](#), parte izquierda (I), podemos observar las decisiones por orden de preferencia de un individuo con respecto al par vida útil-marca. Para estimar el modelo completo habría que analizar también los cinco pares restantes. Por otro lado, en la parte derecha de la tabla (II), figuran tres posibles perfiles de un diseño factorial. Si quisiésemos elaborar un diseño factorial completo habría que tener en cuenta todas las combinaciones, que se calculan como el productorio de los niveles de cada variable. En este caso la variable marca tiene 4 niveles, mientras que vida útil, color y precio tienen 3, por lo tanto, existirían $4 \cdot 3 \cdot 3 \cdot 3 = 108$ posibles combinaciones, cantidad considerable de perfiles que haría plantear la posibilidad de un diseño factorial fraccionado.

Una vez decidida la forma de presentación de los estímulos hay que decidir una manera de cuantificar las preferencias. Las dos más empleadas son la ordenación, poniendo cada uno de los perfiles en orden (medida no métrica) de preferencia, o la puntuación (métrica), siguiendo una escala tipo Likert mediante la cual se asigne un valor escalar a cada perfil. Una vez que los individuos hayan respondido a la encuesta emplearíamos técnicas de regresión para obtener sus preferencias generales, las cuales son descompuestas en valores separados y compatibles entre sí correspondientes a cada atributo.

2.3.2. Análisis conjunto adaptativo (ACA)

El ACA es un método de análisis adaptativo, que fue primeramente introducido en 1985 dentro del software Sawtooth. Durante la década de los 80 fue el método más empleado, pero a día de hoy su uso no está muy extendido (Sawtooth declara que solamente el 2% de los usuarios lo emplean en sus proyectos personales). Sin embargo, presenta ventajas para ciertas investigaciones. Se le llama “adaptativo” porque se personaliza para cada individuo. Se diseñó pensando en situaciones donde el número de atributos excedía la cantidad plausible empleada en los AC tradicionales, por lo que este método se centra en aquellos atributos que resultan más relevantes para el encuestado, evitando así la sobrecarga de información.

Tabla 2.6: Ejemplo de un análisis conjunto tradicional. En I se establecen las preferencias entre pares de niveles para las variables marca y vida útil. En II se exponen tres posible perfiles completos para un diseño factorial

I. Pares de niveles entre marca y vida del neumático				II. Tres perfiles completos			
Marca del neumático	Vida del neumático (millas)			Marca	Perfil 1	Perfil 2	Perfil 3
	30.000	40.000	50.000				
Goodyear	8º	4º	1º		Sears	Firestone	Goodrich
Goodrich	12º	9º	5º	Vida útil	50.000	40.000	50.000
Firestone	11º	7º	3º	Color lateral	Blanco	Gris	Rojo
Sears	10º	6º	2º	Precio (\$)	55	50	55

Por ejemplo, si nos planteásemos un AC tradicional de perfiles completos para estudiar un escenario con 12 atributos y cuatro niveles por atributo, necesitaríamos, como norma general, el triple de perfiles que de parámetros a estimar, en este caso supondría someter a los encuestados a 111 perfiles, cada uno descrito por 12 atributos, lo cual es una tarea demasiado complicada para los encuestados. El ACA resuelve este problema personalizando las encuestas, de manera que a cada individuo se le pregunta en detalle por aquellos atributos que más valora. La encuesta tiene dos partes, la primera consiste en un proceso autoexplicativo donde se determina qué niveles y qué atributos son más valorados por el encuestado clasificándolos mediante escalas Likert, generalmente valores del 1 al 7 que van desde muy indeseable hasta muy deseable. Después, en función de las elecciones que tome la persona que conduzca la encuesta, la entrevista se centrará en aquellos atributos a los que la persona dio mayor importancia y en los niveles de atributos más relevantes. Suele llevarse a cabo por medio de perfiles parciales, donde los niveles y atributos no deseados se descartan. Debido a estas dos partes tan diferenciadas del proceso (la primera composicional y la segunda descomposicional), suele conocerse como “híbrida” a esta técnica de análisis conjunto. Ambas partes deben administrarse por ordenador, de otra manera es imposible calcular el modelo rápidamente para elaborar la segunda parte de la encuesta.

A continuación, mostramos el ejemplo recogido en Sawtooth (2007) para ilustrar esta metodología. En este caso, queremos estudiar el mercado de los ordenadores. Tenemos 7 variables a estudiar: marca, disco duro y software, con cuatro niveles cada una; velocidad del procesador, memoria RAM y tamaño del monitor, con tres niveles; y teclado con dos niveles. Si utilizamos técnicas como el AC tradicional o el ACBE podríamos incurrir en una sobrecarga de información al entrevistado porque sería incapaz de tener en cuenta las 7 variables a la hora de tomar una decisión, lo que provocaría una pérdida de información. Sería conveniente emplear un modelo ACA para determinar cuáles de esos atributos son más importantes. En la [Figura 2.1](#), parte superior izquierda, indicamos una de las siete preguntas que se harían para valorar los niveles de cada atributo. En este caso ejemplificamos cuál sería la pregunta en función de la variable marca.

Una vez conocidos los niveles, se procede a preguntar por los atributos. Sorprenderá que la pregunta para evaluar las preferencias por atributos no es: “Puntúa las siguientes variables en función de su interés por ellas”, y a continuación el listado de los 7 atributos sometidos a juicio según los valores

Figura 2.1: Ejemplo de análisis conjunto adaptativo en dos partes

PRIMERA PARTE

Preferencias por niveles

Indique cuánto interés tiene en las siguientes marcas de ordenador

	Ninguno	Poco	Algo	Mucho
Dell	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
HP	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lenovo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gateway	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Preferencias por atributos

Si el resto de características de estos dos ordenadores fuesen idénticas, ¿cuánta importancia tendría esta diferencia para ti?

	Ninguna	Poca	Algo	Mucha
Dell en lugar de Gateway	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SEGUNDA PARTE

Si el resto de características de estos dos ordenadores fuesen idénticas, ¿cuál escogerías?

<u>Ordenador 1</u>	<u>Ordenador 2</u>
Lenovo	HP
1 GB RAM	2 GB RAM
3,5 GHz	2,66 GHz

Izq.	Neutro	Dcha.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

en la escala Likert. Siguiendo la lógica de la empresa Sawtooth indicada en sus artículos³ técnicos, la importancia de un atributo es claramente dependiente de la diferencia entre las magnitudes con las que se valore. Si todos los billetes de tren para viajar desde la ciudad A hasta la ciudad B costasen entre 100 y 105€, el precio no sería importante a la hora de decidir la compañía de viajes, pero si el precio variase entre 100 y 1000€, seguramente el precio sería un factor decisivo. Por tanto, indican que la mejor manera de medir la importancia de cada atributo es comprobar si a los consumidores les importa decidir entre uno u otro nivel. En el ejemplo preguntamos cómo de importante resulta la diferencia entre las marcas Dell y Gateway.

Por último, procedemos con la segunda parte de la encuesta. En esta sección se evalúan las preferencias por uno u otro perfil en función de los atributos más importantes para cada individuo deducidos en la primera parte. En el ejemplo comprobamos que los atributos más valorados por la persona fueron la marca, la memoria RAM y la velocidad del procesador, cuyos niveles más destacados pueden observarse en la tabla.

2.3.3. Análisis conjunto autoexplicativo

A diferencia de los anteriores métodos, el análisis conjunto autoexplicativo es composicional, aunque cabe destacar que la primera parte del ACA está compuesto por una encuesta de tipo autoexplicativo. La metodología composicional implica que, en lugar de calcular las preferencias de los sujetos mediante perfiles (ya sean completos o parciales), se pide a los individuos que evalúen cada uno de los niveles de todos los atributos, y que posteriormente evalúen la importancia relativa de cada atributo. Una vez hecho esto, se calculan las preferencias en base a una suma ponderada de los intereses mostrados

³Ver Sawtooth (2007) para mejor comprensión y ejemplo completo.

por cada nivel específico de los atributos, donde las ponderaciones se corresponden con la importancia relativa dada a los atributos.

2.4. Selección de la técnica de estudio

En esta sección explicaremos por qué decidimos aplicar el ACBE en este trabajo. De primeras, en nuestro estudio tratamos con ocho variables o menos, dependiendo de si finalmente empleamos todas o no, por lo tanto, no merece la pena usar un ACA, cuyo uso está enfocado a cantidades muy grandes de atributos.

Con respecto al análisis autoexplicativo, en Sattler y Hensel-Borner (2007) realizan una comparación entre dicho método y los análisis descomposicionales. Señalan que las principales ventajas del modelo autoexplicativo son los menores costes en cuanto al diseño de la encuesta y la recolección de los datos, además, resulta más sencillo el análisis de los mismos y las preguntas suelen ser más fáciles de responder para los encuestados. Sin embargo, los modelos descomposicionales (ACBE) cuentan con una mayor similitud respecto a las decisiones reales que se llevan a cabo en el mercado, ya que se comparan perfiles completos (teniendo en cuenta todo el conjunto) y no atributo a atributo, que sería una situación menos realista. Otro punto es que los métodos descomposicionales generan una mejor aproximación a las utilidades parciales de los atributos, debido al tipo de preguntas que se realizan en los análisis autoexplicativos, donde suele preguntarse: “¿Cómo de importante es este atributo para ti?”, generando que cada encuestado interprete de manera diferente el funcionamiento del sistema de puntuación de la pregunta. Una tercera desventaja frente a las técnicas descomposicionales es que las autoexplicativas pueden estar sesgadas por respuestas socialmente aceptadas, por ejemplo, ante la pregunta: “¿Cómo de importante es para ti el precio del producto X ?”, podrían darle una menor importancia de la que realmente estiman para demostrar que el dinero no es lo importante, mientras que las aproximaciones descomposicionales son más consistentes en este aspecto.

En definitiva, no aplicaremos un método composicional en el estudio debido a que, en la teoría, los modelos descomposicionales presentan más ventajas, especialmente en términos de validez de sus predicciones.

Por último, debemos decidir la técnica descomposicional. Como introdujimos en la sección anterior, hay dos métodos principales: el análisis conjunto tradicional y el análisis conjunto basado en elecciones. El AC tradicional presenta una mayor complejidad cognitiva para los encuestados a la hora de realizar evaluaciones y comparaciones entre múltiples perfiles. Por un lado, puedes perder la noción de la escala que emplees para puntuar o no ser consistente en su uso, por otro, supondría una carga de trabajo mental grande si el objetivo consiste en ordenar un número de perfiles elevado. Estas tareas se simplifican en el ACBE, donde sencillamente tienes que elegir entre dos o tres perfiles en cada pregunta, de manera muy similar al mercado.

En (Karniouchina et al., 2009) se realiza un metaanálisis sobre estudios que compararon previamente ambas técnicas, señalando que desde la incorporación de modelos bayesianos jerárquicos para la estimación del análisis conjunto, ha aumentado mucho el uso del ACBE. La conclusión final a la que llegan es que los parámetros estimados por ambas técnicas suelen ser muy parecidos. Aún siendo válidas ambas metodologías, indican que el ACBE presenta una mayor validación a nivel individual que el AC tradicional, además de presentar mayor precisión en las predicciones. Por eso, dejando a un lado los resultados francamente similares, recomiendan el uso de técnicas de análisis conjunto basadas en elecciones.

Capítulo 3

Aplicación del análisis conjunto basado en elecciones

En este capítulo describiremos el proceso seguido para el cálculo del análisis conjunto basado en elecciones aplicado sobre nuestro conjunto de datos. Consistirá en ir detallando paso a paso las etapas que vimos en el [Capítulo 2](#), explicando por qué se ha hecho así e indicando los resultados obtenidos.

3.1. Descripción de los datos

Antes de empezar a aplicar las técnicas de análisis conjunto, conviene conocer cómo se distribuyen las variables para entender las magnitudes con las que vamos a trabajar y cómo se han calculado. En el [Capítulo 1](#) presentamos las variables y su interés, ahora detallaremos cómo se han calculado y cuáles son sus posibles valores.

- *PteModa*. La moda de una variable es aquel valor que más se repite. En una variable discreta es fácil de calcular, pero la pendiente es continua, por lo que resulta imposible que dos valores se repitan. Para su cálculo decidimos emplear intervalos, de forma que el intervalo que más observaciones reúna dentro de una parcela será su moda. Como detallaremos más adelante en este capítulo, realizamos una encuesta a los profesionales del sector maderero y mediante sus respuestas determinamos los intervalos en función de sus preferencias de explotación de las fincas. Los intervalos son $(0;20]$, $(20;40]$ y >40 . Si la parcela tiene una moda de la pendiente entre $(0;20]$ tomará el valor 1, si está entre $(20;40]$ tomará el valor 2 y si es mayor de 40 tomará el valor 3.
- *DistViasMe*, medida en metros. Se calculó como la distancia media desde todos los puntos de la parcela a la vía útil más cercana. Dichas vías vienen recogidas en la Base Topográfica de Galicia 2020, y están conformadas por carreteras, pistas, caminos y vías urbanas. Se excluyen autopistas y autovías porque no pueden servir de cargadero de camión ni sendero.
- *Frag* es una variable politómica, por lo que puede tomar 3 valores diferentes: 1 si el conjunto de parcelas circundantes son de gran tamaño (menos fragmentadas), 2 para parcelas de tamaño medio y 3 para parcelas de pequeño tamaño que, por lo tanto, están muy fragmentadas.
- *Area*. Superficie de la parcela medida en hectáreas.
- *Forma*. Es un índice calculado como

$$Forma = \frac{4 \cdot \pi \cdot area}{perimetro^2},$$

y pretende asignar un valor al perímetro de la parcela en función de su forma. Hemos decidido usar un índice que pondera con un mayor valor las fincas con forma redonda. Toma valores entre 0 y 1, por lo tanto el valor 1 corresponderá a una parcela perfectamente circular.

En la [Tabla 3.1](#) mostramos las primeras diez filas de la matriz de datos para mostrar cómo se ve la base de datos con la que vamos a trabajar.

Tabla 3.1: Disposición de las primeras diez filas de la matriz de datos

	Area	Forma	Frag	DistViasme	PteMo	X	Y
1	0,33	0,71	2	25,95	1	561344,10	4787227,20
2	0,09	0,72	2	19,66	1	556337,00	4784731,70
3	0,08	0,68	2	23,31	1	556527,60	4782189,00
4	0,10	0,64	3	66,27	1	556240,80	4784682,80
5	0,40	0,76	3	24,91	1	556578,10	4782086,00
6	0,21	0,53	3	39,12	1	556307,90	4784669,40
7	0,37	0,46	3	21,25	3	555910,60	4781965,60
8	7,85	0,42	2	34,68	1	555969,30	4784701,80
9	0,18	0,45	3	47,15	1	556090,20	4784527,20
10	0,46	0,77	2	41,27	1	556006,10	4784473,10

3.1.1. Estadísticos descriptivos

Primero mostraremos las principales medidas de posición y dispersión de las variables cuantitativas. No las calculamos para las variables correspondientes a las coordenadas porque no es necesario estudiarlas, así como tampoco se incluyen *PteModa* ni *Frag*, pues son variables cualitativas.

Tabla 3.2: Estadísticos descriptivos de las variables cuantitativas

	Area (ha)	Forma	DistViasMe(m)
Mínimo	0,0000	0,0000	0,00
Q_1	0,0293	0,3100	26,37
Mediana	0,0688	0,4910	59,92
Media	0,2153	0,4705	84,06
Q_3	0,1711	0,6480	116,90
Máximo	2135,3901	1,0000	2317,73
Desv. típ.	2,1562	0,2091	80,1341

Fijémonos en la variable área de la [Tabla 3.2](#). Lo primero de lo que nos damos cuenta es del enorme valor máximo que presenta comparado con el resto de sus medidas. Sin embargo, la medida que más llama la atención es la mediana, que toma un valor de 0,0688 ha, lo que significa que la mitad de las 6.689.822 parcelas privadas tienen menos de 0,0688 hectáreas, lo que equivaldría a una superficie

cuadrada de 26,22 metros. Estos valores señalan el marcado carácter minifundista del terreno forestal gallego. También, cabe señalar que la media triplica a la mediana, siendo esta última una medida más robusta, lo que implica que hay parcelas exageradamente grandes comparadas con las demás. Por otro lado, la variable forma presenta una media muy cercana a 0,5. En la [Subsección 3.1.2](#) veremos más en detalle su distribución. Por último, la distancia media vemos que no toma un valor máximo tan exagerado como la media. De las tres, área parece ser la variable con mayor dispersión debido a su alta desviación típica respecto a la media.

Para las variables *PteModa* y *Frag* mostramos la cantidad de datos asociados a cada uno de sus valores en la [Tabla 3.3](#). Por un lado, observamos que la mayor parte de las parcelas presentan una pendiente muy pronunciada en gran parte de su terreno y que, debido al carácter minifundista del predio gallego, lo más común es que las parcelas circundantes estén muy fragmentadas.

Tabla 3.3: Tabla de frecuencias de las variables cualitativas *Frag* y *PteModa*

Frag			PteModa		
1	2	3	1	2	3
227.536	2.331.491	4.130.376	634.050	1.194.141	4.861.631
3,40 %	34,85 %	61,75 %	9,47 %	17,85 %	72,68 %

3.1.2. Histogramas

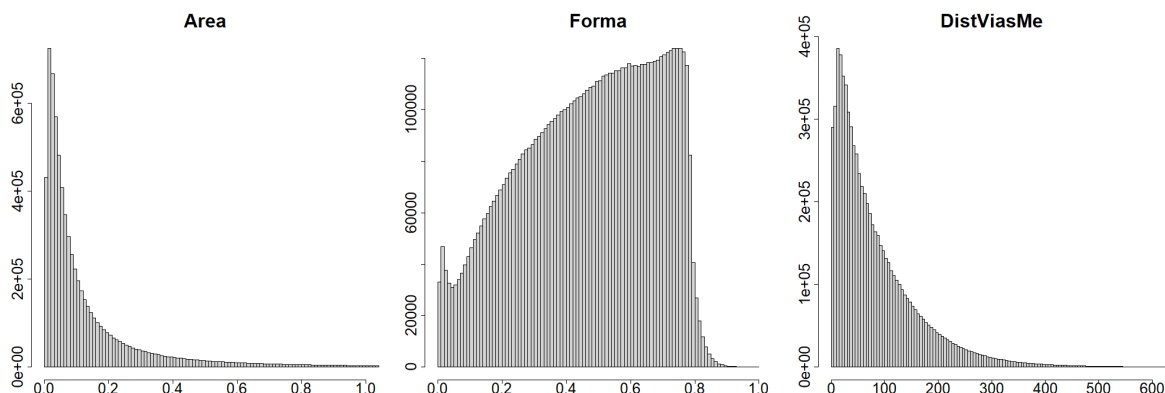
En la [Figura 3.1](#) mostramos los histogramas de las variables continuas para tratar de visualizar la distribución de los datos. Debemos aclarar que los histogramas están acotados, porque si mostramos todo el rango de valores hasta el máximo no se observaría adecuadamente la forma del histograma.

Si nos fijamos en área veremos que sus observaciones están muy concentradas en los primeros valores del histograma. De hecho, gracias a la [Tabla A.1](#) del primer Apéndice, donde se muestran las frecuencias absolutas y relativas de las variables, veremos que en el intervalo $[0;0,5)$ están el 92,11 % de las observaciones, clara muestra del minifundismo gallego. Solamente un 3,11 % de las parcelas forestales superan 1 hectárea de superficie.

Por otra parte, podemos apreciar que la variable forma presenta una tendencia creciente en cuanto a frecuencia de observaciones hasta el valor 0,8, punto en el que empieza a descender y se hace prácticamente nulo para valores superiores. Esto indica que hay muy pocas parcelas con forma de círculo perfecto, lo cual tiene sentido, ya que, por lo general, adoptan una forma rectangular.

Con respecto a la distancia media vemos que presenta una distribución muy parecida a área, pero con una pendiente más suavizada. Tampoco tiene tanta concentración de datos en torno a un valor concreto como sí ocurre con área, donde el intervalo $[0;0,5)$ acumula 6.161.776 observaciones, mientras que la distancia media no alcanza esa cifra hasta los 200 metros. Cabe aclarar que todos estos datos depende de los intervalos tomados en cuenta, pero como explicaremos más adelante, el que tomemos estos intervalos tiene su razón de ser en decisiones de los profesionales de la madera.

Figura 3.1: Histogramas de las variables cuantitativas



3.2. Aplicación del ACBE

Ya hemos planteado el problema, introducido el objeto de estudio, escogido la técnica de trabajo y presentado cómo se distribuyen las variables con las que vamos a trabajar. Ahora, aplicaremos las técnicas del ACBE a las parcelas forestales.

3.2.1. Diseño de la encuesta

Identificación y elección de atributos y niveles

La primera etapa en la aplicación de un análisis conjunto, independientemente del tipo, es la identificación del problema. En nuestro caso, nos planteamos cuáles son las preferencias de los profesionales dedicados a la corta de madera en Galicia a la hora de decidir si realizar un aprovechamiento sobre una parcela forestal o no. Ya que tenemos a nuestra disposición las variables sobre el total de las parcelas forestales en territorio gallego, decidimos emplear técnicas de análisis conjunto para detectar esas preferencias. El siguiente paso consiste en determinar aquellos atributos (variables) que influyan en las decisiones de los consumidores a la hora de elegir. Este es un paso que realizamos previamente a la redacción del trabajo. En un principio barajamos 9 variables de las cuales hemos descartado 4 y nos hemos quedado con las 5 que consideramos más representativas.

El siguiente paso consiste en determinar los niveles de los atributos. Esta es una etapa crucial dentro del diseño del AC, pues escoger entre unos niveles u otros puede alterar con facilidad los resultados del estudio. No existe un criterio objetivo para calcular cuál es el número óptimo de niveles ni los puntos de corte para cada atributo. Puesto que no disponemos de ningún punto de referencia sobre el que basarnos, decidimos elaborar y plantear una encuesta a los agentes implicados en la compra de madera dentro del mercado gallego, y determinar los distintos niveles en los que categorizar las variables según sus respuestas. No se incluyeron preguntas sobre los atributos forma y fragmentación, puesto que son factores poco intuitivos para los encuestados y las preguntas serían difíciles de plantear e interpretar, ya que cada uno las valoraría según sus propios criterios, que no casarían con los que estamos siguiendo en esta investigación.

La encuesta fue respondida por 60 personas y las preguntas planteadas figuran en el [Apéndice B](#). En base a las respuestas recogidas, determinamos que lo óptimo sería considerar cuatro niveles diferentes para la variable área y tres niveles tanto para distancia como para pendiente. Esto es así, mayormente, a causa de la gran varianza que presentan las superficies de las parcelas gallegas, rasgo no tan exagerado, como ya vimos en la [Tabla 3.2](#), en las variables distancia y pendiente. Si redujésemos los niveles de área

a tres, podría generarse una pérdida de información. Para representar cada nivel decidimos trabajar con intervalos en lugar de valores discretos.

Por otro lado, la variable fragmentación ya está categorizada en tres niveles diferentes desde su planteamiento. Sin embargo, no podemos presentar los valores 1, 2 y 3 como indicativos, porque los sujetos encuestados no entenderían la información. Por eso, el nivel 1 se corresponderá con fragmentación baja, el 2 con fragmentación media y el 3 con fragmentación alta. Por último, las categorías del atributo forma fueron decididas siguiendo un criterio técnico, concretamente empleando el primer cuartil 0,3 y el tercer cuartil 0,6, aproximadamente, como umbrales. Decidimos dividir la variable en tres niveles al igual que las demás (salvo área), por tanto el primer intervalo sería $[0;0,3)$ forma muy irregular, el segundo $[0,3;0,6)$ forma irregular y el tercero $[0,6;1]$ forma regular.

En la primera columna de la [Tabla 3.4](#) pueden observarse los atributos incluidos en el estudio, y en la primera fila los valores de los niveles codificados. El único atributo con cuatro niveles, como ya hemos comentado, es el área. Entrecomillado después de cada atributo se indican las unidades de las variables métricas.

Tabla 3.4: Atributos y niveles incluidos en el estudio final

	1	2	3	4
Área (ha)	<0,5	0,5–1	1–5	>5
Forma	Muy irregular	Irregular	Regular	
Fragmentación	Baja	Media	Alta	
Distancia (m)	>500	100–500	<100	
Pendiente (%)	>40	20–40	<20	

Determinación de perfiles y estructura de la encuesta

Para la elaboración de la encuesta, es preciso determinar el número de perfiles que vamos a utilizar como alternativas, así como el número de preguntas a plantear. Como vimos en la [Subsección 2.2.1](#), el primer paso del diseño consiste en la selección de perfiles. Para ello, debemos conocer el número de total de posibles combinaciones. Esto se calcula como el productorio de los niveles de los atributos $4 \times 3 \times 3 \times 3 \times 3$ que da un total 324 perfiles. Es un número demasiado alto como para realizar un diseño factorial completo, ya que, por ejemplo, si utilizásemos tres perfiles como alternativas en cada pregunta, habría que plantear 108 preguntas, un excesivo volumen de cuestiones para los encuestados. Por lo tanto, llevaremos a cabo un diseño factorial fraccionado, el cual presenta una cantidad significativamente menor de perfiles a costa de una pérdida de información. Como ya mencionamos, trataremos de generar un diseño lineal ortogonal y balanceado, es decir, 100 % eficiente. Para conocer si existe tal posibilidad bajo las características de nuestro diseño, emplearemos el lenguaje de programación SAS, que contiene numerosas macros de suma utilidad para el campo del análisis conjunto. Toda la programación se detalla en el [Apéndice C](#).

En la [Tabla 3.5](#) se muestran los resultados de la macro `%MktRuns`, la cual calcula el número de perfiles o tamaño del diseño necesario para formar una matriz ortogonal, en nuestro caso, con un atributo de 4 niveles y cuatro atributos de 3 niveles, así como otros posibles tamaños y el número de

infracciones que cometen. Vemos que hay dos posibles diseños ortogonales (indicados con un asterisco), uno con tamaño 36 y otro 72. El diseño saturado es aquel que contiene el número mínimo de perfiles, que coincide con el número de parámetros a estimar en el modelo, en este caso $1 + \sum_{j=1}^k (m_j - 1) = 1 + 3 + 2 + 2 + 2 + 2 = 12$, siendo k el número de atributos y m_j los niveles del atributo j -ésimo. Para que un diseño sea ortogonal y balanceado, su tamaño debe ser divisible por los diferentes números de niveles (3, 4), el cuadrado del nivel mínimo ($3^2 = 9$) y por el producto de los niveles ($3 \times 4 = 12$), en conjunto: 3, 4, 9 y 12. Las infracciones recogen el número de veces que cada tamaño incumple las divisiones.

Tabla 3.5: Tamaños de diseños ortogonales

Tamaño del diseño	Infracciones	No puede ser dividido por
36 *	0	
72 *	0	
18	5	4, 12
27	5	4, 12
45	5	4, 12
54	5	4, 12
63	5	4, 12
12 S	6	9

* – Diseño 100 % eficiente

S – Diseño saturado

Las infracciones recogen el número de veces que cada diseño incumple con las divisiones. Si un diseño presenta infracciones no puede ser ortogonal y balanceado. Entre los diseños ortogonales escogeremos el de tamaño 36, ya que 72 es un número elevado que alargaría la encuesta sin compensar con un alto aporte de información.

El siguiente paso consiste en la determinación de los 36 perfiles. Utilizaremos la macro `%MktEx`, que busca y devuelve un diseño lineal que cumpla con los atributos, niveles y número de perfiles seleccionados, tratando de maximizar la eficiencia D. Esta macro emplea el algoritmo de Fedorov modificado para la búsqueda del diseño final (en el [Apéndice D](#) se amplía información sobre este algoritmo de búsqueda). Hemos incluido una semilla en las llamadas de las macros que emplean algoritmos para generar siempre el mismo resultado en caso de que quieran replicarse los resultados. En la [Tabla 3.6](#) mostramos el resultado y, tal y como se indicó en la [Tabla 3.5](#), existe un diseño 100 % eficiente, por tanto balanceado y ortogonal, para el tamaño 36. Este será el diseño lineal codificado que emplearemos a continuación para construir el diseño de elección (no lineal) organizado por bloques.

Para asegurar que el diseño es ortogonal y balanceado ejecutamos la macro `%MktEval`, que devuelve una tabla donde se muestra la frecuencia de los niveles de los atributos y de los pares de niveles. Como vemos en la parte superior de la [Tabla 3.7](#), la frecuencia es nueve para cada uno de los cuatro niveles de la variable área, en total suman los 36 perfiles del diseño. Para el resto de atributos la frecuencia por nivel también es la misma, doce, porque sólo tienen tres niveles. Con esto se demuestra que, efectivamente, el diseño cumple la propiedad de balance. Por otro lado, en la parte inferior observamos las frecuencias por parejas de atributos. El número de la izquierda se corresponde con el nivel del primer atributo y el de la derecha con el del segundo atributo, por ejemplo, la frecuencia de la pareja Área–Distancia y el par 4-1, representa cuántas veces coinciden en un mismo perfil el nivel 4 del atributo área con el nivel 1 del atributo Distancia. Como para todos los pares de atributos la frecuencia

Tabla 3.6: Perfiles del diseño lineal

Perfiles	Área	Forma	Fragmentación	Distancia	Pendiente
1	1	1	1	1	1
2	1	1	2	1	3
3	1	1	3	2	3
4	1	2	1	3	1
5	1	2	2	2	2
6	1	2	3	2	1
7	1	3	1	3	2
8	1	3	2	1	2
9	1	3	3	3	3
10	2	1	1	1	1
11	2	1	2	1	3
12	2	1	3	2	3
13	2	2	1	3	1
14	2	2	2	2	2
15	2	2	3	2	1
16	2	3	1	3	2
17	2	3	2	1	2
18	2	3	3	3	3
19	3	1	1	2	2
20	3	1	2	3	1
21	3	1	3	3	2
22	3	2	1	1	3
23	3	2	2	3	3
24	3	2	3	1	2
25	3	3	1	2	3
26	3	3	2	2	1
27	3	3	3	1	1
28	4	1	1	2	2
29	4	1	2	3	1
30	4	1	3	3	2
31	4	2	1	1	3
32	4	2	2	3	3
33	4	2	3	1	2
34	4	3	1	2	3
35	4	3	2	2	1
36	4	3	3	1	1

<i>Eficiencia D</i>	<i>Eficiencia A</i>	<i>Eficiencia G</i>
100,00	100,00	100,00

por pares de niveles es la misma, el diseño es ortogonal.

Tabla 3.7: Frecuencias de los niveles de los atributos y de sus combinaciones

Atributos	Frecuencias por niveles											
	1	2	3	4								
Área	9	9	9	9								
Forma	12	12	12									
Fragmentación	12	12	12									
Distancia	12	12	12									
Pendiente	12	12	12									
Parejas de atributos	Frecuencias por pares de niveles											
	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3	4-1	4-2	4-3
Área–Forma	3	3	3	3	3	3	3	3	3	3	3	3
Área–Fragmentación	3	3	3	3	3	3	3	3	3	3	3	3
Área–Distancia	3	3	3	3	3	3	3	3	3	3	3	3
Área–Pendiente	3	3	3	3	3	3	3	3	3	3	3	3
Forma–Fragmentación	4	4	4	4	4	4	4	4	4			
Forma–Distancia	4	4	4	4	4	4	4	4	4			
Forma–Pendiente	4	4	4	4	4	4	4	4	4			
Fragmentación–Distancia	4	4	4	4	4	4	4	4	4			
Fragmentación–Pendiente	4	4	4	4	4	4	4	4	4			
Distancia–Pendiente	4	4	4	4	4	4	4	4	4			

Recordemos que la meta del diseño de elección consiste en definir un conjunto de bloques, dado un valor anticipado β , que minimice el “tamaño” de la matriz de covarianzas, Σ , definida en la ???. En nuestro estudio asumiremos que $\beta = 0$, lo que supone adoptar la hipótesis nula de que las utilidades parciales de los diferentes atributos son 0. La otra opción sería dar una estimación inicial de la importancia de los atributos, pero si finalmente la estimación es incorrecta nos conduciría a resultados erróneos.

Para determinar la mejor división en bloques de alternativas en base a los perfiles obtenidos, empleamos la macro *%ChoiceEff*. Esta macro emplea el ya mencionado algoritmo de Fedorov modificado para encontrar un diseño de elección que minimice las varianzas de los parámetros estimados (error D), dado un conjunto de perfiles y un vector de parámetros β asumido. Además de los perfiles, se debe indicar la disposición de las alternativas, es decir, cuántos bloques y cuántas alternativas por bloque. El número de alternativas mayormente empleado en la práctica es de tres o cuatro por bloque, dependiendo del número de atributos que se presenten, pues a mayor número de atributos y de alternativas, mayor saturación de información para el encuestado. En nuestro caso, al contar con 5 atributos consideramos que lo mejor sería tres alternativas por bloque y, por lo tanto, 12 bloques (preguntas). Barajamos la posibilidad de incluir una cuarta alternativa constante a todos los bloques, que tomase la forma: “Ante estas tres alternativas prefiero no escoger ninguna”. Sin embargo, descartamos esta idea

porque se perdería información si los individuos eligen esta opción con frecuencia, además habría que calcular un parámetro adicional. Una vez conocida la distribución del modelo, procedemos a ejecutar la macro, que se encargará de buscar el diseño más eficiente para los 36 perfiles divididos en 12 bloques de 3 alternativas.

Tabla 3.8: Resultados del diseño de elección con menor Error D tras 100 iteraciones

n	Atributo	Nivel	Varianza	df	Error estándar
1	Area	<0,5	0,1161	1	0,34074
2	Area	0,5–1	0,10496	1	0,32398
3	Area	1–5	0,09611	1	0,31001
4	Forma	muy irregular	0,09487	1	0,30801
5	Forma	irregular	0,10116	1	0,31805
6	Fragmentacion	baja	0,08607	1	0,29337
7	Fragmentacion	media	0,09153	1	0,30255
8	Distancia	>500	0,10139	1	0,31841
9	Distancia	100–500	0,10154	1	0,31865
10	Pendiente	>40	0,09531	1	0,30873
11	Pendiente	20–40	0,10138	1	0,3184

Resultados

Diseño	61
Bloques	12
Alternativas	3
Parámetros	11
Parámetros máximos	24
Eficiencia D	10,1977
Eficiencia D relativa	84,9812
Error D	0,0981
1/Bloques	0,0833

Podemos observar los niveles, sus varianzas asociadas y la eficiencia del modelo en la [Tabla 3.8](#). Elegimos que la macro generase 100 iteraciones, un número razonable a partir del cual es complicado mejorar significativamente la eficiencia del modelo. Observamos diferencias en las varianzas, por ejemplo, la variable área tiene valores ligeramente superiores al resto de atributos debido a su mayor número de niveles, pero las diferencias son muy pequeñas, de manera que no supondrá un problema. La segunda parte de la tabla indica que en la iteración número 61 se encontró el modelo con menor Error D, el cual equivale a $1/\text{Eficiencia D}$. Si este fuese un diseño de elección perfecto, el cual puede alcanzarse sólo para un número genérico y limitado de diseños, entonces la eficiencia D sería igual a $1/\text{Bloques}$, en este caso $1/12=0,0833$. La eficiencia D relativa equivale a la eficiencia D multiplicada por 100 y dividida por el número de bloques. Como vemos, para esta combinación de atributos y niveles no existe un diseño de elección 100 % eficiente, obteniendo una eficiencia relativa de 84,98 %. Si hubiésemos escogido sólo 3 niveles para la variable área habríamos obtenido un diseño 100 % eficiente, ya que habiendo 3 alternativas por bloque y 3 niveles para cada atributo, podría generarse un diseño con bloques balanceados y ortogonales, pero al haber una variable con 4 niveles es imposible generar un bloque de tres alternativas que presente los 4 valores de área. Decidimos aumentar el número de parámetros de área y obtener información más concreta sobre esta variable costeándolo con una

Tabla 3.9: Matriz de varianzas covarianzas del diseño de elección

	<0,5	0,5–1	1–5	muy irregular	irregular	baja	media	>500	100–500	>40	20–40
<0,5	0,1161	0,0098	-0,0122	-0,0037	0,0077	-0,0018	-0,0031	-0,0042	-0,0001	-0,0027	-0,0035
0,5–1	0,0098	0,105	-0,0069	-0,0001	0,0002	0,0031	0,0054	0,0026	0,0077	-0,006	-0,0026
1–5	-0,0122	-0,0069	0,0961	0,0003	-0,001	0	0	-0,0031	0,0051	0,0035	-0,0049
muy irregular	-0,0037	-0,0001	0,0003	0,0949	0,0052	0,0001	0,0001	0,0049	-0,0028	-0,0036	-0,0053
irregular	0,0077	0,0002	-0,001	0,0052	0,1012	-0,0001	-0,0002	-0,0029	-0,0052	-0,0003	-0,0004
baja	-0,0018	0,0031	0	0,0001	-0,0001	0,0861	0,0047	0,0002	0,0002	-0,0001	0
media	-0,0031	0,0054	0	0,0001	-0,0002	0,0047	0,0915	0,0003	0,0004	-0,0002	0
>500	-0,0042	0,0026	-0,0031	0,0049	-0,0029	0,0002	0,0003	0,1014	-0,0003	-0,0081	0,0025
100–500	-0,0001	0,0077	0,0051	-0,0028	-0,0052	0,0002	0,0004	-0,0003	0,1015	0,0023	-0,0052
>40	-0,0027	-0,006	0,0035	-0,0036	-0,0003	-0,0001	-0,0002	-0,0081	0,0023	0,0953	0,0054
20–40	-0,0035	-0,0026	-0,0048	-0,0053	-0,0004	0	0	0,0025	-0,0052	0,0054	0,1014

pequeña pérdida de eficiencia en el modelo.

En la [Tabla 3.9](#) se muestra la matriz de varianzas covarianzas de la matriz de diseño. Observamos covarianzas mayores que cero debidas al desbalance interno dentro de cada bloque, pero con valores despreciables por debajo de 0,01.

A continuación, representamos en la [Tabla 3.10](#) el diseño de elección para 12 bloques y 3 alternativas generado por la macro. Mostramos los valores codificados en la parte izquierda, al igual que hicimos con el diseño lineal, y sin codificar en la parte derecha, por ser el diseño final y para facilitar su interpretación.

Tabla 3.10: Diseño de elección codificado y sin codificar

Bloque	Area	Forma	Fragmentacion	Distancia	Pendiente	Area	Forma	Fragmentacion	Distancia	Pendiente
1	4	2	3	1	2	>5	Irregular	Alta	>500	20-40
	1	3	1	3	2	<0,5	Regular	Baja	<100	20-40
	2	1	2	1	3	0,5-1	Muy irregular	Media	>500	<20
2	2	3	3	3	3	0,5-1	Regular	Alta	<100	<20
	1	2	2	2	2	<0,5	Irregular	Media	100-500	20-40
	2	1	1	1	1	0,5-1	Muy irregular	Baja	>500	>40
3	3	2	2	3	3	1-5	Irregular	Media	<100	<20
	2	3	1	3	2	0,5-1	Regular	Baja	<100	20-40
	1	2	3	2	1	<0,5	Irregular	Alta	100-500	>40
4	4	2	1	1	3	>5	Irregular	Baja	>500	<20
	3	3	2	2	1	1-5	Regular	Media	100-500	>40
	4	1	3	3	2	>5	Muy irregular	Alta	<100	20-40
5	4	2	2	3	3	>5	Irregular	Media	<100	<20
	3	2	3	1	2	1-5	Irregular	Alta	>500	20-40
	1	1	1	1	1	<0,5	Muy irregular	Baja	>500	>40
6	4	3	2	2	1	>5	Regular	Media	100-500	>40
	3	1	3	3	2	1-5	Muy irregular	Alta	<100	20-40
	1	2	1	3	1	<0,5	Irregular	Baja	<100	>40
7	3	3	1	2	3	1-5	Regular	Baja	100-500	<20
	2	2	3	2	1	0,5-1	Irregular	Alta	100-500	>40
	1	1	2	1	3	<0,5	Muy irregular	Media	>500	<20
8	2	3	2	1	2	0,5-1	Regular	Media	>500	20-40
	4	3	1	2	3	>5	Regular	Baja	100-500	<20
	3	1	2	3	1	1-5	Muy irregular	Media	<100	>40
9	3	1	1	2	2	1-5	Muy irregular	Baja	100-500	20-40
	4	3	3	1	1	>5	Regular	Alta	>500	>40
	2	2	2	2	2	0,5-1	Irregular	Media	100-500	20-40
10	1	3	2	1	2	<0,5	Regular	Media	>500	20-40
	1	1	3	2	3	<0,5	Muy irregular	Alta	100-500	<20
	2	2	1	3	1	0,5-1	Irregular	Baja	<100	>40
11	4	1	2	3	1	>5	Muy irregular	Media	<100	>40
	3	2	1	1	3	1-5	Irregular	Baja	>500	<20
	2	1	3	2	3	0,5-1	Muy irregular	Alta	100-500	<20
12	3	3	3	1	1	1-5	Regular	Alta	>500	>40
	4	1	1	2	2	>5	Muy irregular	Baja	100-500	20-40
	1	3	3	3	3	<0,5	Regular	Alta	<100	<20

3.2.2. Recogida de datos

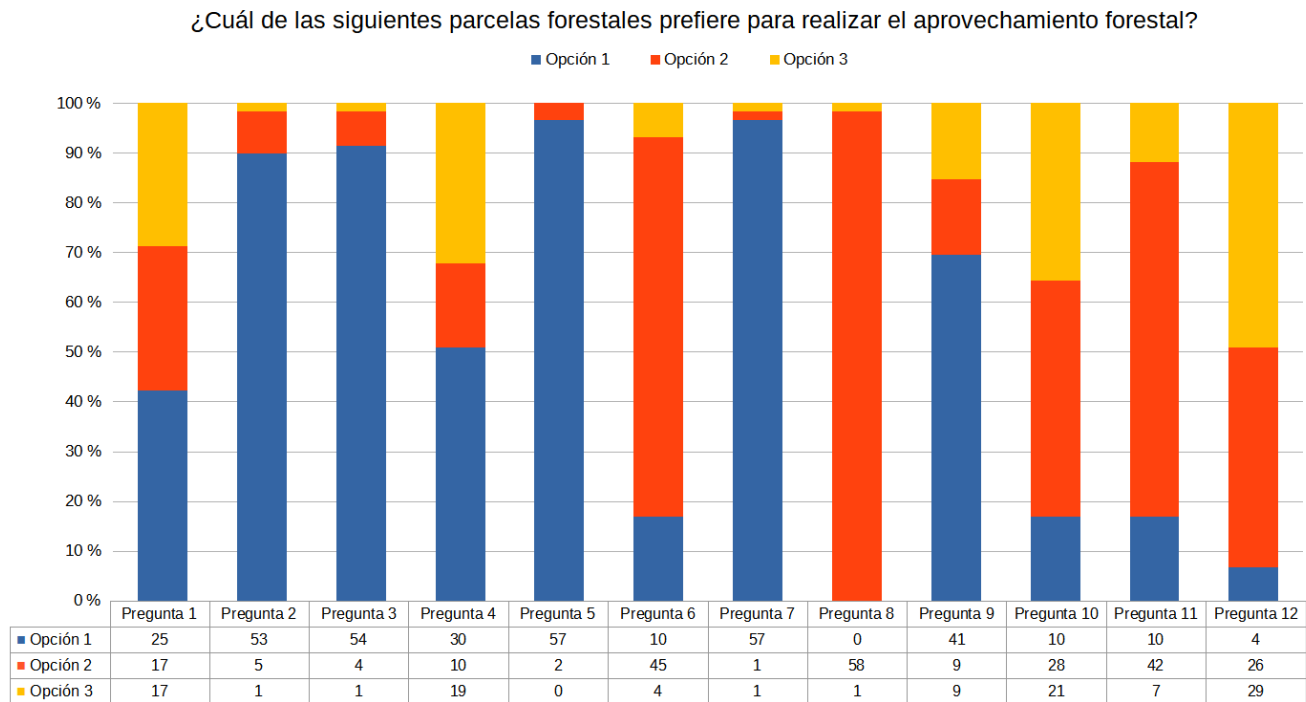
Teniendo el diseño de la encuesta, procedemos a su presentación. Decidimos hacerlo online por la comodidad que supone para los encuestados, por lo que buscamos un software que permitiese, por una parte, formatear y editar la encuesta, y por otra distribuirla con facilidad. Para esta tarea hemos usado el software Formularios de Google.

En la primera página del cuestionario redactamos una introducción de la encuesta donde se indica que es un cuestionario anónimo, así como que no hay respuestas correctas e incorrectas. Con esto tratamos de acomodar al encuestado para que responda con tranquilidad y sinceridad. En la siguiente hoja recopilamos una serie de aclaraciones sobre el formato de la encuesta, adelantando que habrá 12 preguntas conformadas por 3 opciones cada una. Estas opciones representan una parcela forestal descrita por una serie de variables y su tarea será escoger una por pregunta. A continuación, indicamos brevemente la definición de cada variable y sus unidades. Lo siguiente que se muestra son las 12 preguntas. El cuestionario terminará una vez se hayan respondido todas, sin poder dejar preguntas en blanco. En el [Apéndice E](#) hemos incorporado el cuestionario generado con Formularios de Google al que han respondido los encuestados.

La cuestión planteada para cada pregunta fue: ¿Cuál de las siguientes parcelas forestales prefiere para realizar un aprovechamiento forestal? Para representar las opciones de manera visual para los encuestados concluimos que lo mejor era mostrar, en primer lugar, los intervalos que cumplen los atributos para cada opción y acompañarlo de un ejemplo concreto que cumpla dichas características y en el que se indiquen asimismo sus valores puntuales. Para localizar los ejemplos lo primero que hicimos fue filtrar las parcelas por perfiles, por ejemplo, para la primera opción de la primera pregunta nos quedamos con aquellos terrenos que cumpliesen con los niveles vistos en la [Tabla 3.10](#), que son: área mayor de 5 hectáreas, forma irregular, fragmentación alta, distancia mayor a 500 metros y pendiente entre el 20 y el 40 %. Una vez filtradas, escogimos una parcela al azar y buscamos su referencia catastral en la Sede Electrónica del Catastro, donde obtuvimos una imagen del terreno. Repetimos esta operación para las 36 opciones. Consideramos importante resaltar al final de cada pregunta que las imágenes de las parcelas no comparten la misma escala, pues la diferencia de tamaño entre unas y otras imposibilitó la tarea.

La encuesta se dirigió a las principales empresas y asociaciones profesionales del sector forestal gallego. En total respondieron 59 personas a la encuesta. En la [Figura 3.2](#) plasmamos las repuestas obtenidas en valores absolutos en el eje horizontal y en valores relativos en el eje vertical. Observamos que en las preguntas número 2, 3, 5, 7 y 8 más del 90 % de los encuestados respondieron la misma opción. Por ejemplo, si nos fijamos en las opciones de la pregunta 2, veremos que todos los atributos de la opción 1 (perfil 1) son iguales o mejores que los de las otras dos opciones, por lo que tiene sentido la gran preferencia demostrada. De hecho, que haya preguntas donde la gran mayoría escojan el mismo perfil es un buen indicativo de que los encuestados están respondiendo con seriedad y atención a la encuesta.

Figura 3.2: Resultados de la encuesta



3.2.3. Estimación de las utilidades parciales e interpretación

A continuación aplicaremos la regresión logística multinomial al conjunto de resultados recogidos. Nuevamente, nos valdremos del software SAS para el cálculo. Utilizaremos la función PROC PHREG, para la cual se ha tenido que elaborar una matriz de datos donde figuren las 36 opciones del cuestionario para cada individuo, es decir $59 \times 36 = 2.124$ filas, caracterizadas por sus respectivos atributos y una columna extra donde se indica si esa opción se escogió (2) o no (1), que será la variable dependiente del modelo. Los pasos para la obtención de esta matriz se indican en el [Apéndice C](#) paso a paso. Utilizaremos como parámetros base para cada atributo aquel que debería haber sido peor valorado por los encuestados, es decir, área menor de 0,5 ha, forma muy irregular, fragmentación baja, distancia mayor a 500 metros y pendiente superior a 40 grados. De esta forma, si los resultados son como esperamos, será más sencillo interpretar los coeficientes estimados porque irán de menor a mayor.

En la [Tabla 3.11](#) está toda la información referida al modelo. Lo primero que vemos son los resultados de los test de bondad de ajuste. El primero es el test de la razón de verosimilitud, el cual es una prueba de la suficiencia de un modelo más pequeño frente a un modelo más complejo. La hipótesis nula de la prueba indica que el modelo más pequeño proporciona un ajuste tan bueno para los datos como el modelo complejo. Si la hipótesis nula es cierta, el estadístico de test seguirá una distribución χ^2 . El estadístico de la razón de verosimilitud ($-2\ln(L)$) fue 892,321, y su estadístico chi-cuadrado 663,31 con 11 grados de libertad rechaza la hipótesis nula con un p-valor menor que 0,0001. La tabla también presenta los valores de los criterios de Akaike y Bayesiano.

En cuanto a las utilidades parciales (coeficientes) estimadas vemos que todas son estadísticamente significativas salvo la fragmentación media, que comparado con el resto de niveles tiene un estadístico χ^2

Tabla 3.11: Resultados de la regresión logística multinomial

Test de bondad de ajuste		Test de hipótesis nula: $\beta = 0$				
$-2\ln(L)$		Test	χ^2	df	p-valor	
AIC	892,321					
BIC	914,321					
	964,508	Ratio de verosimilitud	663,31	11	< 0,0001	
		Puntuación	575,93	11	< 0,0001	
		Wald	294,67	11	< 0,0001	

Resultados estimación por máxima verosimilitud							
Atributo	Nivel	Utilidad parcial	Importancia	Desv. típica	df	χ^2	p-valor
Área	0,5–1	0,6008	2,2836	0,1949	1	9,4993	0,0021
	1–5	2,2836		0,2381	1	91,9944	< 0,0001
	>5	2,2628		0,2237	1	102,2739	< 0,0001
Forma	Irregular	0,4137	0,6358	0,1561	1	7,0215	0,0081
	Regular	0,6358		0,1597	1	15,8378	< 0,0001
Fragmentación	Media	-0,1450	0,3291	0,1557	1	0,8680	0,3515
	Alta	-0,3291		0,1272	1	6,6849	0,0097
Distancia	100–500	0,8008	1,5029	0,1783	1	20,1693	< 0,0001
	<100	1,5029		0,1695	1	78,5402	< 0,0001
Pendiente	20–40	1,1512	2,4675	0,1934	1	35,4348	< 0,0001
	<20	2,4675		0,1916	1	165,7588	< 0,0001
Total			7,2189				

más pequeño, y un p-valor por encima del 5 %. Además, la fragmentación alta es el segundo parámetro (nivel) del modelo con menor χ^2 . Ante esto, interpretamos que, de los 5 atributos analizados, la fragmentación es el menos relevante. De hecho, los coeficientes de la fragmentación media (-0,1450) y alta (-0,3291) tienen una utilidad parcial menor que la fragmentación baja (0), un resultado contrario al planteamiento de partida, donde supusimos que una fragmentación alta favorecería la utilidad de la parcela. El resto de atributos sí obtuvieron los coeficientes que esperábamos, viendo incrementada su utilidad parcial con respecto al nivel de referencia a medida que aumenta su nivel.

La variable área presenta valores positivos para todos los niveles. Los individuos demostraron mayor preferencia por parcelas entre 0,5 y 1 hectárea que por parcelas menores de 0,5 hectáreas (0,6008). Sin embargo, esta preferencia es mucho mayor por parcelas entre 1 y 5 hectáreas (2,2836) y mayores de 5 hectáreas (2,2628), siendo el valor prácticamente igual entre ambos niveles, por lo que, en base a los resultados, podemos interpretar que los encuestados son indiferentes entre el tamaño de las parcelas cuando éstas superan 1 hectárea de superficie, con el matiz de que el nivel 1–5 tiene un rango muy amplio de hectáreas, por tanto 1 ha podría no ser el valor umbral.

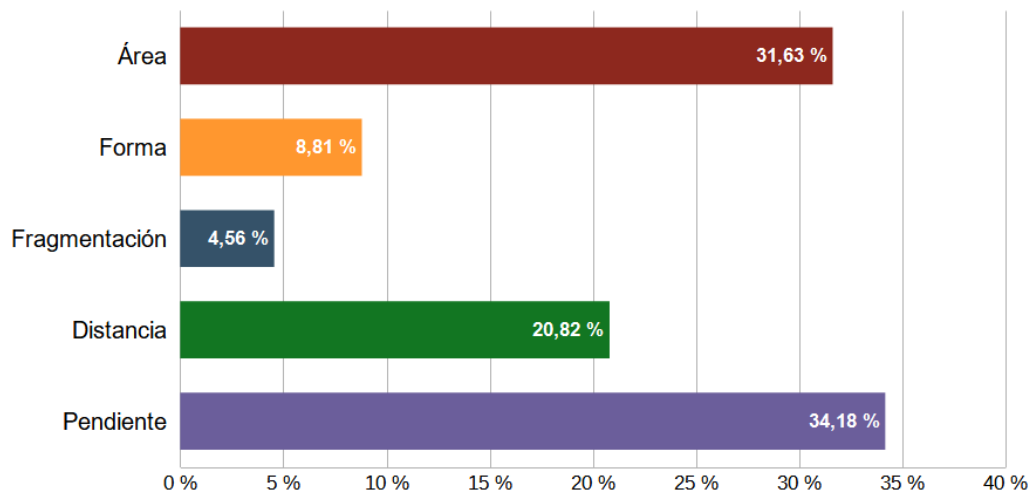
El atributo forma mostró un incremento de la utilidad parcial a medida que se regulariza la forma de la parcela, por tanto, se ha comportado como esperábamos. Sin embargo, los individuos no demostraron grandes preferencias por las formas irregular y regular (0,4137 y 0,6358) con respecto al valor de

referencia (muy irregular), por lo que suponemos que los distintos niveles del atributo forma no son determinantes a la hora de decidirse por una parcela.

En cuanto a la distancia, presenta mayores utilidades parciales con respecto al nivel de referencia y lo mismo pasa con la pendiente, la cual presenta el nivel con mayor utilidad parcial de todo los atributos. Es decir, que la pendiente sea menor de 20 grados tiene una preferencia con respecto a que tenga más de 40 grados (2,4675) mayor a cualquier otra preferencia del modelo, lo cual resulta sorprendente pues esperábamos que la mayor utilidad parcial se encontrase en el atributo área.

Es importante recalcar que no pueden compararse las utilidades parciales de un atributo frente a las de otro, ya que están calculadas en base al nivel de referencia de cada atributo. Entonces, para comparar las preferencias por atributo debemos calcular la importancia de ese atributo. Comúnmente se computa como la diferencia entre los niveles con mayor y menor utilidad parcial de cada atributo. En nuestro caso la utilidad parcial mínima es el nivel de referencia 0 (salvo fragmentación, donde 0 es la mayor utilidad), por tanto la importancia será el valor absoluto de la mayor utilidad parcial. Por ejemplo, el nivel con mayor utilidad de área es 2,2836, por lo tanto esa será su importancia. Puede contemplarse en la columna “Importancia” de la [Tabla 3.11](#). Luego, la importancia relativa será el porcentaje de la importancia total que represente cada atributo. Esta importancia relativa no indica que un atributo sea estrictamente más preferido que otro, sino que señala qué diferencias entre el peor y el mejor nivel de un atributo generan una mayor respuesta por parte del individuo. En la [Figura 3.3](#) representamos gráficamente la importancia relativa de los distintos atributos. Observamos que el atributo con mayor importancia es la pendiente, seguido de área y distancia, acumulando entre las tres el 86,63 % de la importancia relativa. El restante 13,37 % se reparte entre forma y fragmentación.

Figura 3.3: Importancia relativa de los atributos



Podemos desagregar las importancias relativas en los distintos niveles para entender en mayor profundidad las preferencias de los individuos. Para ello debemos centrar las utilidades relativas de los niveles de cada atributo sobre 0. Incluimos un ejemplo del cálculo para la variable área en la [Tabla 3.12](#), donde se ve cómo el primer paso consiste en centrar las utilidades parciales, luego se dividen entre el sumatorio del valor absoluto de las utilidades centradas y por último se multiplican por la importancia relativa del atributo, de forma que la suma de esta última cantidad esté centrada sobre 0. En la [Figura 3.4](#) observamos las preferencias relativas por nivel. Lo más llamativo es la gran preferencia que muestran los individuos por pendientes menores a 20 grados con respecto a los otros dos niveles de

Tabla 3.12: Cálculo de la preferencia por nivel del atributo área

Nivel	Imp. rel	U. parcial	Media	U. parcial – media	$\Sigma (U. \text{parcial} - \text{media})$	$\frac{U. \text{parcial} - \text{media}}{\Sigma(U. \text{parcial} - \text{media})}$	$\frac{U. \text{parcial} - \text{media}}{\Sigma(U. \text{parcial} - \text{media})} \times Imp.rel$
< 0,5	31,63	0	1,2868	-1,2868	3,9456	-0,3261	-10,31 %
0,5–1		0,6008		-0,6860		-0,1738	-5,49 %
1–5		2,2836		0,9968		0,2526	7,99 %
> 5		2,2628		0,9760		0,2473	7,82 %

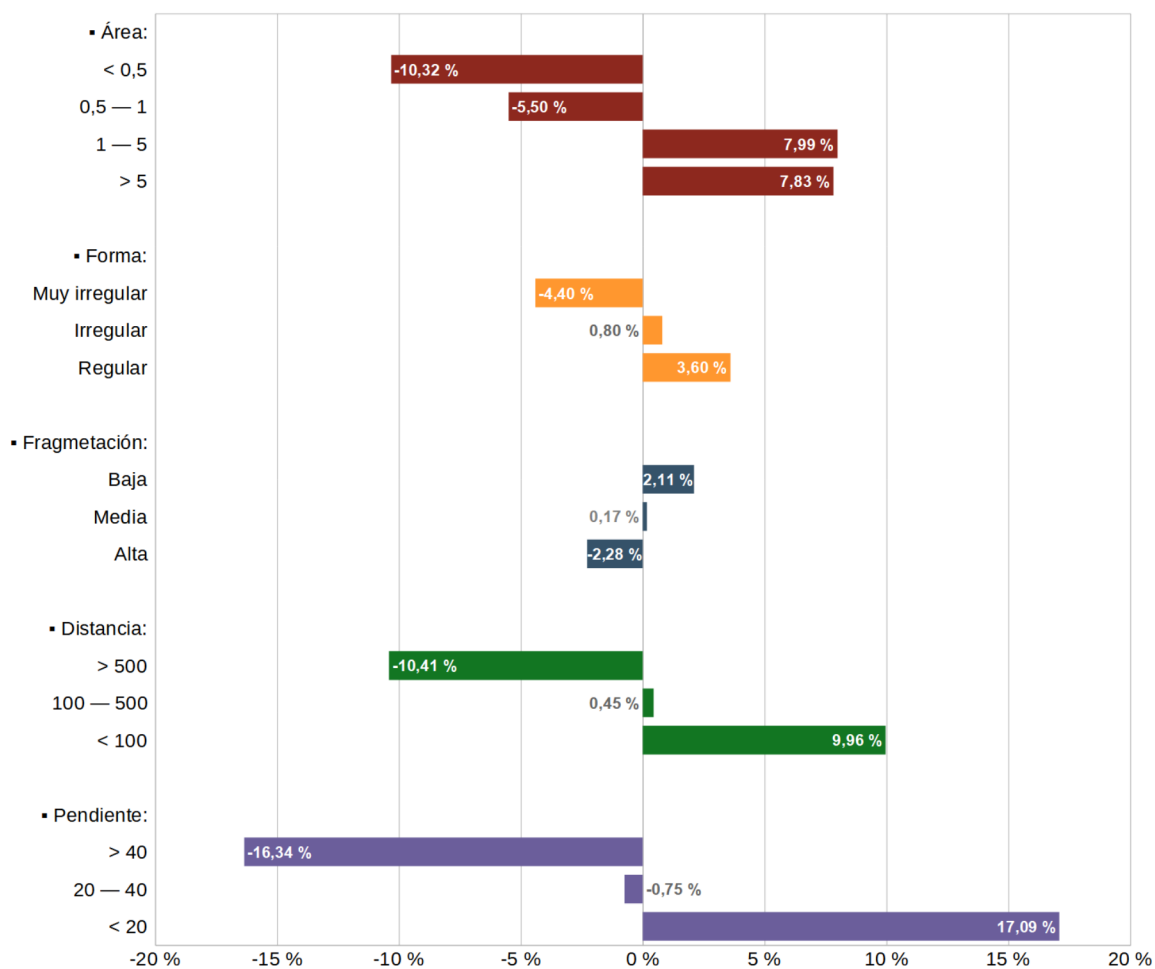
pendiente. Cabe aclarar que un porcentaje negativo no indica que ese nivel influya negativamente en la utilidad total del encuestado, sino que la utilidad parcial de ese nivel es inferior a la media de las utilidades parciales del atributo.

Ahora que conocemos las utilidades parciales de los niveles, podemos calcular la utilidad total del conjunto de los 324 perfiles posibles y ordenarlos por orden de mayor a menor utilidad. Esto nos permitiría conocer los perfiles de parcela forestal más deseados por los profesionales de la madera. En la [Tabla 3.13](#) mostramos los 10 perfiles con mayor utilidad de los 324 posibles. En la tabla simplemente sustituimos el valor del nivel por la utilidad parcial si ese perfil toma ese nivel para ese atributo. Si todos los niveles de un atributo son 0 implica que el perfil toma el valor del nivel de referencia, por lo tanto su utilidad parcial es 0.

Tabla 3.13: 10 perfiles con mayor utilidad del total de 324

Perfil	Área			Forma		Fragmentación		Distancia		Pendiente		Utilidad
	0,5–1	1–5	>5	Irregular	Regular	Media	Alta	100–500	<100	20–40	<20	
1	0	2,2836	0	0	0,6358	0	0	0	1,5029	0	2,4675	6,8898
2	0	0	2,2628	0	0,6358	0	0	0	1,5029	0	2,4675	6,8690
3	0	2,2836	0	0	0,6358	-0,1450	0	0	1,5029	0	2,4675	6,7448
4	0	0	2,2628	0	0,6358	-0,1450	0	0	1,5029	0	2,4675	6,7240
5	0	2,2836	0	0,4137	0	0	0	0	1,5029	0	2,4675	6,6677
6	0	0	2,2628	0,4137	0	0	0	0	1,5029	0	2,4675	6,6469
7	0	2,2836	0	0	0,6358	0	-0,3291	0	1,5029	0	2,4675	6,5607
8	0	0	2,2628	0	0,6358	0	-0,3291	0	1,5029	0	2,4675	6,5399
9	0	2,2836	0	0,4137	0	-0,1450	0	0	1,5029	0	2,4675	6,5227
10	0	0	2,2628	0,4137	0	-0,1450	0	0	1,5029	0	2,4675	6,5399

Figura 3.4: Importancia relativa de los niveles (los porcentajes negativos indican el porcentaje de una utilidad inferior a la media)



3.2.4. Conclusiones

Este trabajo ha tenido por objeto el estudio y cuantificación de las preferencias de los profesionales del sector de la silvicultura gallega a la hora de elegir una parcela forestal para su explotación, tomando como motivo la gran improductividad y minufundismo del monte gallego. Para ello, lo primero que hicimos fue indentificar cinco variables físicas, presentes en todos los terrenos, que intuimos serían ponderadas por los profesionales a la hora de su elección. Luego, empleamos técnicas de análisis conjunto ayudándonos del software estadístico SAS para estimar las utilidades parciales de los encuestados.

Lo primero que hicimos fue un estudio previo para determinar aquellos atributos físicos del terreno que serían más valorados por los profesionales del sector de la madera. Al final nos decantamos por el área o superficie del terreno, la forma de la parcela, la fragmentación circundante, la distancia media a la vía más cercana y la pendiente modal de la parcela. El siguiente paso consistió en determinar los niveles de dichos atributos, para ello lanzamos una encuesta dentro del sector de la madera en la que se pedía a los encuestados que identificasen el valor umbral de los atributos mencionados, a partir del cual no estarían dispuestos a realizar la compra o aprovechamiento forestal de una parcela, y al contrario, el valor a partir del cual considerarían la compra o aprovechamiento como favorables. Una vez escogidos los atributos y niveles, decidimos aplicar técnicas de análisis conjunto basado en elecciones para estimar las preferencias, considerando sus ventajas frente a técnicas más clásicas como el análisis conjunto tradicional o el más simple análisis autoexplicativo. Dicho método consiste en analizar los resultados de una encuesta formada por una serie de preguntas. En dichas preguntas deben figurar combinaciones de los diferentes atributos y niveles que estudiamos previamente. Para elaborar la encuesta empleamos conjuntamente el diseño experimental y el algoritmo modificado de Fedorov implementado en SAS. Con estas herramientas encontramos un diseño lo más óptimo posible bajo el criterio de maximizar la eficiencia, lo que a su vez minimizó las varianzas y covarianzas de los niveles estimados. En base a este diseño construimos una encuesta en Formularios de Google y la distribuimos entre las principales empresas y asociaciones profesionales del sector forestal gallego. En total respondieron 59 personas. A partir de sus respuestas estimamos por máxima verosimilitud un modelo multinomial logístico que permitió identificar las preferencias de los encuestados por las 5 variables estudiadas y sus diversos niveles.

Los resultados de la estimación fueron muy cercanos a lo esperado. Las variables más valoradas por los profesionales de la madera fueron la pendiente del terreno, su superficie y la distancia a la vía más cercana. Básicamente, una menor pendiente facilita el trabajo y el desplazamiento por el terreno, a la vez que permite la entrada de vehículos de carga en la finca, por lo que esta variable repercute en la productividad del aprovechamiento forestal. Por su parte, la superficie de la parcela repercute directamente en los beneficios obtenidos, ya que, por lo general, a mayor tamaño más cantidad de árboles y más madera. Por último la distancia a la vía más cercana desde la que transportar la madera extraída también influye en la productividad de la extracción. Las otras dos variables incluidas en el estudio fueron el grado de fragmentación de las parcelas colindantes y la forma del terreno. Sobre la fragmentación, el modelo mostró que el nivel intermedio no era estadísticamente significativo, pero los individuos sí que demostraron preferencia por una fragmentación baja frente a una fragmentación alta. La forma del terreno también la toman en cuenta, pero tiene menor importancia que la pendiente, la superficie o la distancia.

Una posible aplicación para los resultados obtenidos vendría por el lado de construir un indicador que permita identificar la utilidad de cualquier parcela en Galicia. Tal y como mostramos en la [Tabla 3.13](#) conocemos las utilidades totales de cada uno de los 324 perfiles posibles de parcela, y ya que conocemos los valores de las variables estudiadas para la totalidad de las 6.689.822 parcelas forestales en Galicia, podemos categorizar cada una dentro de los 324 perfiles, y por lo tanto asignarle una utilidad. Además, conocemos la ubicación de cada parcela, con lo que podrían ubicarse fácilmente en el mapa. Esto serviría como herramienta de gestión para la autoridad competente, ya que, a día de

hoy, no existe ninguna herramienta en el ámbito forestal que permita identificar el potencial de una parcela basándose en las características físicas del terreno. Por ejemplo, podría detectarse una zona del mapa con acumulación de terrenos con bajas utilidades comparados con los de las zonas aledañas. Esto sería un indicativo de posible improductividad en dicho lugar, con lo que podrían proponerse planes de reestructuración parcelaria que permitan reunificar los terrenos para hacerlos más eficientes. También podría utilizarse para detectar zonas con alto potencial forestal, es decir, con altas utilidades, que no se estén explotando en el momento. A nivel individual también podría servirle a las empresas y asociaciones profesionales del sector para detectar parcelas dentro de su radio de acción y ponerse en contacto con el dueño del terreno para ofrecerle un aprovechamiento forestal.

Apéndice A

Tablas de frecuencias

Para complementar los histogramas de la [Figura 3.1](#), decidimos mostrar una tabla de frecuencias de las variables cuantitativas separándolas por intervalos que permitan apreciar la segmentación más en detalle.

Tabla A.1: Tabla de frecuencias

Area				Forma			
Intervalos	n_i	f_i (%)	F_i (%)	Intervalos	n_i	f_i (%)	F_i (%)
[0; 0,5)	6.161.776	92,11	92,11	[0; 0,1)	324.498	4,85	4,85
[0,5; 1)	320.058	4,78	96,89	[0, 1; 0, 2)	533.894	7,98	12,83
[1; 1,5)	94.987	1,42	98,31	[0, 2; 0, 3)	736.808	11,01	23,84
[1,5; 2)	40.933	0,61	98,92	[0, 3; 0, 4)	861.449	12,88	36,72
[2; 2,5)	21.387	0,32	99,24	[0, 4; 0, 5)	976.647	14,60	51,32
[2,5; 3)	12.899	0,19	99,43	[0, 5; 0, 6)	1.070.951	16,01	67,33
[3; 3,5)	8.197	0,12	99,55	[0, 6; 0, 7)	1.094.079	16,35	83,68
[3,5; 4)	5.527	0,08	99,63	[0, 7; 0, 8)	1.017.041	15,20	98,88
[4; 4,5)	3.994	0,06	99,69	[0, 8; 0,9)	73.375	1,10	99,98
[4,5; 5)	2.950	0,04	99,73	[0, 9; 1)	1.076	0,02	100
[5, 2140)	17.113	0,26	100				

DistViasMe			
Intervalos	n_i	f_i (%)	F_i (%)
[0; 50)	2.925.475	43,73	43,73
[50; 100)	1.700.060	25,41	69,14
[100; 150)	959.690	14,35	83,49
[150; 200)	524.194	7,84	91,33
[200; 250)	278.258	4,16	95,49
[250; 300)	145.995	2,18	97,67
[300; 350)	74.988	1,12	98,79
[350; 400)	38.717	0,58	99,37
[400; 450)	19.760	0,30	99,67
[450; 500)	9.999	0,15	99,82
[500; 2802)	12.685	0,19	100

Apéndice B

Encuesta de sondeo previa

La encuesta fue respondida por 60 personas. Las preguntas planteadas fueron las siguientes:

1. ¿Por debajo de qué tamaño de parcela (en hectáreas) no consideraría realizar la compra o el aprovechamiento?
2. Aunque algunas veces realice aprovechamientos en parcelas menores, ¿a partir de qué tamaño (en hectáreas) considera que sería óptimo realizar el aprovechamiento?
3. ¿Por debajo de qué valor de pendiente (en %) es óptimo realizar un aprovechamiento?
4. ¿Cuál es el valor de pendiente (en %) a partir del cual considera inviable realizar el aprovechamiento de una parcela?
5. ¿Cuál es el valor de pendiente (en %) a partir del cual realizaría el aprovechamiento solo en el caso de que el resto de condiciones fuesen fantásticas?
6. ¿Cuál es la distancia a cargadero (en metros) por debajo de la cual considera el aprovechamiento forestal cómodo de realizar?
7. ¿A partir de qué distancia a cargadero (en metros) no haría el aprovechamiento en ningún caso?

Apéndice C

Código empleado en SAS

```
%mktruns(4 3 3 3 3) /* matrices ortogonales para estos niveles de atributos */

%mktdes(factors=x1=4 x2-x5=3, n=36, maxiter=1000, examine=i v) /* matriz de información*/
proc print; run;

%mktex(4 3 3 3 3, /* niveles y número de atributos */
n=36) /* número de perfiles */
proc print; run; /* mostrar el diseño resultante */

%mktlab(data=design)

%mkteval(data=design) /* frecuencias de los atributos */

%choiceff(data=design, /* conjunto de perfiles a utilizar */
model=class(x1-x5/sta), /* modelo codificado con ortogonalidad estándar */
nsets=12, /* número de bloques */
flags=3, /* número de alternativas */
seed=185, /* semilla aleatoria */
maxiter=100, /* número máximo de diseños a iterar */
options=relative, /* mostrar eficiencia D relativa */
beta=zero); /* vector beta asumido, Ho: b=0 */

proc print; var x1-x5; id set; by set; run; /* mostrar el diseño final codificado */

proc print data=bestcov label; /* mostrar la matriz var-cov */
    title 'Variance-Covariance Matrix';
    id __label;
    label __label = '00'x;
    var x:;
    run;
title;

proc format; /* valores reales de los niveles codificados */
    value area 1 = '<0.5ha' 2 = '0.5-1ha' 3 = '1-5ha' 4 = '>5ha';
    value forma 1 = 'Muy irregular' 2 = 'Irregular' 3 = 'Regular';
    value fragmentacion 1 = 'Baja' 2 = 'Media' 3 = 'Alta';
    value distancia 1 = '>500m' 2 = '100-500m' 3 = '<100m';
    value pendiente 1 = '>40%' 2 = '20-40%' 3 = '<20%';
run;
```

```

data ChoiceDesign;          /* el mejor diseño generado por el algoritmo */
set best;
format x1 area. x2 forma. x3 fragmentacion. x4 distancia. x5 pendiente.;
label x1 = Area x2 = Forma x3 = Fragmentacion x4 = Distancia x5 = Pendiente;
run;

proc print label; var x1-x5; id set; by set; run;

proc SQL;                  /* creación de la tabla que contiene las respuestas */
create table Diseño AS
SELECT SET, X1 AS Area, X2 AS Forma, X3 AS Fragmentacion, X4 AS Distancia, X5 AS Pendiente
FROM ChoiceDesign;

title 'Respuestas del cuestionario';

data results; /*resultados de la encuesta*/
input Sujeto (Pregunta_1-Pregunta_12) (1.) @@;
datalines;

1 311112121223 2 111312121222 3 111312121322
4 111312121312 5 111312121322 6 311212121321
7 211112121312 8 211112121233 9 111112121222
10 111211121113 11 311112121223 12 212211123122
13 211312121222 14 111212123321 15 311112123123
16 311312121221 17 212311121332 18 311112121223
19 312113121223 20 111312121122 21 211311121223
22 221111121312 23 211312121233 24 111112121132
25 311112121322 26 211112123323 27 111312122112
28 311312121312 29 311112121223 30 311112121223
31 311113123323 32 211221123322 33 311112123333
34 211112121223 35 111112121312 36 211213121212
37 111312122123 38 111112121222 39 111311122233
40 131212222321 41 311113121223 42 111212121223
43 111312121223 44 221211121322 45 323122132332
46 111112123213 47 311112121222 48 111112121123
49 222212121223 50 111112121223 51 111312122122
52 221312121312 53 111111122222 54 111112121223
55 111112123223 56 211312121223 57 311312121322
58 211112322123 59 111111122323
;

data Respuestas;
set results;
if _N_ <= 59 then output;
run;

/* %mktmerge genera las 36 opciones para los 59 individuos encuestados
y señala con un 1 las opciones escogidas en cada pregunta */

%mktmerge(design=Diseño, data=Respuestas, out=res2,
nsets=12, nalts=3, setvars=Pregunta_1-Pregunta_12)

proc print data=res2(obs=39);
id Sujeto set; by Sujeto set;
run;

```



```

/* La siguiente sentencia calcula la estimación del modelo. La opción design=5000
permite que se procese la información en bloques de 5000 filas. La opción de
nonzeroconstant no aplica para este problema, pero es recomendable añadirla.
Provoca que, si se genera una variable constante, esta no pueda ser igual a 0.
La parte del class() declara que las variables son categóricas. La opción zero=first
indica que se deben tomar como niveles de referencia los primeros. Por último
el output señala el conjunto de datos de salida y el drop elimina las variables
del modelo que no queremos, como el intercepto.*/

proc transreg design=5000 data=res2 nonzeroconstant norestoremissing;
model class(Area Forma Fragmentacion Distancia Pendiente / zero=first);
output out=coded(drop=_type_ _name_ intercept);
id Sujeto set c;
run;

proc phreg data=coded outest=betas brief;
model c*c(2) = &_trgind / ties=breslow;
strata Sujeto set;
run;

/* A partir de aquí es código adicional no mencionado en el cuerpo del TFM.
Las siguientes sentencias permiten estimar, según los resultados del
modelo, las probabilidades que tiene un individuo de elegir las
respuestas de las 12 preguntas. Podrían estimarse también las
probabilidades para nuevas preguntas generadas. */

proc score data=coded(where=(Sujeto=1) drop=c)
score=betas type=parms out=p;
var &_trgind;
run;

data p2;
set p;
p = exp(c);
run;

proc means data=p2 noprint;
output out=s sum(p) = sp;
by set;
run;

data p;
merge p2 s(keep=set sp);
by set;
p = p / sp;
keep brand set price p;
run;

proc print data=p;
title2 ?Probabilidades de elección por bloque?;
id set; by set;
run;

```


Apéndice D

Algoritmo del diseño

En este trabajo estamos realizando un análisis conjunto que cuenta con cinco atributos, uno de cuatro niveles y el resto de tres niveles. Para el estudio queremos construir un diseño de 12 bloques con 3 perfiles cada uno. El total de perfiles posibles para esta combinación de atributos y niveles es de 324, teniendo en cuenta las condiciones del diseño habría un total de $\frac{324^{324}}{3!12!}$ diseños posibles, por lo que resulta imposible calcularlos todos. En su lugar, el código de SAS implementa el algoritmo de Fedorov modificado¹. El proceso empieza por construir una lista con todas las alternativas candidatas a formar parte del diseño. A continuación, elige un diseño inicial aleatorio. Entonces, el algoritmo se encarga de ir alterando el diseño inicial intercambiando sus perfiles por las alternativas candidatas hasta encontrar la mejor (si existe) para el primer perfil del diseño. La primera iteración termina cuando el algoritmo encuentre secuencialmente los mejores intercambios para todas las alternativas del diseño inicial. A continuación empezaría una nueva iteración, volviendo a buscar el mejor intercambio para el primer perfil del diseño resultante de la primera iteración, hasta que no haya ninguna mejora sustancial posible en la eficiencia del diseño. Para evitar malos óptimos locales puede cambiarse la semilla, lo que alteraría el diseño inicial del problema. Aunque no haya garantía de alcanzar un óptimo, los resultados de su aplicación son buenos.

En nuestro caso, la lista de candidatos está formada por 324 perfiles y en el diseño figurarán $12 \times 3 = 36$ perfiles, por lo que cada iteración requerirá evaluar 11.664 intercambios. El algoritmo intenta maximizar $|X'X|$ en vez de minimizar $|(X'X)^{-1}|$ (véase que $|(X'X)^{-1}| = |X'X|^{-1}$). Cada intercambio requiere calcular el determinante de la matriz $X'X$. Sin embargo, no partimos de cero cuando se calcula el determinante para cada intercambio, ya que $|X'X + x'x| = |X'X| |I + x(X'X)^{-1}x'|$.


¹Para más información sobre el algoritmo original acudir a Fedorov (1972). Para ampliar información sobre el algoritmo de Fedorov modificado ir a Cook y Nachtsheim (1980).

Apéndice E

Encuesta

En este apéndice incluimos el formulario google generado en base al diseño obtenido en SAS. Con motivo de cuadrar mejor las imágenes hemos excluido la sección, para cada pregunta, donde se permite marcar la opción escogida.

Inventario Forestal Continuo de Galicia

 No compartido



Cuestionario sobre factores en aprovechamientos forestales

Este cuestionario es anónimo.

No hay respuestas correctas o incorrectas. Por lo tanto, le pedimos que responda de manera espontánea y honesta a todas las preguntas.

Puede retirarse de la encuesta libremente y en cualquier momento si lo desea.

Gracias por su colaboración.

Aclaraciones

Este cuestionario está formado por 12 preguntas, en cada una figuran 3 opciones que se corresponden con parcelas diferentes. Tanto las parcelas mostradas como las colindantes forman parte de un entorno forestal. Se asume que todas las parcelas tienen la misma tipología de madera y mismas existencias por hectárea.

Para cada opción elegible hay una descripción de 5 variables acompañadas por un ejemplo ilustrativo de una parcela concreta. Las variables en cuestión son:

- Área: superficie de la parcela, medida en hectáreas (los ejemplos no están representados a la misma escala).
- Distancia a vías: indica la distancia promedio de la parcela al cargadero más cercano, en metros.
- Pendiente: representa la pendiente más frecuente en la parcela, indicada con porcentaje.
- Forma: describe la forma de la parcela.
- Fragmentación: esta variable se refiere al tamaño de las parcelas colindantes. Si la fragmentación es baja, las parcelas cercanas son mayores; si la fragmentación es alta, habrá muchas parcelas de muy reducido tamaño.

En cada pregunta debe seleccionar aquella opción que preferiría para realizar el aprovechamiento forestal.

1. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?
 Seleccione la opción escogida:

OPCIÓN 1

Área: mayor de 5 ha
 Distancia a vías: más de 500 m
 Pendiente: entre el 20 y el 40%
 Forma: irregular
 Fragmentación: alta
 Ejemplo:



Área: 13 ha
 Distancia a vías: 598 m
 Pendiente: 35%

OPCIÓN 2

Área: menor de 0,5 ha
 Distancia a vías: menos de 100 m
 Pendiente: entre el 20 y el 40%
 Forma: regular
 Fragmentación: baja
 Ejemplo:



Área: 0,17 ha
 Distancia a vías: 21,42 m
 Pendiente: 22,09%

OPCION 3

Área: entre 0,5 y 1 ha
 Distancia a vías: más de 500 m
 Pendiente: menos del 20%
 Forma: muy irregular
 Fragmentación: media
 Ejemplo:



Área: 0,92 ha
 Distancia a vías: 600,51 m
 Pendiente: 10,56%

Los ejemplos están representados a escalas diferentes

2. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?
 Seleccione la opción escogida:

OPCIÓN 1

Área: entre 0,5 y 1 ha
 Distancia a vías: menos de 100 m
 Pendiente: menos del 20%
 Forma: regular
 Fragmentación: alta
 Ejemplo:



Área: 0,72 ha
 Distancia a vías: 49,24 m
 Pendiente: 6,12%

OPCIÓN 2

Área: menos de 0,5 ha
 Distancia a vías: entre 100 y 500 m
 Pendiente: entre el 20 y el 40%
 Forma: irregular
 Fragmentación: media
 Ejemplo:



Área: 0,17 ha
 Distancia a vías: 180,79 m
 Pendiente: 30,55%

OPCION 3

Área: entre 0,5 y 1 ha
 Distancia a vías: más de 500 m
 Pendiente: más del 40%
 Forma: muy irregular
 Fragmentación: baja
 Ejemplo:



Área: 0,79 ha
 Distancia a vías: 923,15 m
 Pendiente: 60,98%

Los ejemplos están representados a escalas diferentes

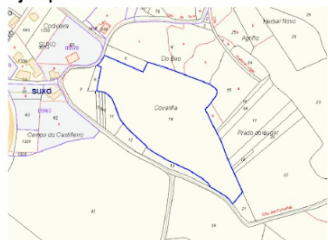
3. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?

Seleccione la opción escogida:

OPCIÓN 1

Área: entre 1 y 5 ha
Distancia a vías: menos de 100 m
Pendiente: menos del 20%
Forma: irregular
Fragmentación: media

Ejemplo:



Área: 2,66 ha
Distancia a vías: 60,93 m
Pendiente: 18,50%

OPCIÓN 2

Área: entre 0,5 y 1 ha
Distancia a vías: menos de 100 m
Pendiente: entre el 20 y el 40%
Forma: regular
Fragmentación: baja

Ejemplo:



Área: 0,62 ha
Distancia a vías: 28,06 m
Pendiente: 28,09%

OPCIÓN 3

Área: menos de 0,5 ha
Distancia a vías: entre 100 y 500 m
Pendiente: más de 40%
Forma: irregular
Fragmentación: alta

Ejemplo:



Área: 0,24 ha
Distancia a vías: 118,59 m
Pendiente: 52,8%

Los ejemplos están representados a escalas diferentes

4. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?

Seleccione la opción escogida:

OPCIÓN 1

Área: más de 5 ha
Distancia a vías: más de 500 m
Pendiente: menos del 20%
Forma: irregular
Fragmentación: baja

Ejemplo:



Área: 5,13 ha
Distancia a vías: 825,72 ha
Pendiente: 19,87%

OPCIÓN 2

Área: entre 1 y 5 ha
Distancia a vías: entre 100 y 500 m
Pendiente: más del 40%
Forma: regular
Fragmentación: media

Ejemplo:

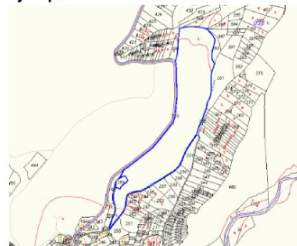


Área: 1,59 ha
Distancia a vías: 131,39
Pendiente: 42,87%

OPCIÓN 3

Área: más de 5 ha
Distancia a vías: menos de 100 m
Pendiente: entre el 20 y el 40%
Forma: muy irregular
Fragmentación: alta

Ejemplo:



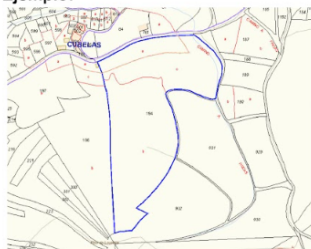
Área: 5,64 ha
Distancia a vías: 55,15 m
Pendiente: 33,94%

Los ejemplos están representados a escalas diferentes

5. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?
 Seleccione la opción escogida:

OPCIÓN 1

Área: más de 5 ha
 Distancia a vías: menos de 100 m
 Pendiente: menos del 20%
 Forma: irregular
 Fragmentación: media
Ejemplo:



Área: 6,96 ha
 Distancia a vías: 57,40 m
 Pendiente: 14,62%

OPCIÓN 2

Área: entre 1 y 5 ha
 Distancia a vías: más de 500 m
 Pendiente: entre el 20 y el 40%
 Forma: irregular
 Fragmentación: alta
Ejemplo:



Área: 1,72 ha
 Distancia a vías: 553,23 m
 Pendiente: 24,32%

OPCION 3

Área: menos de 0,5 ha
 Distancia a vías: más de 500 m
 Pendiente: más del 40%
 Forma: muy irregular
 Fragmentación: baja
Ejemplo:



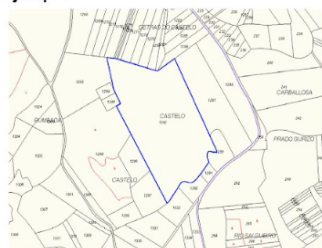
Área: 0,32 ha
 Distancia a vías: 520,89 m
 Pendiente: 40,73%

Los ejemplos están representados a escalas diferentes

6. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?
 Seleccione la opción escogida:

OPCIÓN 1

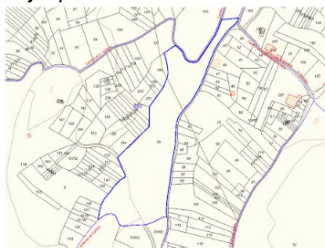
Área: más de 5 ha
 Distancia a vías: entre 100 y 500 m
 Pendiente: más del 40%
 Forma: regular
 Fragmentación: media
Ejemplo:



Área: 5,68 ha
 Distancia a vías: 145,32 m
 Pendiente: 55,93%

OPCIÓN 2

Área: entre 1 y 5 ha
 Distancia a vías: menos de 100 m
 Pendiente: entre el 20 y el 40%
 Forma: muy irregular
 Fragmentación: alta
Ejemplo:



Área: 2,80 ha
 Distancia a vías: 22,28 m
 Pendiente: 32,80%

OPCION 3

Área: menos de 0,5 ha
 Distancia a vías: menos de 100 m
 Pendiente: más del 40%
 Forma: irregular
 Fragmentación: baja
Ejemplo:



Área: 0,11 ha
 Distancia a vías: 67,05 m
 Pendiente: 41,96%

Los ejemplos están representados a escalas diferentes

7. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?

Seleccione la opción escogida:

OPCIÓN 1

Área: entre 1 y 5 ha
Distancia a vías: entre 100 y 500 m
Pendiente: menos del 20%
Forma: regular
Fragmentación: baja
Ejemplo:



Área: 2,37 ha
Distancia a vías: 155,22 m
Pendiente: 19,87%

OPCIÓN 2

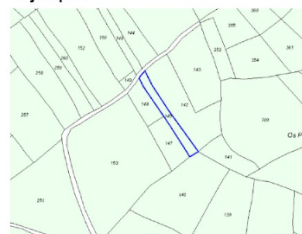
Área: entre 0,5 y 1 ha
Distancia a vías: entre 100 y 500 m
Pendiente: más del 40%
Forma: irregular
Fragmentación: alta
Ejemplo:



Área: 0,98 ha
Distancia a vías: 108,74 m
Pendiente: 41,72%

OPCION 3

Área: menos de 0,5 ha
Distancia a vías: más de 500 m
Pendiente: menos del 20%
Forma: muy irregular
Fragmentación: media
Ejemplo:



Área: 0,08 ha
Distancia a vías: 502,91 m
Pendiente: 11,08%

Los ejemplos están representados a escalas diferentes

8. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?

Seleccione la opción escogida:

OPCIÓN 1

Área: entre 0,5 y 1 ha
Distancia a vías: más de 500 m
Pendiente: entre el 20 y el 40%
Forma: regular
Fragmentación: media
Ejemplo:



Área: 0,75 ha
Distancia a vías: 614,76 m
Pendiente: 35,90%

OPCIÓN 2

Área: más de 5 ha
Distancia a vías: entre 100 y 500 m
Pendiente: menos del 20%
Forma: regular
Fragmentación: baja
Ejemplo:



Área: 8 ha
Distancia a vías: 103,78 m
Pendiente: 12,50%

OPCION 3

Área: entre 1 y 5 ha
Distancia a vías: menos de 100 m
Pendiente: más del 40%
Forma: muy irregular
Fragmentación: media
Ejemplo:



Área: 2,98 ha
Distancia a vías: 40,93 m
Pendiente: 40,01%

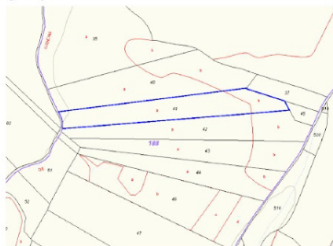
Los ejemplos están representados a escalas diferentes

9. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?

Seleccione la opción escogida:

OPCIÓN 1

Área: entre 1 y 5 ha
Distancia a vías: entre 100 y 500 m
Pendiente: entre el 20 y el 40%
Forma: muy irregular
Fragmentación: baja
Ejemplo:



Área: 2,22 ha
Distancia a vías: 354,12 m
Pendiente: 29,39%

OPCIÓN 2

Área: más de 5 ha
Distancia a vías: más de 500 m
Pendiente: más del 40%
Forma: regular
Fragmentación: alta
Ejemplo:



Área: 5,44 ha
Distancia a vías: 531,68 m
Pendiente: 40,33%

OPCION 3

Área: entre 0,5 y 1 ha
Distancia a vías: entre 100 y 500 m
Pendiente: entre el 20 y el 40%
Forma: irregular
Fragmentación: media
Ejemplo:



Área: 0,64 ha
Distancia a vías: 234,22 ha
Pendiente: 38,62%

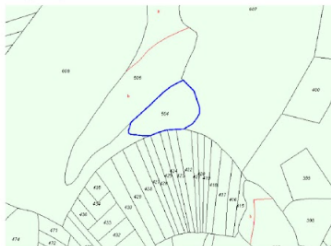
Los ejemplos están representados a escalas diferentes

10. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?

Seleccione la opción escogida:

OPCIÓN 1

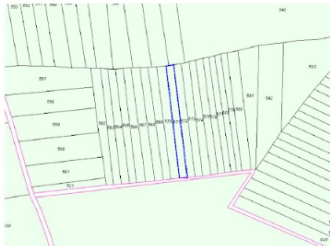
Área: menos de 0,5 ha
Distancia a vías: más de 500 m
Pendiente: entre el 20 y el 40%
Forma: regular
Fragmentación: media
Ejemplo:



Área: 0,21 ha
Distancia a vías: 609,54 m
Pendiente: 38,49%

OPCIÓN 2

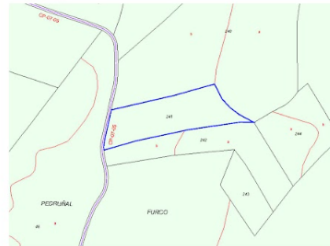
Área: menos de 0,5 ha
Distancia a vías: entre 100 y 500 m
Pendiente: menos del 20%
Forma: muy irregular
Fragmentación: alta
Ejemplo:



Área: 0,02 ha
Distancia a vías: 107,96 m
Pendiente: 19,14%

OPCION 3

Área: entre 0,5 y 1 ha
Distancia a vías: menos de 100 m
Pendiente: más del 40%
Forma: irregular
Fragmentación: baja
Ejemplo:



Área: 0,66 ha
Distancia a vías: 70,09 m
Pendiente: 47,39%

Los ejemplos están representados a escalas diferentes

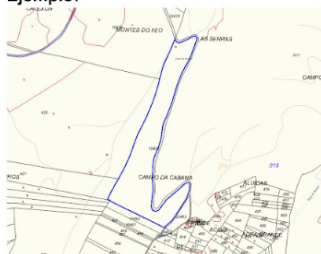
11. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?

Seleccione la opción escogida:

OPCIÓN 1

Área: más de 5 ha
Distancia a vías: menos de 100 m
Pendiente: más del 40%
Forma: muy irregular
Fragmentación: media

Ejemplo:



Área: 6,36 ha
Distancia a vías: 31,97 m
Pendiente: 45,59%

OPCIÓN 2

Área: entre 1 y 5 ha
Distancia a vías: más de 500 m
Pendiente: menos del 20%
Forma: irregular
Fragmentación: baja

Ejemplo:

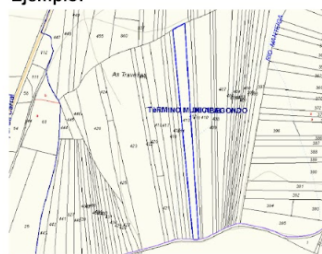


Área: 3,39 ha
Distancia a vías: 599,60 m
Pendiente: 18,15%

OPCIÓN 3

Área: entre 0,5 y 1 ha
Distancia a vías: entre 100 y 500 m
Pendiente: menos del 20%
Forma: muy irregular
Fragmentación: alta

Ejemplo:



Área: 0,92 ha
Distancia a vías: 221,36 m
Pendiente: 14,67%

Los ejemplos están representados a escalas diferentes

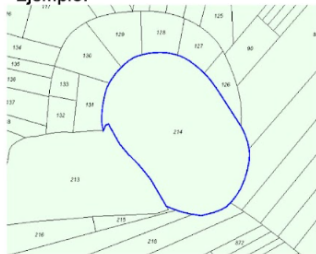
12. ¿Cuál de las siguientes parcelas prefiere para realizar el aprovechamiento forestal?

Seleccione la opción escogida:

OPCIÓN 1

Área: entre 1 y 5 ha
Distancia a vías: más de 500 m
Pendiente: más del 40%
Forma: regular
Fragmentación: alta

Ejemplo:



Área: 1,59 ha
Distancia a vías: 567,34 m
Pendiente: 45,53%

OPCIÓN 2

Área: más de 5 ha
Distancia a vías: entre 100 y 500 m
Pendiente: entre el 20 y el 40%
Forma: muy irregular
Fragmentación: baja

Ejemplo:

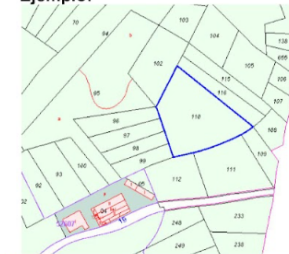


Área: 5,38 ha
Distancia a vías: 254,19 m
Pendiente: 31,27%

OPCIÓN 3

Área: menos de 0,5 ha
Distancia a vías: menos de 100 m
Pendiente: menos del 20%
Forma: regular
Fragmentación: alta

Ejemplo:



Área: 0,08 ha
Distancia a vías: 31,74 m
Pendiente: 17,70%

Los ejemplos están representados a escalas diferentes

Bibliografía

- [1] Anderson, N. H. (1970). Functional measurement and psychological judgment. *Psychological Review*, 77, 153-170.
- [2] Área de Inventario y Estadísticas Forestales. (2011, Noviembre). *Cuarto Inventario Forestal Nacional: Galicia*. Dirección General de Medio Natural y Política Forestal y Ministerio de Medio Ambiente, y Medio Rural y Marino. https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/galicia_tcm30-531786.pdf
- [3] Baarsma, B. (2003). The valuation of the IJmeer nature reserve using conjoint analysis. *Environmental and Resource Economics*, 25(3), 343-356.
- [4] Balboa López, X.L., B. Besteiro Rodríguez, X. Fernández Leiceaga, L. Fernández Prieto, M. Jordán Rodríguez, E. López Iglesias, D. Soto Fernández, y P. Viso Outeiriño. (2013). La política forestal gallega en los montes vecinales en mano común. *Ambienta*, 104, 114-125.
- [5] Corstjens, M. L., y Gautschi, D. A. (1983). A comparative analysis of specification tests for the utility function. *Management Science*, 29(12), 1393-1413.
- [6] Crecente Maseda, R., Rodríguez Vicente, V., y Marey Pérez, M.F. (2007). Perfil del propietario forestal individual en Galicia: Objetivos y prácticas de gestión en el noreste de la comunidad. *Revista Galega de Economía*, 16(1), 47-70.
- [7] DeSarbo, W. S., Ramaswamy, V., y Cohen, S. H. (1995). Market segmentation with choice-based conjoint study. *Marketing Letters*, 6(2), 137-147.
- [8] Dufhues, T., Heidhues, F., y Buchenrieder, G. (2004). Participatory product design by using conjoint analysis in the rural financial market of Northern Vietnam. *Asian Economic Journal*, 18(1), 81-114.
- [9] Finn, A., y Louviere, J. J. (1992). Determining the appropriate response to evidence of public concern: The case of food safety. *Journal of Public Policy and Marketing*, 11, 2-25.
- [10] Fishbein, M. (1967). *Readings in attitude theory and measurement*. John Wiley and Sons Inc.
- [11] Green, P. E., y Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, 8, 355-363.
- [12] Green, P. E., y Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, 5, 103-123.
- [13] Goldberg, S. M., Green, P. E., y Wind, Y. (1984). Conjoint analysis of price premiums for hotel amenities. *Journal of Business*, 57(1), 111-132.
- [14] Igarashi, A., Nakano, Y., y Yoneyama-Hirozane, M. (2022). Public preferences and willingness to accept a hypothetical vaccine to prevent a pandemic in Japan: a conjoint analysis. *Expert Review of Vaccines*, 21(2), 241-248.

- [15] Instituto Galego de Estatística. (2021, Febrero). *Análise da cadea forestal-madeira*. http://www.ige.eu/estatico/pdfs/s3/publicaciones/AIE_Analise_Cadea_Forestal.pdf
- [16] Karniouchina, E. V., Moore, W. L., van der Rhee, B., y Verma, R. (2009). Issues in the use of ratings-based versus choice-based conjoint analysis in operations management research. *European Journal of Operational Research*, 197, 340-348.
- [17] Ley 7/2012, de 28 de junio, de montes de Galicia (BOE núm. 217, de 08 de septiembre de 2012).
- [18] Luce, D., y Tukey, J. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 1, 1-27.
- [19] Ministerio para la Transición Ecológica y el Reto Demográfico. (2021). *Anuario de Estadística Forestal 2019*. https://www.miteco.gob.es/es/biodiversidad/estadisticas/aef2019_completo_estandar_tcm30-534526.pdf
- [20] Molina, F. (1997). *Condicionantes estructurales para el aprovechamiento del monte en Galicia y Portugal*. Congreso Empresarial Técnico-Científico Forestal del Arco Atlántico. Pontevedra.
- [21] Ong, A.K.S., Yogi, Y.T., Chuenyindee, T., Young, M.N., Doma, B.T., Caballes, D.G., Centeno, R.S., Morfe, A.S. y Bautista, C.S. (2022). Preference analysis on the online learning attributes among senior high school students during the COVID-19 pandemic: A conjoint analysis approach. *Evaluation and Program Planning*, 92, 102100.
- [22] Ramantswana, T., Cheruiyot, K. y Azasu, S. (2020). Corporate Headquarters Location Preferences in the South African Metropolitan Areas: A Conjoint Analysis Approach. *Journal of Real Estate Literature*, 27(2), 227-244.
- [23] Rao Vithala, R. (2014). *Applied Conjoint Analysis*. Springer.
- [24] Sattler, H. y Hensel-Börner, S. (2007). A comparison of conjoint measurement with self-explicated approaches. *Conjoint Measurement: Methods and Applications*; Gustafsson, A., Herrmann, A., Huber, F., Eds. Springer: Berlin, Germany, pp 121-133.
- [25] Sawtooth Software, Inc. (2007). ACA System for Adaptive Conjoint Analysis. <https://sawtoothsoftware.com/resources/technical-papers/aca-technical-paper>.
- [26] Skjoldborg, U. S., y Gyrd-Hansen, D. (2003). Conjoint analysis. The cost variable: An Achilles heel? *Health Economics*, 12, 479-491.
- [27] Tariq, A. (2018) Pilot Study on the Use of Adaptive Conjoint Analysis to Assess Risk Factors for Highly Pathogenic Avian Influenza out breaks on Commercial Poultry Holdings of Pakistan. *Appro Poult Dairy and Vet Sci*. 4(1).