UNIVERSIDAD SANTO TOMAS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA

**Facultad de Estadística**
Trabajo de Grado
Junio 2019

# Análisis conjunto mediante modelos lineales jerárquicos y modelos jerárquicos Bayesianos. Una aproximación mediante análisis multivariado

## Conjoint analysis using hierarchical linear models and hierarchical Bayesian models. An approach using multivariate analysis

Mauricio Alarcón Granados[a]
mauricio.alarcon@usantotomas.edu.co

Director: Wilmer Pineda Ríos[b]
wilmerpineda@usantotomas.edu.co

**Resumen**

Con el fin de evaluar la estrategia de mercado a seguir en el relanzamiento de un producto financiero vigente en el mercado, la compañía consultó a sus clientes actuales y potenciales a través de una encuesta, por el producto financiero ideal. Los resultados fueron recogidos, procesados y analizados mediante al análisis conjunto o Conjoint Analysis. La primera evaluación se realiza mediante el análisis conjunto para datos ordenados, tradicionalmente utilizado en la investigación de mercados. Seguidamente, y con el objeto de evaluar los resultados a través de distintas metodologías como el análisis conjunto lineal, análisis conjunto lineal jerárquico y análisis conjunto jerárquico bayesiano se realiza la transformación de los datos tipo RANKING a datos tipo SCORE utilizando el análisis de homogeneidad.

En la primer parte de este documento se caracterizará la problemática a desarrollar y se abordará el estado del arte de las distintas metodologías mencionadas. Posteriormente, se realiza el análisis descriptivo de los datos para dar paso a los resultados obtenidos y las conclusiones de los mismos.

***Palabras clave*:** Conjoint, Análisis de homogeneidad, Datos ordenados .

**Abstract**

In order to evaluate the market strategy to be followed in the relaunch of a current financial product in the market, the company consulted its current and potential customers through a survey of the ideal financial product. The results were collected, processed and analyzed through Análisis Conjunto or Conjoint Analysis. The first evaluation is carried out through a conjoint analysis for ordered data, traditionally used in market research. Next, and in order to evaluate the results through different methodologies such as linear conjoint analysis, hierarchical linear conjoint analysis and Bayesian hierarchical conjoint analysis, the transformation of RANKING type data to SCORE type data is performed using homogeneity analysis.

In the first part of this document the problem to be developed will be characterized and the state of the art of the different methodologies mentioned will be addressed. Subsequently, the descriptive analysis of the data is carried out to give way to the results obtained and the conclusions thereof.

***Keywords*:** Conjoint, Homogeneity Analysis, Ordered data.

[a]Economista. Estudiante de Estadística. Universidad Santo Tomás, sede Bogotá
[b]Docente. Facultad de Estadistica. Universidad Santo Tomás, sede Bogotá

# 1. Introducción

El análisis conjunto (Conjoint) es un método estadístico ampliamente reconocido por diferentes disciplinas como la psicología, el marketing, la publicidad, el diseño y la economía debido a la enorme utilidad y a la facilidad con la que se puede adaptar esta metodología a diferentes escenarios en distintos campos del conocimiento. Precisamente, es su adaptabilidad, la que ha llamado la atención de la compañía contratante del estudio, para identificar aquellas características que más valoran sus clientes en un producto financiero.

Existen estudios que evaluan la pérdida de información en el uso de modelos conjoint para datos tipo ranking, en los que se solicita al encuestado, evaluar y organizar en orden de preferencia los productos observados. Edlira Shehu (2005) realiza una investigación de los supuestos generales alrededor del análisis Conjoint basado en ranking para explicar los problemas de la pérdida de información por el uso de métodos de rankeo. En especial, da cuenta de las dudosas estimaciones de las utilidades (part-worth) obetnidas por esta metodología.

Autores como Maydeu-Olivares  Bockenholt destacan como los métodos de Rankeo y los de comparación por pares gozan de cierta popularidad debido a que imponen menos restricciones a las respuestas de los entrevistados, aún cuando las diferencias entre los distintos perfiles o alternativas evaluadas es mínima, estos métodos pueden llegar a proveer más información acerca de las preferencias individuales frente a los métodos de puntuación (rating).

Este documento es el resultado de la apuesta por simplificar los procedimientos y cálculos comúnmente utilizados en el análisis conjunto sobre los datos rankeados puesto que como señala (Lam, Koning Franses, 2010) los modelos para ranking por lo general resultan mucho más complejos. En este sentido también los autores hacen referencia al uso de escalamiento multidimensional como técnica para derivar las preferencias de los resultados. Es por ello que se recurre al uso de métodos multivariados como el análisis de homogeneidad o de correspondencias múltiples para transformar los datos tipo RANKING en datos tipo SCORE para a partir de ellos realizar un análisis conjunto lineal, una análisis conjunto lineal jerárquico y finalmente un modelo conjunto jerárquico bayesiano.

En la primer parte de este documento se aborda el marco teórico, en él se realiza la descripción metodológica de las técnicas, métodos, modelos y procedimientos que configuran la propuesta de estudio central. Entre los temas que serán abordados se encuentra el análisis conjunto, el análisis de homogeneidad, el modelo lineal jerárquico, la fundamentación de la teoría bayesiana y el modelo jerárquico bayesiano. Posteriormente se procede con la descripción de los datos utilizados y el analisis sociodemografico de los mismos. Por último se realiza el análisis de los resultados obtenidos de los modelos mencionados mediante la transformación de los datos, utilizando el análisis de homogeneidad. Los resultados del modelo lineal jerárquico frecuentista y los resultados del modelo jerárquico bayesiano serán presentados en dicha sección.

# 2. Marco Teórico

## 2.1. Análisis Conjoint

El Conjoint o análisis conjunto es una metodología estadística que nace en los años 70 como resultado de la propuesta metodológica hecha por Paul Green de la Universidad de Pensilvania basada en el trabajo de Luce  Tukey (1964) al preguntar sobre los productos compuestos para medir atributos. Más tarde en los años 80, Jordán Louveniere de la Universidad de Iowa propone una mejora a los postulados de

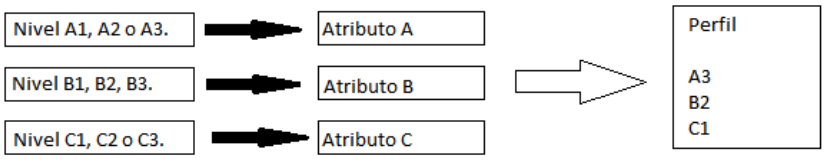
Green lo que significa el desarrollo de una técnica que da lugar a las comparaciones entre diferentes productos en lugar de medir la preferencia del consumidor mediante escalas numéricas como lo propuso Green inicialmente.

Más tarde, una década después el Conjoint cobra una mayor relevancia en el campo de la investigación con los nuevos desarrollos computacionales y la llegada de mejores procesadores informáticos que mejoran la eficiencia de los cálculos, que hasta el momento resultaban ser medianamente complejos. Así las cosas, el Conjoint recibió gran aceptación de profesionales de distintas ramas del conocimiento como economistas, psicólogos, mercaderistas y sobretodo de los investigadores de mercado.

La principal razón de esta aceptación, de acuerdo con McCullough (1979), siguiendo a Johnson(1975), puede rastrearse hasta la capacidad de la medición Conjoint para producir resultados relativamente "sofisticados", a partir de datos que los autores mencionados califican como "primitivos", que normalmente consisten en preferencias ordenadas o datos de utilidades. En este sentido McCullough afirma que la técnica "permite al investigador hacer predicciones del comportamiento de elecciones individuales".

Para entender mejor el análisis conjunto, es preciso definir algunos términos que a lo largo de este documento usaremos con frecuencia: El primero de ellos es el término atributo, el cual hace referencia a las características generales del producto (ej. Color, Tamaño, Sabor). Los niveles, son entonces las diferentes alternativas que puede tomar un atributo (ej. Color: Rojo, Negro, Azul). Y el perfil, se refiere a la combinación de distintos niveles de atributos que configuran lo que conocemos como un producto, (ej. Un bebida de Color rojo, tamaño 330ml y sabor a manzana). Lo anterior se resume a continuación:

Figura 1: Niveles, atributos y perfiles.




**Fuente:** Propia basada en (*Ramírez-Hurtado. 2010*)

El objetivo principal del Conjoint es medir el valor que da un consumidor a uno o un grupo de atributos que conforman un producto o servicio. Por lo general y sobretodo en la investigación de mercados el Conjoint es utilizado para determinar la importancia de distintas variables o características de un producto buscando identificar la dinámica del mercado ante variaciones en dichos atributos. Inclusive el análisis conjoint permite observar cambios en la demanda ante variaciones de precio. Al respecto Orme,B (2010) sobre el análisis conjunto manifiesta que : "La característica clave del análisis conjunto es que los encuestados evalúan perfiles de productos compuestos de múltiples elementos (atributos o características). En función de cómo los encuestados evalúan los elementos combinados (los conceptos del producto), deducimos los puntajes de preferencia que podrían haber asignado a los componentes individuales del producto que habrían resultado en esas evaluaciones generales" (p.29).

Existen varias metodologías de análisis conjunto, entre los más conocidos se encuentra el análisis conjunto de perfil completo (Full-Profile Conjoint), el análisis conjunto basado en elecciones también conocido como Choice Based Conjoint (CBC), el análisis conjunto adaptativo (Adaptive Conjoint Analysis, ACA):

- **Análisis Conjunto de Perfil Completo:** Este análisis cosiste en ordenar o puntuar un grupo de perfiles según la preferecia de cada individuo.

- **Análisis Conjunto Basado en Elecciones:** Es tal vez el tipo de Conjoint más popular debido a su capacidad de imitar una situación de compra real para los consumidores llevándolos a elegir distintos productos con ciertas características.
- **Análisis Conjunto Adaptativo:** Como su nombre indica, el análisis conjunto adaptativo cambia los conjuntos de opciones que son presentados a los encuestados en función de sus preferencias. Esta adaptación está enfocada en los gustos y preferencias del encuestado hacia las características y los niveles presentados. Esta metodología se considera una alternativa del análisis de Perfiles Completos.

A continuación se muestran las diferencias entre el Análisis Conjunto de Perfil Completo y el CBC:

Figura 2: Diferencias entre el Análisis Conjunto tradicional y el Análisis conjunto basado en la Elección

| | **Análisis Conjunto tradicional** (perfiles completos) | **Análisis Conjunto Basado en la Elección** |
|---|---|---|
| **Base teórica comportamental** | ? | RUT |
| **Método de recogida de datos** | Ordenación o puntuación de perfiles | Elección de la alternativa más preferida |
| **Forma del modelo** | Aditivo (efectos principales) | Aditivo + interacciones |
| **Método de estimación** | Regresión OLS | Logit o Probit Multinomial |
| **datos producidos** | A nivel individual | A nivel agregado (Bayesiano a nivel individual) |

**Fuente:** Fuente: Picón, E., Varela, J., Braña, T. 2006.

### 2.1.1. Análisis Conjunto tradicional (Perfiles Completos)

Como se ha mencionado previamente, el análisis conjunto busca determinar las utilidades parciales (part-worth) o coeficientes $\beta$ para todos los factores evaluados teniendo en cuenta los datos rankeados. Para ello el modelo se define de la siguiente manera:

$$y_k = \sum_{j=1}^{J} \sum_{m=1}^{Mj} \beta_{jm} x_{jm} \tag{1}$$

Donde, $y_k$ es la utilidad total estimada para el perfil k. $\beta_{jm}$ es la utilidad parcial del atributo $m$ en el nivel $j$ y $x_{jm}$ toma valor 1 si el perfil k contiene valores en la categoría m del nivel j y será 0 en otro caso.

## 2.2. Análisis de homogeneidad

De acuerdo Tenenhaus  Young (1985), diferentes países han llamado de distintas maneras una misma metodología. Por ejemplo la literatura Estadounidense le llama Escalamiento Óptimo, Canadá le conoce como Escalamiento Dual, en Alemania se refiere a él como el Análisis de Homogeneidad pero tal vez el más popularizado es la propuesta francesa de llamarlo Análisis de Correspondencia Múltiple.

Estos métodos son comúnmente utilizados para la cuantificación de datos categóricos multivalentes. En esencia el análisis de correspondencias múltiples es un método para maximizar la homogeneidad de un número de variables(Van Der Burg, De Leeuw, Verdegaal, 1988). Siguiendo a Van Der Burg et.al para

definir el análisis de homogeneidad es preciso conocer alguna notación importante. Supóngase que se tiene una matriz multivariante de datos de tamaño $n \times m$, donde las columnas corresponden a las variables y las filas a los objetos. Se asume que la variable $j$ toma $k_j$ valores o categorías diferentes y se define la matriz $G_j$ como la matriz indicadora de tamaño $n \times k$ cuya función es indicar precisamente cuáles categorías son puntuadas por cuales objetos.

Usualmente, el análisis de datos categóricos por cuantificación es hecha con una expresión de criterio, configurada para que la cuantificación se ajuste al propósito analítico (Okamoto, 2015).En este sentido, retomando a Van Der Burg et.al el análisis de homogeneidad determina las cuantificaciones o transformaciones de las categorías de cada variable de tal forma que la homogeneidad es maximizada. Para tal fin se busca minimizar la siguiente función conocida como función de pérdida.

$$Min\sigma(X, Y) = \frac{1}{m} \sum_{j=1}^{m} SSQ(X - G_j Y_j) on_i \tag{2}$$

Sujeto a que $X'X = nI$ y $u'X = 0$ donde $u$ es una columna con $n$ elementos iguales a 1, SSQ es la sumatoria de los cuadrados de los elementos de un vector o matriz y los elementos de X son llamados objetos de puntuación.

## 2.3. Modelo Lineal Jerárquico

Los modelos lineales jerárquicos (HLM) son una forma compleja de regresión de mínimos cuadrados ordinarios (OLS) que se utiliza para analizar la varianza en las variables dependientes cuando las variables predictoras o independientes se encuentran en diferentes niveles jerárquicos. (Woltman, Feldstain, MacKay, Rocchi, 2012, p.52). Estos modelos también son conocido como modelos multinivel debido a que los individuos y la muestra total están en diferentes niveles.

Existen varios algoritmos propuestos para la estimación de este tipo de modelos, por lo general se utiliza un algoritmo en el que el modelo se ajusta a todos los datos, para luego ajustar el modelo a cada observación.

A manera general este modelo requiere que el conjunto de datos este conformado por múltiples observaciones para cada individuo.

Los modelos jerárquicos identifican dos tipos de efectos. El primero se conoce como efecto fijo, el cual, al igual que en los modelos lineales son efectos estáticos para todos los individuos respondientes. Los segundos son los efectos aleatorios, que deben su nombre debido a que son estimados como variables aleatorias que siguen una distribución alrededor de los efectos fijos estimados. A su vez los modelos multinivel cuentan con un subconjunto de modelos conocidos como los modelos de efectos mixtos que deben su nombre a que combinan los efectos fijos y los efectos aleatorios para cada respondiente.

De los modelos de efectos mixtos se desprenden los modelos anidados o nested models en los que el factor de interés únicamente sucede entre los subgrupos de la muestra general.

La especificación matemática de estos modelos parte una regresión simple para cada individuo $i$ para lo cual,

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + r_{ij} \tag{3}$$

Donde, $Y_{ij}$ es la variable dependiente medida en el i-ésimo nivel 1 con el j-ésimo nivel 2. $X_{ij}$ es la observación individual para el nivel 1. Por supuesto, $r_{ij}$ es el error aleatorio asociado a la i-ésima unidad del nivel 1 con j-ésima unidad del segundo nivel que sigue una distribución normal con media cero y varianza $\sigma^2$. (Sullivan, Dukes  Losina, 1999).

$$E(r_{ij}) = 0; var(r_{ij}) = \sigma^2 \tag{4}$$

Entre tanto $\beta_{0j}$ y $\beta_{1j}$ corresponde al intercepto y el coeficiente de la regresión asociado a las unidades j-ésimas del segundo nivel.

En el modelo del segundo nivel, los coeficientes de regresión del nivel 1 ($\beta_{0j}$ y $\beta_{1j}$) son usadas como variables de resultado y se relacionan con cada predictor del nivel-2. (Woltman et.al, 2012, pp. 52-69)

$$\beta_{0j} = \gamma_{00} + \gamma_{01} G_j + U_{0j} \tag{5}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} G_j + U_{1j} \tag{6}$$

Donde, $\beta_{0j}$ es el intercepto para la unidad j-ésima del nivel-2. $\beta_{1j}$ es la pendiente para la unidad del j-ésimo unidad del nivel 2. $G_j$ es el valor del predictor en el nivel-2. $\gamma_{00}$ es la media total del intercepto ajustada para $G_j$ y $\gamma_{01}$ es el coeficiente de regresión asociado con G relativo al intercepto del nivel-1. $\gamma_{11}$ es el coeficiente de regresión asociado con G relativo a la pendiente del nivel-1. $U_{0j}$ y $U_{1j}$ corresponde a los efectos aleatorios para la j-ésima unidad del nivel-2 ajustado al intercepto y la pendiente respectivamente.

## 2.4. Fundamentos de la Teoría Bayesiana

El desarrollo estadístico de la corriente Bayesiana se origina hacia 1973 tras la publicación de un documento realizado por Thomas Bayes donde postula una idea revolucionaria para el campo matemático de la época al considerar que la probabilidad condicional de un evento A dado que sucedio un evento B está vinculado con la probabilidad de que ocurra el evento B dado un evento A. En términos matemáticos esto es:

$$\begin{aligned} P(A \mid B) &= \frac{P(B \mid A)P(A)}{P(B)} \\ &\propto P(B \mid A)P(A) \end{aligned} \tag{7}$$

donde $P(A)$ es la probabilidad a priori, $P(B \mid A)$ es la verosimilitud y $P(A \mid B)$ es la probabilidad a posteriori.

Antes de explicar de qué trata la probabilidad a priori, a posteriori, la verosimilitud y cómo trabajan en conjunto, es conveniente explicar cómo en la inferencia bayesiana existen dos distinciones importantes que se deben tener presentes. Lo primero es diferenciar las cantidades observables $X$ es decir los datos y las cantidades inobservables $\theta$ donde $\theta$ puede ser cualquier parámetro, una variable latente o información faltante. En el caso bayesiano los parámetros son tratados como variables aleatorias y sobre ellos es que se realiza la inferencia. Por el contrario en la estadística frecuentista los parámetros se consideran fijos y la inferencia se realiza sobre los datos.

### 2.4.1. Verosimilitud

La verosimilitud es una cantidad que sera usada para calcular la probabilidad posteriori. Al respecto Brewer (s.f) define la verosimilitud de la siguiente manera:

"La probabilidad de una hipótesis es la probabilidad de que usted hubiera observado los datos, si esa hipótesis fuera cierta. Los valores se pueden encontrar repasando cada hipótesis, imaginando que es cierta y preguntando: "¿Cuál es la probabilidad de obtener los datos que observé?"(p.13)

### 2.4.2. Distribuciones a priori

Dado que $\theta$ es desconocido se debe postular una distribución de probabilidad que refleje la incertidumbre que se tiene antes de ver los datos (Economic  Social Research Council - ESRC, s.f).
Por lo tanto se especifica la distribución a priori como $p(\theta)$ y siguiendo a ESRC por ende la distribución a priori expresa la incertidumbre acerca de $\theta$ antes de observar los datos.

### 2.4.3. Distribuciones a posteriori

Aquí es donde el teorema de bayes entra a formar parte esencial de la estimación bayesiana, puesto que permite obtener la distribución condicional para cantidades observadas dado los datos que conocemos. Como $X$ es conocida debe estar condicionada. De tal manera que la distribución posterior expresa la incertidumbre de $\theta$ después de observar los datos (ESRC, p.12)

### 2.4.4. Método de Cadenas de Markov Monte-Carlo

Las Cadenas de Markov Monte-Carlo (Markov Chain Monte-Carlo [MCMC]) son un método de muestreo computacional (Gamerman and Lopes, 2006; Gilks et al. 1996) utilizado en la inferencia Bayesiana para estimar la distribución posterior. Siguiendo a van Ravenzwaaij, D., Cassey, P.  Brown, S.D. Psychon Bull Rev (2018) al respecto afirman que el método MCMC:

"Permite caracterizar una distribución sin conocer todas las propiedades matemáticas de la distribución mediante el muestreo aleatorio de valores fuera de la distribución. Una ventaja particular de MCMC es que se puede usar para extraer muestras de distribuciones, incluso cuando todo lo que se sabe acerca de la distribución es cómo calcular la densidad para diferentes muestras".(pp. 143-154).

#### 2.4.4.1 Cadenas de Markov

Antes de describir los algoritmos MCMC es necesario abordar una definición importante de la teoría de las Cadenas de Markov.

- **Definicion 1:** Una cadena de Markov es una secuencia de variables aleatorias $X_1, ..., X_n$ tal que, dado el estado presente, el futuro y el pasado son independientes. Es decir, que la distribución condicional de $X_{n+1}$ en el futuro, depende únicamente del estado presente $X_n$ (Sharma, 2006, p.16).

$$Prob(X_{n+1} = x \mid X_1 = x_1, X_2 = x_2, ..., X_n = x_n) = Prob(X_{n+1} = x \mid X_n = x_n) \tag{8}$$

Por lo tanto, si la probabilidad de transición es independiente de $n$, se considera que la cadena es homogénea en el tiempo. Esta cadena se define especificando las probabilidades de transición de un estado a otro.(Sharma, 2006, p.16).

$$K(x, y) = Prob(X_{n+1} = y \mid X_n = x) \tag{9}$$

Esta probabilidad de transición, es válida para espacios continuos y discretos. En el caso de espacios discretos la probabilidad de transición es una matriz escrita como $K_{xy}$. Siguiendo a Sharma (2006), dado un espacio de estado, una cadena de Markov homogénea en el tiempo tiene una distribución estacionaria $\pi$ si:

$$\pi(y) = \int \pi(x)K(x, y)dx \tag{10}$$

A este respecto, una anotación importante para entender este planteamiento es: "Una cadena de Markov es irreducible si puede pasar de cualquier estado $x$ de un espacio de estado discreto a cualquier otro estado $y$ en un número finito de pasos, es decir, existe un entero n tal que $K_{xy}^n > 0$. Si una cadena teniendo una distribución estacionaria es irreducible, la distribución estacionaria es única y la cadena es recurrente positiva. Para una cadena aperiódica, positiva recurrente con distribución estacionaria , la distribución es limitante (distribución de equilibrio). Esto significa que si comenzamos con cualquier distribución inicial (un vector de fila que especifica la probabilidad sobre los estados de un espacio de estado discreto) y aplicamos el operador de transición $K$ (una matriz) muchas veces, la distribución final se aproximaba a la distribución estacionaria (un vector de fila )"(Sharma, p.17)

$$\lim_{n\to\infty} || \lambda K^n - \pi ||= 0 \tag{11}$$

Utilizando la ley de los grandes números la cual, para una cadena de Markov con una distribución estacionaria única, dice :

$$E_\pi[g(x)] = \int g(x)\pi(x)dx = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} g(x_i) \tag{12}$$

Las técnicas que utilizan esta propiedad que permite hacer estimaciones de Monte Carlo de cantidades específicas de una cadena de Markov, son conocidas como las Cadenas de Markov Monte Carlo - MCMC.

#### 2.4.4.2 Algoritmo Metropolis–Hastings

El algoritmo inventado por Metropolis, Rosenbluth, Rosenbluth, Teller and Teller (1953) fue generalizado más tarde por Hastings (1970). El algoritmo de Metropolis-Hastings (MH) simula muestras de una distribución de probabilidad haciendo uso de la función de densidad conjunta completa y propuestas de distribuciones (independientes) para cada una de las variables de interés.(Yildirim I, 2012b, pp.1-2).

Figura 3: Algoritmo Metropolis–Hastings

**Algorithm 1** Metropolis-Hastings algorithm

Initialize $x^{(0)} \sim q(x)$
**for** iteration $i = 1, 2, \ldots$ **do**
  Propose: $x^{cand} \sim q(x^{(i)}|x^{(i-1)})$
  Acceptance Probability:
    $\alpha(x^{cand}|x^{(i-1)}) = \min\{1, \frac{q(x^{(i-1)}|x^{cand})\pi(x^{cand})}{q(x^{cand}|x^{(i-1)})\pi(x^{(i-1)})}\}$
  $u \sim$ Uniform $(u; 0, 1)$
  **if** $u < \alpha$ **then**
    Accept the proposal: $x^{(i)} \leftarrow x^{cand}$
  **else**
    Reject the proposal: $x^{(i)} \leftarrow x^{(i-1)}$
  **end if**
**end for**

**Fuente:** Yildirim, 2012, p.2

El primer paso como describe la tabla, es dar un valor inicial de muestra para cada variable aleatoria. Posteriormente se crea un bucle o ciclo que se repetirá determinadas veces hasta que se cumpla la

condición asignada al bucle. Este ciclo comienza con proponer una muestra $x^{cand}$ a partir de $q(x^{(i)} \mid x^{i-1})$ donde $q$ es una distribución igualmente propuesta. Se establece la probabilidad de aceptación determinada por la función $\alpha(x^{(cand)} \mid x^{i-1})$ la cual tiene en cuenta la distribución propuesta anteriormente y la densidad conjunta $\pi$. Finalmente con una distribución uniforme se establece la regla de decisión, donde se acepta la muestra candidata con probabilidad $\alpha$ o se rechaza con probabilidad $1 - \alpha$.

#### 2.4.4.3 Algoritmo Muestreador de Gibbs

El muestreador de Gibbs es otro de los métodos MCMC mas populares. La idea en el muestreo de Gibbs es generar muestras posteriores barriendo cada variable (o bloque de variables) para muestrear desde su distribución condicional con las variables restantes fijas a sus valores actuales.(Yildirim I, 2012a, p.1)

Figura 4: Algoritmo Metropolis–Hastings

**Algorithm 1** Gibbs sampler

Initialize $x^{(0)} \sim q(x)$

**for** iteration $i = 1, 2, \ldots$ **do**

$x_1^{(i)} \sim p(X_1 = x_1 | X_2 = x_2^{(i-1)}, X_3 = x_3^{(i-1)}, \ldots, X_D = x_D^{(i-1)})$

$x_2^{(i)} \sim p(X_2 = x_2 | X_1 = x_1^{(i)}, X_3 = x_3^{(i-1)}, \ldots, X_D = x_D^{(i-1)})$

$\vdots$

$x_D^{(i)} \sim p(X_D = x_D | X_1 = x_1^{(i)}, X_2 = x_2^{(i)}, \ldots, X_D = x_{D-1}^{(i)})$

**end for**

**Fuente:** Yildirim, 2012, p.2

El algoritmo para este método considera un grupo de variables aleatorias $X_1, X_2, X_3$. Para comenzar, se establecen los valores iniciales $x_1^{(0)}$, $x_2^{(0)}$, $x_3^{(0)}$. Para cada iteración se realiza $x_1^{(i)} \sim p(X_1 = x_1 \mid X_2 = x_2^{(i-1)}, X_3 = x_3^{(i-1)})$. Este proceso se realiza hasta conseguir la convergencia, la cual se determina si los valores muestreado tienen la misma distribución que la verdadera distribución posterior.

## 2.5. Modelo Lineal Jerárquico Bayesiano

En años recientes la introducción de la estadística bayesiana a métodos comúnmente frecuentistas no ha sido ajena a los modelos de investigación de mercados en referencia al análisis conjunto.

En un documento titulado "A Hierarchical Bayes Model for Ranked Conjoint Data" escrito por Thorsten Teichert, quien se desempeña como profesor de la universidad de hamburgo, establece a través de una comparación entre las estimaciones de modelos de Mínimos Cuadrados Ordinarios (OLS), la técnica de programación lineal para el análisis multidimensional de preferencias (LINMAP, por sus siglas en inglés) y las estimaciones Jerárquicas Bayesianas (HB) que son estas últimas, es el más robusto eficiente y acertado método de estimación

Una de las desventajas de la estimación mediante métodos frecuentistas como los modelos Logit, OLS, LINMAP entre otros, es postular teorías de "utilidades constantes" en la que se establece que los consumidores son probabilísticos, es decir que al momento de realizar sus elecciones no escogen la alternativa de mayor utilidad sino que sus elecciones están condicionadas a ciertas probabilidades para cada alternativa. En otros términos, cada individuo tiene una probabilidad asignada de elegir una alternativa de acuerdo a una función de distribución establecida.(Factum Mercadotécnico, 2009).

Para dar tratamiento a este inconveniente frecuentista, la teoría bayesiana, mediante el análisis jerárquico considera que cada individuo es único y, como tal, cada uno asignará un porcentaje de su consumo a la marca, de acuerdo con su experiencia previa. La gran ventaja de este método de estimación, se debe a que

aplica dos tipos de análisis a los datos, uno a nivel de los individuos mediante una distribución normal multivariada y otra a nivel global sin llegar a considerar los parámetros estimados como constantes.

La estimación de los modelos jerárquicos Bayesianos utiliza probabilidades a priori y verosimilitudes de los datos para establecer las probabilidades posterior en un proceso iterativo de optimización mediante el uso de simulaciones de cadenas de Markov y Monte Carlo (MCMC). En el proceso, el conocimiento previo(a priori) de las características, preferencias individuales y las preferencias conjuntas de todos quienes participan en el estudio permite obtener a través de la jerarquización Bayesiana los puntajes de utilidades para cada individuo que mediante metodos clasicos no es posible.

El concepto matemático de este modelo parte del modelo que describe la variación en las respuestas y de las utilidades de los individuos alrededor de la población.(Park C, 2004, p.1093)

$$\begin{aligned} Y_i &= X\beta_i + \epsilon_i \\ \beta_i &= \theta z_i + \delta_i \end{aligned} \tag{13}$$

para $i = 1, 2, ..., n$.

Donde $\beta_i$ es un vector p- dimensional de partworths para el sujeto i-ésimo y describe la heterogeneidad de las utilidades mediante una regresión multivariada con covariables q-dimensionales, $z_i$ y una matriz de parámetros de la regresión de tamaño $p \times q$ denominada $\theta$.

Los términos de error $\epsilon_i$ y $\delta_i$ son asumidos mutuamente independientes de una distribución normal,

$$\varepsilon_i \sim Np(0, \sigma_i I) \tag{14}$$

$$\delta_i \sim Np(0, \wedge) \tag{15}$$

donde $I$ es la matriz identidad y $\wedge$ es una matriz definida positiva de tamaño $p \times p$.

# 3. Descripción de los datos

Los datos corresponden a 409 individuos que respondieron la encuesta. La recolección de los datos se realizó mediante un diseño muestral no probabilístico por cuotas en Bogotá, Cali, Medellín, Barranquilla y Bucaramanga para clientes y no clientes de la marca. Para este fin se realizó una encuesta autoaplicada online.

Se evaluaron los diferentes beneficios que tiene una tarjeta de crédito como el tipo de *oferta*, la *modalidad* de la cuota de manejo, el *precio* de la cuota de manejo y la *tasa*, los cuales tienen las siguientes categorías:

Tabla 1: Atributos y categorías.

| Atributo | Categoría |
|---|---|
| Oferta | Devolución del 5 %. |
| | Devolución exclusiva. |
| | Devolución del IVA. |
| Modalidad | Cobro de la cuota de manejo solo si usa la tarjeta. |
| | Exoneración de la cuota de manejo si realiza transacciones por más de $150.000 al mes. |
| | Exoneración de la cuota si se realizan más de 5 transacciones al mes. |
| Precio | $7.000 de cuota de manejo. |
| | $15.000 de cuota de manejo. |
| | $19.000 de cuota de manejo. |
| Tasa | Tasa de interés preferencial. |
| | Acumulación de puntos. |
| | Sin cobro de intereses por compras diferidas a 1 mes. |

**Fuente:** Elaboración propia

A cada individuo se le solicitó organizar por orden de preferencia 9 perfiles propuestos para el proyecto:

- **Perfil 1:** Devolución del 5 %, cobro de la cuota de manejo solo si usa la tarjeta, $7.000 de cuota de manejo y una tasa de interés preferencial.
- **Perfil 2:** Descuentos exclusivos, exoneración de la cuota de manejo si realiza transacciones por más de $150.000 al mes, $7.000 de cuota de manejo y una tasa de interés preferencial.
- **Perfil 3:** Devolución del IVA, exoneración de la cuota si se realizan más de 5 transacciones al mes, $15.000 de cuota de manejo y una tasa de interés preferencial.
- **Perfil 4:** Devolución del 5 %, exoneración de la cuota si se realizan más de 5 transacciones al mes, $19.000 de cuota de manejo y una tasa de interés preferencial.
- **Perfil 5:** Devolución del IVA, exoneración de la cuota si se realizan más de 5 transacciones al mes, $7.000 de cuota de manejo y acumulación de puntos.
- **Perfil 6:** Descuentos exclusivos, exoneración de la cuota de manejo si realiza transacciones por más de $150.000 al mes, $15.000 de cuota de manejo y acumulación de puntos.
- **Perfil 7:** Descuentos exclusivos, exoneración de la cuota si se realizan más de 5 transacciones al mes, $19.000 de cuota de manejo y acumulación de puntos.
- **Perfil 8:** Devolución del 5 %„ exoneración de la cuota si se realizan más de 5 transacciones al mes, $15.000 de cuota de manejo y sin cobro de intereses por compras diferidas a 1 mes.
- **Perfil 9:** Devolución del IVA, exoneración de la cuota de manejo si realiza transacciones por más de $150.000 al mes, $19.000 de cuota de manejo y sin cobro de intereses por compras diferidas a 1 mes.

A la hora de recolectar los datos se obtuvo una base de la siguiente forma:

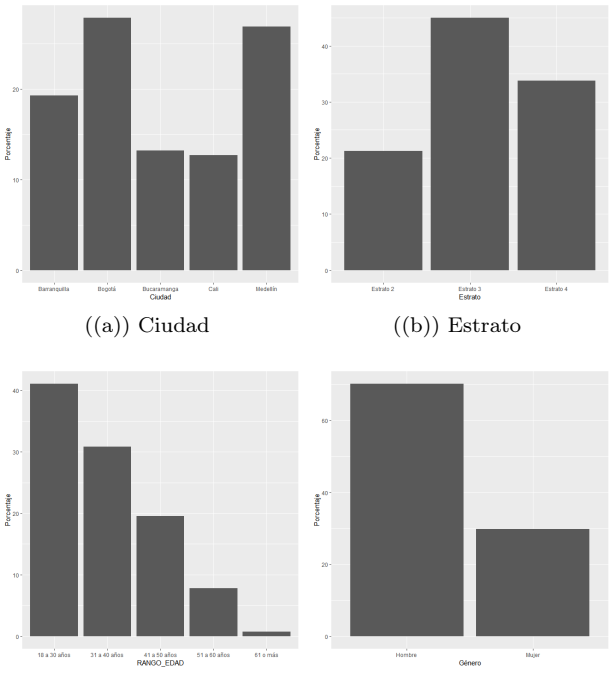
Tabla 2: Estructura de la base de datos

| Individuo | Perfil | Ranking |
|---|---|---|
| 1 | Perfil 1 | 1-9 |
| 1 | Perfil 2 | 1-9 |
| ... | ... | ... |
| 1 | Perfil 8 | 1-9 |
| 1 | Perfil 9 | 1-9 |
| ... | ... | ... |
| n | Perfil 1 | 1-9 |
| n | Perfil 2 | 1-9 |
| ... | ... | ... |
| n | Perfil 8 | 1-9 |
| n | Perfil 9 | 1-9 |

**Fuente:** Elaboración propia.

## 3.1. Análisis sociodemográfico

En la *Figura 5* se observa que la mayoría de los encuestados viven en Bogotá y Medellín, esto representa el 27.87 % y el 26.89 % de la muestra respectivamente. Además el 44.99 % de las personas vive en estrato 3. Por otro lado el 41.08 % de los encuestados tienen entre 18 y 30 años de edad y solo el 0.73 % tienen 61 años o más. Al ver la *figura d* se observa que la mayoría de los respondientes son hombres y representan el 70.17 % del total de encuestados, lo cual puede indicar que hubo un sesgo a la hora de realizar las encuestas. No obstante esta mayor diferencia puede deberse a condiciones de mercado en el uso de las tarjetas.




((a)) Ciudad ((b)) Estrato

((c)) Rango de edad ((d)) Género

Figura 5: Sociodemográficos

## 3.2. Transformación de los datos - Análisis de Homogeneidad

Con el fin de contrastar la hipótesis planteada en inicio de este documento acerca de la posibilidad de simplificar y de robustecer las estimaciones de un Conjoint con datos rankeados a través de la utilización de un proceso multivariado conocido como análisis de homogeneidad se procede a explicar el procedimiento y los recursos utilizados para tal fin.

Para llevar a cabo la transformación de los datos, se recurre al uso del paquete "homals" desarrollado para el software estadístico R por Patrick Mair y Jan De Leeuw quienes en la literatura académica son ampliamente reconocidos por sus aportes en la psicometría. Mair es profesor titular de estadística en el Departamento de Psicología de la Universidad de Harvard. Cuenta con un doctorado en estadística y una maestría en psicología de la Universidad de Viena. Sus aportes de investigación se centran en las estadísticas computacionales y aplicadas, con énfasis en la psicometría. Entre tanto De Leeuw es profesor emérito distinguido de Estadística y presidente fundador del Departamento de Estadística de la Universidad de California en Los Ángeles. También es autor de alrededor de 761 escritos biográficos en estadística, análisis de datos y psicometría.

El paquete homals es conocido por realizar análisis de homogeneidad (análisis de correspondencia múltiple) y varias extensiones del mismo. En este caso utilizaremos la función "homals" del paquete para llevar a cabo el análisis de homogeneidad el cual nos permite cuantificación del ranking dado por los respondientes. Con la ejecución de la función sobre los datos se obtiene una matriz que contiene la matriz de datos reproducidos basada en puntuaciones de categoría. Así por ejemplo el resultado obtenido es el siguiente:

Tabla 3: Resultado - Análisis de Homogeneidad

| **Respondiente** | **Perfil Evaluado** | **Ranking** | **Score** |
|---|---|---|---|
| 1 | Perfil 1 | 8 | 0.0013751967 |
| 1 | Perfil 2 | 1 | -0.0027033984 |
| 1 | Perfil 3 | 4 | 0.0001929864 |
| 1 | Perfil 4 | 5 | -0.0002299645 |
| 1 | Perfil 5 | 9 | 0.0003272461 |
| 1 | Perfil 6 | 3 | -0.0002494756 |
| 1 | Perfil 7 | 6 | 0.0011455514 |
| 1 | Perfil 8 | 7 | 0.0018128218 |
| 1 | Perfil 9 | 2 | -0.0016709641 |

**Fuente:** Elaboración propia.

La columna Score es el resultado del análisis de correspondencias llevado a cabo con la función homals. A partir de ella se realiza una nueva transformación para convertirla en un puntaje de 0 a 1 utilizando la siguiente fórmula:

$$\frac{R_i - min(R)}{max(R) - min(R)} \tag{16}$$

Donde R es la columna Ranking y $R_i$ es el elemento i-ésimo de la columna R.

## 3.3. Resultados del modelo jerárquico lineal

Como se había anotado previamente la diferencia entre un modelo lineal y el modelo lineal jerarquizado radica en que el segundo a diferencia del primero, no solo estima los efectos fijos sino que estima efectos individuales por niveles.

En la tabla, los coeficientes indican la asociación de las categorías con la preferencia (Score). Es importante aclarar que en este caso los coeficientes no pueden ser relacionados directamente a un cambio en la variable dependiente sino que expresan numéricamente la utilidad (part-worth) sobre cada factor. Al respecto del concepto de utilidad cabe aclarar que "Es la base conceptual para medir el valor. No tiene una equivalencia directa a alguna otra medida, ni siquiera dentro del mismo conjoint, dado que la utilidad depende de los atributos o niveles que se miden y por supuesto de la escala o nivel de medición empleado."(Andrade J, 2010).

Por ende, de la tabla se puede interpretar que la mejor tarjeta de crédito podría tener como caracteristica descuentos exclusivos todos los días del año, dado que los valores de Oferta_Devol_5 % y Oferta_Devol_iva son negativos. Adicionalmente las personas preferirían no tener cobros de intereses por compras diferidas a un mes. Para el caso de la Modalidad y el Precio, las personas prefieren una tarjeta que les exonere el cobro de cuota de manejo si realizan transacciones en el mes por más de $150.000. y resulta extraño que el modelo arroje la preferencia por el cobro de cuota de manejo por $19.000 cuando la lógica del consumidor pagar lo menos posible por cualquier bien, objeto o servicio.

Tabla 4: Resultados - Efectos fijo del modelo lineal jerárquico

| **Efectos Fijos** | **Estimación** | **Std. Error** | **Valor t** |
|---|---|---|---|
| (Intercepto) | 0.709609 | 0.034602 | 19.533 |
| Oferta_Devol_5 % | -0.023833 | 0.023814 | -1.001 |
| Oferta_Devol_iva | -0.010568 | 0.017332 | -0.610 |
| Modalidad_Exonerar_150 | -0.019541 | 0.023557 | -0.830 |
| Modalidad_Exonerar_5tr | 0.003087 | 0.017992 | 0.172 |
| Precio_19Mil | 0.033681 | 0.011242 | 2.996 |
| Precio_7Mil | -0.199212 | 0.015360 | -10.777 |
| Tasa_Puntos | -0.022597 | 0.021332 | -1.059 |
| Tasa_Int_Prefer | -0.021560 | 0.014508 | -1.486 |

**Fuente:** Elaboración propia.

## 3.4. Resultados del modelo jerárquico Bayesiano

La aproximación al modelo bayesiano para la estimación de modelos jerárquicos puede proporcionar estimaciones más certeras con la ventaja de trabajar con pocas observaciones e inclusive realizar estimaciones para cada individuo.

En esta sección, se aplica el método jerárquico Bayesiano para trabajar datos basados en puntuación que se obtuvieron tras la transformación mediante correspondencias múltiples. Para esta estimación se utilizó el paquete MCMCPack el cual contiene funciones para realizar inferencia bayesiana mediante simulaciones. Dentro de este paquete se encuentra la función MCMCregress la cual genera una muestra a partir de la distribución posterior de un modelo de regresión lineal con errores de Gauss usando el muestreador de Gibbs. Este modelo asume que cada entrevistado tiene un conjunto de preferencias determinadas por un distribución mayor que define el rango de posibles preferencias.

Al momento de observar las estimaciones realizadas, es posible determinar que estas son casi idénticas a las obtenidas en el modelo frecuentista antes presentado. La diferencia entre ambas estimaciones, la frecuentista y la bayesiana radica en que la segunda calcula los coeficientes para cada individuo. Es decir, que es posible para una agencia de investigación de mercado realizar una segmentación de las preferencias individuales de los entrevistados. Esto permite orientar los esfuerzos hacia públicos específicos con cierto nivel de referencias similares.

Tabla 5: Resultados - Modelo jerárquico Bayesiano

| | **Media** | **SD** |
|---|---|---|
| (Intercepto) | 0.709630 | 0.0010495 |
| OFERTADevol_5 | -0.023845 | 0.0006841 |
| OFERTADevol_Iva | -0.010577 | 0.0004908 |
| MODALIDADExonera_150 | -0.019570 | 0.0008065 |
| MODALIDADExonerar_5tr | 0.003074 | 0.0006322 |
| PRECIO15mil | -0.033681 | 0.0003499 |
| PRECIO7mil | -0.199216 | 0.0005097 |
| TASAPuntos | -0.022608 | 0.0006239 |
| TASATI_pref | -0.021557 | 0.0004279 |

**Fuente:** Elaboración propia.

## 3.5. Resultados Análisis Conjunto Tradicional

Tabla 6: Resultados - Modelo análisis conjunto tradicional

| **Fuente** | **Utilidad** |
|---|---|
| Intercepto | 5.0000 |
| Oferta-Desc-Exclusivos | 0.002 |
| Oferta-Devol-5 % | -0.071 |
| Oferta-Devol-Iva | 0.069 |
| Modalidad-Cuota-Uso | -0.026 |
| Modalidad-Exonerar-150 | 0.052 |
| Modalidad-Exonerar-5 | -0.027 |
| Precio-7mil | 1.036 |
| Precio-15mil | -0.219 |
| Precio-19mil | -0.817 |
| Tasa-No-int | 0.016 |
| Tasa-puntos | -0.036 |
| Tasa-T.int-pref | 0.020 |

**Fuente:** Elaboración propia.

.

Los resultados del análisis conjunto tradicional obtenidos usando el paquete estadistico XLSTAT mediante el método de Mínimos Cuadrados Ordinarios (OLS) para la respuesta tipo ranking dan cuenta de una mayor preferencia de las personas hacia un producto con una cuota de manejo de 7 mil pesos, con una tasa de interés preferencial, devolución del IVA por compras realizadas en alimentos y exoneración de cuota de manejo si realiza transacciones en el mes por más de $150.000.

Figura 6: Resultados por Modalidad

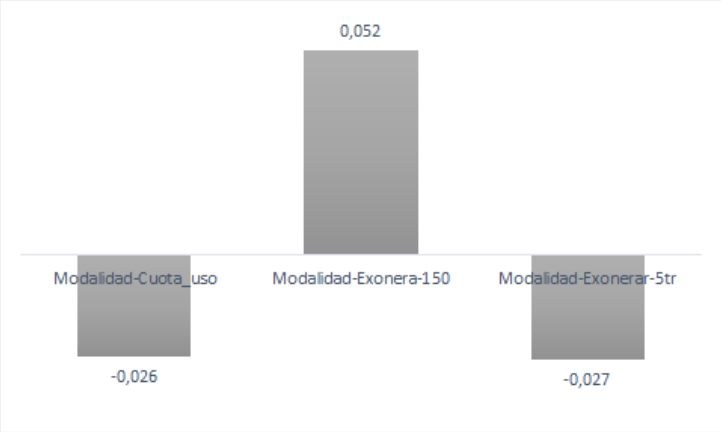



Figura 7: Resultados por Oferta

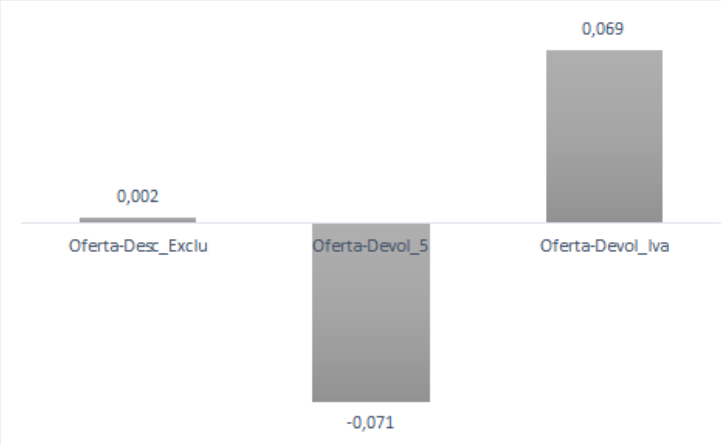



Figura 8: Resultados por Precio

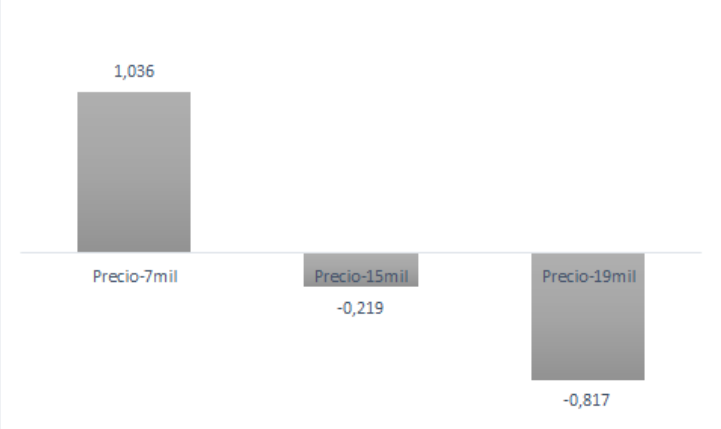
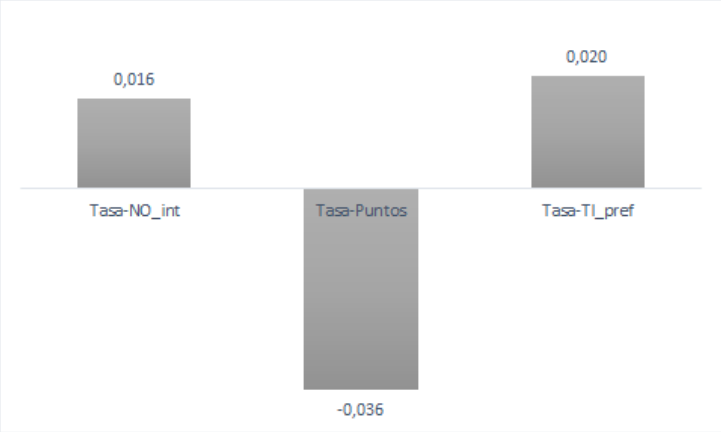


Figura 9: Resultados por Oferta




## 3.6. Conclusiones

Los modelos evaluados en este documento son similares pero no idénticos. Tratar de responder la pregunta acerca de cuál de los dos modelos es más pertinente dependerá en gran medida de la experticia del investigador o ejecutor del modelo. Será precisamente la experiencia la encargada de establecer la coherencia de los resultados con la realidad.

Como se ha notado en este documento, las estimaciones del modelo lineal jerárquico frecuentista y el modelo jerárquico bayesiano son aproximadamente idénticos. Por lo tanto desde este único punto de vista no sería posible determinar diferencias significativas entre los resultados, sin embargo es importante destacar que el modelo bayesiano permite calcular las utilidades para cada uno de los individuos de la muestra lo que genera un grado mayor de información para la toma de decisiones frente al lanzamiento de nuevos productos adaptados a las necesidades de los clientes.

Como menciona Park (2004), los datos de tipo ranking contienen menos información y violan los supuestos de normalidad del modelo Jerárquico Bayesiano. No obstante, aplicar un modelo Jerárquico Bayesiano a la variable dependiente de tipo ranking puede llegar a ser tan robusta como un modelo de Mínimos Cuadrados Ordinarios (OLS)(p.1096). El hallazgo de este documento, es que no existe una ganancia de información al transformar los datos vía Análisis de Componentes Principales, de hecho existe una diferencia interpretativa entre los resultados de ambas metodologías.

## 3.7. Referencias Bilbliográficas

Andrade, J., 2010. *Conjoint analysis a pie*. Market Variance. Recuperado de http://marketvariance.com/blog/2010/01/12/conjoint-analysis-a-pie.

Bak, A. , Bartlomowicz, T. 2012. *Conjoint analysis method and its implementation in conjoint R package*. Wroclaw University of Economics.

Brewer, B. J., (s.f). *Introduction to Bayesian Statistics*. Recuperado de https://www.stat.auckland.ac.nz/brewer/stats331.pdf

Chapman, C. 2015. *R for marketing research and analytics*. Springer.

De la Cerna Villavicencio, J. F. 2016. *Identificación de preferencias académicas universitarias en alumnos de los últimos años de educación secundaria en el colegio particular "Bella Unión" mediante el uso del Análisis Conjunto -Perfil Completo- con el aplicativo estadístico R*. Lima, Perú.

Economic Social Research Council. sf. *An introduction to MCMC methods and Bayesian Statistics*. Recuperado de https://www.ukdataservice.ac.uk/media/307220/presentation4.pdf.

Edlira ,S. 2005. *A Hierarchical Bayes Model for Ranked Conjoint Data*. University of Hamburg.

Factum Mercadotécnico. 2009. *Análisis de conjuntos (Conjoint Analysis*. Recuperado de www.factum-marketing.com/download.php?file=CONJOINT.ppt.

Ferreira Lopes, S. (2011). Análisis Conjunto. *Teoría, campos de aplicación y conceptos inherentes*. Estudios y Perspectivas en Turismo, *20* (2), 341-366.

Lam, K.Y. Koning, A.J. Franses, Ph.H.B.F., 2010. *Ranking Models in Conjoint Analysis,.Econometric Institute Research Papers EI 2010-51, Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute..*

Lenk, P. J. , DeSarbo, W. S. , Green, P. E. and Young, M. R. 1996. *Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs*. Informs.

Luce, R. D. , Tukey, R. D. 1994. *Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement*. , Journal of Mathematical Psychology. 1, 1-27.

Maydeu-Olivares, A, Bockenholt, U. 2005. *Structural Equation Modeling of Paired-Comparison and Ranking Data*.Psychological Methods. Vol. 10, No. 3, pp. 285-304.

McCullough, J, Best, R. 1979. *Conjoint Measurement: Temporal Stability and Structural Reliability*.Journal of Marketing Research. Vol. 16, No. 1 (Feb., 1979), pp. 26-31.

Okamoto, Y. 2015. *A Two-Step Analysis with Common Quantification of Categorical Data*. Japan Women's University Journal: Faculty of Integrated Arts and Social Sciences, 2015, 26, 99-112.

Orme, B. 2010. *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*. Second Edition, Madison, Wis.: Research Publishers LLC.

Park, C. 2004. *The robustness of hierarchical Bayes conjoint analysis under alternative measurement scales*. Journal of Business Research *57*, 1092-1097.

Ramírez, J. M. 2010. Measuring Preferences: from Conjoint Analysis to Integrated Conjoint Experiments. *Revista de métodos cuantitativos para la economía y la empresa*. 28-43.

Ravenzwaaij, D. , Cassey, P. , Brown, S. (2018). A simple introduction to Markov Chain Monte-Carlo sampling. *Psychonomic Bulletin  Review. 25*, 143-154.

Sullivan, L., Dukes, K.A.,  Losina, E. (1999). *Tutorial in biostatistics. An introduction to hierarchical linear modelling.*Statistics in medicine, 18 7, 855-88.

Sharma, S. 2017. *Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy.*Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW.

Teichert, T. 2005. *A Hierarchical Bayes Model for Ranked Conjoint Data.* University of Hamburg.

Tenenhaus, M. , Young, F. W. 19855. *An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data.* Psychometrika 50(1):91-119.

Woltman, H. , Feldstain, A. , MacKay, J. C. , Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology . 8*(1), 52-69.

Yildirim, I. 2012. *Bayesian Inference: Metropolis-Hastings Sampling.* Department of Brain and Cognitive Sciences, University of Rochester. Rochester, NY.

Yildirim, I. 2012. *Bayesian Inference: Gibbs Sampling.* Department of Brain and Cognitive Sciences, University of Rochester. Rochester, NY.

Van Der Burg, De Leeuw, Verdegaal, 1988. *Homogeneity analysis with k sets of variables: an alternating least squares method with optimal scaling features.Tutorials in Quantitative Methods for Psychology . 8*(1), 52-69. Psychometrika, University of Rochester. Rochester, NY.