

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

ESTIMACIÓN BAYESIANA DE ECUACIONES ESTRUCTURALES

**PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO
MATEMÁTICO**

CHRISTIAN ALEJANDRO HIDALGO CORREA
cristian118@hotmail.com

DIRECTOR: LUIS JAIME ANDRADE GONZÁLEZ MSc.
jaime.andrade@epn.edu.ec

Quito, julio 2012

DECLARACIÓN

Yo, **CHRISTIAN ALEJANDRO HIDALGO CORREA**, declaro que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluye en este documento.

La Escuela Politécnica Nacional, puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

Christian Alejandro Hidalgo Correa

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por **CHRISTIAN ALEJANDRO HIDALGO CORREA**, bajo mi supervisión.

Jaime Andrade G. MSc.

DIRECTOR DE PROYECTO

AGRADECIMIENTOS

Muchas gracias a todos: a mis papás, a mis hermanos, a mis amigos y profesores, sobre todo a mi novia Susi, todo este trabajo es fruto de cada uno de sus consejos y motivaciones, de cada palabra y gesto hacia mi, sin ustedes no habría sido posible todo esto.

Gracias Susi por esos momentos maravillosos, por tu amor, tu cariño y comprensión.

MUCHÍSIMAS GRACIAS

DEDICATORIA

Se lo dedico a un sueño roto, a ese que ya no se podrá cumplir, y a dado paso a otros sueños, las circunstancias me llevaron por otro camino el cual culmina aquí.

Una vida sin sueños no es vida, y esta vida es para intentar cumplir nuestros sueños.

CONTENIDO

CONTENIDO	1
ÍNDICE DE FIGURAS	4
ÍNDICE DE CUADROS	7
1. ESTIMACIÓN BAYESIANA	11
1.1. INTRODUCCIÓN	11
1.2. TEOREMA DE BAYES	12
1.3. ESTIMACIÓN BAYESIANA	16
1.4. INFERENCIA BAYESIANA	20
1.4.1. EJEMPLO: INFERENCIA BAYESIANA DE LA PROBABILI- DAD DE ÉXITO DE UNA VARIABLE BERNOULLI	23
2. MÉTODOS DE MONTE CARLO VÍA CADENAS DE MARKOV	28
2.1. INTRODUCCIÓN	28
2.2. MÉTODOS DE MONTE CARLO VÍA CADENAS DE MARKOV	29
2.2.1. RESEÑA HISTÓRICA	29
2.2.2. CONCEPTOS BÁSICOS Y ALGORITMO BASE	30
2.3. MÉTODOS DE MONTECARLO VÍA CADENAS DE MAR- KOV: ALGORITMOS	37
2.3.1. ALGORITMO HASTING METRÓPOLIS	37

	2
2.3.2. EJEMPLO: MUESTRA DE LA DISTRIBUCIÓN BETA UTILIZANDO EL ALGORITMO DE HASTING METRÓPOLIS	38
2.3.3. ALGORITMO DE HASTING METRÓPOLIS COMPONENTE A COMPONENTE	40
2.3.4. EJEMPLO: MUESTRA DE LA DISTRIBUCIÓN NORMAL BIVARIADA UTILIZANDO EL ALGORITMO DE HASTING METRÓPOLIS COMPONENTE A COMPONENTE	41
2.3.5. MUESTREADOR DE GIBBS	44
2.3.6. EJEMPLO: ESTIMACIÓN DE MOMENTOS DE LA FUNCIÓN DE DISTRIBUCIÓN BINOMIAL UTILIZANDO EL MUESTREADOR DE GIBBS	45
3. ECUACIONES ESTRUCTURALES	48
3.1. INTRODUCCIÓN	48
3.2. ECUACIONES ESTRUCTURALES: CONCEPTOS BÁSICOS	49
3.2.1. ANÁLISIS FACTORIAL EXPLORATORIO	53
3.2.2. ANÁLISIS FACTORIAL CONFIRMATORIO	54
3.2.3. ANÁLISIS FACTORIAL DE ORDEN SUPERIOR	56
3.3. ESTIMACIÓN BAYESIANA DE ECUACIONES ESTRUCTURALES	59
3.3.1. ANÁLISIS FACTORIAL CONFIRMATORIO USANDO ESTIMACIÓN BAYESIANA	60
4. APLICACIÓN: CONSTRUCCIÓN DE UN ÍNDICE DE SATISFACCIÓN DEL CLIENTE CONSIDERANDO EL MODELO ACSI	63
4.1. INTRODUCCIÓN	63
4.2. FILOSOFÍA DE LA EMPRESA: EL ENFOQUE MARKETING	64
4.2.1. EL ENFOQUE MARKETING	65

	3
4.3. SATISFACCIÓN DEL CLIENTE: DEFINICIÓN Y BENEFICIOS DE LOGRARLA	66
4.3.1. ¿QUÉ ES LA SATISFACCIÓN DEL CLIENTE?	68
4.3.2. BENEFICIOS DE LOGRAR ALTOS NIVELES DE SATISFACCIÓN EN EL CLIENTE	69
4.4. MODELO ACSI: <i>AMERICAN CUSTOMER SATISFACTION INDEX</i>	70
4.4.1. MODELOS DE ECUACIONES ESTRUCTURALES PARA LA ESTIMACIÓN DE LA SATISFACCIÓN DEL CLIENTE	71
5. CONCLUSIONES Y RECOMENDACIONES	117
5.1. CONCLUSIONES	117
5.2. RECOMENDACIONES	119
BIBLIOGRAFÍA	121
ANEXOS	126

Índice de figuras

1.1. Distribución a priori de θ	25
1.2. Distribución a posteriori de θ	26
2.1. Histograma obtenido después de realizar 10000 iteraciones	39
2.2. Distribución Beta con parámetros (1.5;3.25)	39
2.3. Componentes generados de la distribución normal bivariada	42
2.4. Distribución normal bivariada teórica	42
2.5. Gráfico para la media de X según el número de iteraciones	46
2.6. Gráfico para la media de Y según número de iteraciones	46
2.7. Histogramas de las variables X e Y	47
3.1. Representación gráfica de un modelo de ecuaciones estructurales	52
3.2. Representación gráfica del análisis factorial exploratorio	54
3.3. Representación gráfica del análisis factorial confirmatorio	56
3.4. Representación gráfica del modelo estructural para la elaboración de un análisis factorial confirmatorio de orden superior	58
4.1. Representación gráfica del modelo <i>ACSI</i>	72
4.2. Representación gráfica del modelo planteado utilizando la metodología <i>ACSI</i>	73
4.3. Representación gráfica del modelo utilizado para la realización del análisis factorial exploratorio.	76
4.4. Representación gráfica de resultado obtenido en la aplicación del AFE	78

4.5. Representación gráfica del modelo resultante después de realizar un cambio de nombres en las variables	79
4.6. Ejemplo de distribuciones a priori utilizadas en la estimación	81
4.7. Autocorrelogramas de las estimaciones bayesianas del Análisis Factorial Exploratorio considerando 10000 iteraciones	83
4.8. Autocorrelogramas de las estimaciones bayesianas del Análisis Factorial Exploratorio considerando 10000 iteraciones	84
4.9. Convergencia de las estimaciones en el Análisis Factorial Exploratorio . . .	85
4.10. Comparación de series de iteraciones de estimaciones bayesianas en el análisis factorial exploratorio	86
4.11. Comparación de series de iteraciones de estimaciones bayesianas en el análisis factorial exploratorio	87
4.12. Comparación de series de iteraciones de estimaciones bayesianas en el análisis factorial exploratorio (continuación)	88
4.13. Modelo resultante después de haber realizado el análisis factorial exploratorio utilizando estimación bayesiana	92
4.14. Modelo considerado para el análisis factorial confirmatorio	96
4.15. Autocorrelogramas de las estimaciones bayesianas del Análisis Factorial Confirmatorio considerando 10000 iteraciones	99
4.16. Autocorrelogramas de las estimaciones bayesianas del Análisis Factorial Confirmatorio considerando 10000 iteraciones	100
4.17. Convergencia de las estimaciones en el Análisis Factorial Confirmatorio . .	102
4.18. Comparación de series de iteraciones estimaciones bayesianas en el análisis factorial confirmatorio	103
4.19. Comparación de series de iteraciones estimaciones bayesianas en el análisis factorial confirmatorio (continuación)	104
4.20. Comparación de series de iteraciones estimaciones bayesianas en el análisis factorial confirmatorio (continuación)	105
4.21. Usos de la satisfacción del cliente estimada	110

4.22. Ejemplo de lectura del modelo de ecuaciones estructurales	111
4.23. Modelo resultante	112
4.24. Mecanismo de generación de satisfacción del cliente	116
5.1. Comparación de histogramas de estimaciones bayesianas en el análisis factorial exploratorio	146
5.2. Comparación de histogramas de estimaciones bayesianas en el análisis factorial exploratorio (continuación)	147
5.3. Comparación de histogramas de estimaciones bayesianas en el análisis factorial exploratorio (continuación)	148
5.4. Comparación de histogramas de estimaciones bayesianas en el Análisis Factorial Confirmatorio	149
5.5. Comparación de histogramas de estimaciones bayesianas en el Análisis Factorial Confirmatorio (continuación)	150
5.6. Comparación de histogramas de estimaciones bayesianas en el Análisis Factorial Confirmatorio (continuación)	151

Índice de cuadros

1.1. Datos: aplicación del Teorema de Bayes	15
1.2. Comparación de distribuciones a priori y a posteriori	26
2.1. Comparación de distribuciones teórica y simulada	39
2.2. Medidas descriptivas de las variables simuladas	46
4.1. Estadísticos Descriptivos	77
4.2. Estimaciones de máxima verosimilitud para el Análisis Factorial Exploratorio	80
4.3. Cálculo del período de calentamiento	84
4.4. Estimaciones bayesiana y estadísticos de convergencia para el Análisis Fac- torial Exploratorio	85
4.5. Resultados (AFE)	89
4.6. Matriz de varianzas-covarianzas	94
4.7. Matriz de correlaciones	94
4.8. Ajuste global del modelo para combinaciones de covarianzas representadas	97
4.9. Estimaciones de máxima verosimilitud para el Análisis Factorial Confirmatorio	98
4.10. Cálculo del período de calentamiento para el Análisis Factorial Confirmatorio	101
4.11. Estimaciones bayesiana y estadísticos de convergencia para el Análisis Fac- torial Confirmatorio	102
4.12. Resultados (AFC)	106
4.13. Asociación de conceptos	111
5.1. Cuestionario: Satisfacción del cliente	144

5.2. Cuestionario: Satisfacción del cliente (Parte 2)	145
5.3. Cambio de nomenclatura	145

RESUMEN

El objetivo de este trabajo es la estimación de variables inobservables, para ello se utiliza la técnica de modelización de ecuaciones estructurales y se combina la estimación de máxima verosimilitud con la estimación bayesiana, y con esto lograr mejores estimaciones, inicialmente se considera el Teorema de Bayes y sus fundamentos como una introducción, la teoría desarrollada está fundamentada en el trabajo realizado por Thomas Bayes el cuál fue publicado dos años después de su muerte, el capítulo inicial aborda la teoría de estimación e inferencia bayesiana, además, se ha fundamentado su uso con ejemplos prácticos, se concluye el capítulo con la aplicación del Teorema de Bayes.

Los métodos de Monte Carlo surgen en las últimas décadas, debido al advenimiento de grandes avances en la computación, esto ha permitido que se puedan desarrollar simulaciones que difícilmente se hubieran podido realizar “a mano”, estos métodos se estudian en el capítulo 2, haciendo una breve reseña histórica del origen y sus creadores, además, se desarrollan los conceptos básicos y el algoritmo a partir del cual mediante modificaciones surgen los algoritmos de: Hasting-Metrópolis, Hasting-Metrópolis componente a componente, y el muestreador de Gibbs. Dichos algoritmos son los más comúnmente utilizados, se ha incluido ejemplos de su funcionamiento, estos algoritmos son la base para la creación de nuevos algoritmos adaptados a cualquier tipo de problema.

La modelización de ecuaciones estructurales es una metodología que permite estimar variables inobservables, esta técnica es la única que permite este tipo de estudio, actualmente se encuentra en auge debido a las miles de aplicaciones posibles en diversos campos de la

ciencia, ésta técnica dio sus primeras luces con los trabajos de Spearman y Thrustone en el campo del análisis psicométrico y es tratada en el capítulo 3 presentando los siguientes modelos: Análisis Factorial Exploratorio, Análisis Factorial Confirmatorio y su generalización. Al final del capítulo la modelización de ecuaciones estructurales utiliza la teoría bayesiana para la estimación de sus parámetros.

En el capítulo cuarto se utiliza la teoría tanto bayesiana, así como, las ecuaciones estructurales para la construcción de un índice de satisfacción del cliente para una empresa ecuatoriana, pero, ¿qué es la satisfacción?, a ésta y otras preguntas que surgen de esta aplicación en el campo de la investigación de mercados, se responden mediante los conocimientos incluidos. En este capítulo se encuentra una rápida revisión de lo que se conoce como visión de la empresa, dada por, Enfoque Marketing, después, se aborda a la satisfacción del cliente como uno de los aspectos importantes en cuanto a manejo de clientes se refiere, ¿qué es la satisfacción?, ¿cuáles son los beneficios de lograr satisfacer al cliente?, forman parte del desarrollo del modelo y ayudarán a una mejor comprensión de la metodología *ACSI* para construir la variable objetivo (satisfacción del cliente). Para terminar el presente trabajo, se realizan las conclusiones y recomendaciones obtenidas a lo largo del desarrollo de cada uno de los temas planteados.

Capítulo 1

ESTIMACIÓN BAYESIANA

1.1. INTRODUCCIÓN

La estadística bayesiana, tiene un enfoque muy singular, debido a que se encuentra basada en la llamada probabilidad subjetiva, la cual puede ser alterada por el investigador en función de los datos observados o experiencias en fenómenos similares al estudiado. El objetivo principal de la estimación bayesiana consiste en un método para evaluación de valores estimados de parámetros desconocidos de un modelo estocástico utilizando muestras y, su principal fundamento, la estadística bayesiana. En este capítulo se da una introducción a la estimación bayesiana, inicialmente se enuncia el Teorema de Bayes, como inducción a este punto de vista, a continuación, se desarrolla la teoría de inferencia bayesiana como el principal método para identificar las características de los parámetros considerados, según la visión bayesiana, como variables aleatorias.

A lo largo del capítulo se desarrollan ejemplos tanto teóricos como prácticos: de la utilización del Teorema de Bayes, del uso las distribuciones a priori, el criterio de selección de parámetros para expresar incertidumbre o certeza y la obtención de la distribución a posteriori tanto para el caso de distribuciones conjugadas y no conjugadas. El uso de distribuciones conjugadas es visto en los ejemplos, principalmente, en el modelo de regresión para datos normales.

1.2. TEOREMA DE BAYES

Uno de los objetivos de la estadística es la construcción de modelos estadísticos, lo más simple posibles debido al principio de parsimonia, que sean una aproximación de la realidad, bajo supuestos e hipótesis propias del fenómeno de estudio. En estadística inferencial sobresalen dos escuelas: la Frecuentista y la Bayesiana.

Escuela Frecuentista: los métodos frecuentistas son los más comúnmente utilizados, dichos métodos asumen que cualquier experimento puede ser repetido infinito número de veces y es necesario considerar todas las posibles observaciones del experimento con el fin de realizar la inferencia estadística. La incertidumbre en los resultados del experimento es usada para describir la incertidumbre en los parámetros del modelo estadístico (Knight, 2000).

Escuela Bayesiana: este punto de vista no es mejor ni peor que el frecuentista, se fundamenta en el conocimiento de una distribución sobre un parámetro θ , $f(\theta)$, denominada distribución a priori. Tal distribución se supone completamente conocida, no tiene parámetros desconocidos e involucra probabilidades subjetivas. Además, existe el método bayesiano empírico, el cual depende también de la existencia de una distribución a priori $f(\theta)$ a la cual se le puede dar una interpretación frecuencial y se puede estimar usando observaciones apropiadas, es por esto que el método bayesiano empírico puede ser esencialmente no bayesiano, en el sentido de no involucrar probabilidades subjetivas, y sería un error confundir con el estrictamente bayesiano (Prieto, Quintero, Rodríguez, 2008).

Definición 1. *Modelo Estadístico (Knight, 2000)*

Sean Y_1, \dots, Y_n variables aleatorias, se tienen las siguientes observaciones y_1, \dots, y_n de las variables aleatorias Y_1, \dots, Y_n . Se supone que la distribución conjunta de $Y = (Y_1, \dots, Y_n)$ es desconocida pero pertenece a una familia de distribuciones. Dicha familia de distribuciones es conocida como modelo estadístico y se puede escribir de la siguiente forma:

$$Y = (Y_1, \dots, Y_n) \rightsquigarrow F_\theta \quad \text{para } \theta \in \Theta$$

Donde F_θ es la distribución de probabilidad conjunta de Y ; Θ es el conjunto de posibles valores del parámetro θ . Se asume que Θ es un subconjunto de un espacio euclidiano, por tanto, θ puede ser un valor real o un vector (Knight K, 2000).

Definición 2. Probabilidad Condicional (Ross, 2007)

Sea el espacio de probabilidad (Ω, \mathcal{A}, P) y un evento $H \in \mathcal{A}$ tal que $P(H) > 0$, se define

$$P(A|H) = \frac{P(A \cap H)}{P(H)} \quad (1.1)$$

TEOREMA 1. Regla del producto (Ross, 2007)

Sea el espacio de probabilidad (Ω, \mathcal{A}, P) y un conjunto de sucesos $A_1, \dots, A_n \in \mathcal{A}$ se tiene

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdots P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right) \quad (1.2)$$

TEOREMA 2. Teorema de la Probabilidad Total (Ross, 2007)

Sea el espacio de probabilidad (Ω, \mathcal{A}, P) y dada la partición $A_1, \dots, A_n \in \mathcal{A}$ de Ω , donde n puede ser tan grande como se desee, se tiene

$$P(B) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i) \quad (1.3)$$

para cualquier evento $B \in \mathcal{A}$

Thomas Bayes (Londres, Inglaterra, 1702 - Tunbridge Wells, 1761) fue un matemático británico. Posiblemente De Moivre, autor del libro “La doctrina de las probabilidades”, fue su maestro particular, pues se sabe que por ese entonces ejercía como profesor en Londres. Thomas Bayes muere en el año 1761, dos años después se publicó su obra “*Essay Towards Solving a Problem in the Doctrine of Chances*”, en aquella memoria aparece por primera vez la determinación de la probabilidad de las causas a partir de los efectos que han podido ser observados. El cálculo de dichas probabilidades recibe el nombre de Teorema de Bayes, y resuelve el problema conocido como “de la probabilidad inversa”. Thomas Bayes fue uno de los pioneros en el uso de la probabilidad de forma inductiva y en establecer una base matemática para la inferencia probabilística. Actualmente, con base a su obra, se ha desarrollado una poderosa teoría que ha conseguido notables aplicaciones en las más diversas áreas del conocimiento (Wikipedia, 2012).

TEOREMA 3. *Teorema de Bayes (Ross, 2007)*

Sea el espacio de probabilidad (Ω, \mathcal{A}, P) , dada la partición $A_1, \dots, A_n \in \mathcal{A}$ de Ω , donde n puede ser tan grande como se desee, se tiene

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{i=1}^n P(A_i) \cdot P(B|A_i)} \quad (1.4)$$

para cualquier suceso $B \in \mathcal{A}$

Demostración. De la definición 2 se tiene

$$\begin{aligned} P(A_i \cap B) &= P(B) \cdot P(A_i|B) \\ P(A_i \cap B) &= P(A_i) \cdot P(B|A_i) \end{aligned}$$

Igualando las dos ecuaciones anteriores y despejando

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{P(B)}$$

Aplicando el teorema 2, se concluye la demostración

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{i=1}^n P(A_i) \cdot P(B|A_i)}$$

□

La ecuación (1.4) se puede expresar de una forma más simple para dos posibles eventos A y B

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)} \propto P(A) \cdot P(B|A) \quad (1.5)$$

A la ecuación 1.5 también se la conoce como la *Regla de Bayes* (Ross, 2007). En el caso de los modelos estadísticos, se puede notar que, el teorema 3 ofrece un mecanismo probabilístico de aprendizaje, es decir, después de haber observado la muestra (y_1, \dots, y_n) se puede calcular la distribución a posteriori $F(\theta|y_1, y_2, \dots, y_n)$, la cual combina toda la información disponible (datos recolectados y distribución a priori de θ) (Hair y otros, 2005).

Se debe notar que la distribución a posteriori es el elemento clave en la inferencia bayesiana.

Ejemplo 1. *Aplicación del Teorema de Bayes*

Se desea calcular la probabilidad de que un fumador desarrolle cáncer de pulmón, se conoce los siguientes datos¹:

Cuadro 1.1: Datos: aplicación del Teorema de Bayes

categorias	muerte con cáncer de pulmón	muerte sin cáncer de pulmón	Total
fumador	51	23	74
no fumador	32	47	79
Total	83	70	153

Eventos:

A_1 : tener cáncer de pulmón

A_2 : morir sin cáncer de pulmón

B : ser fumador

Probabilidad de los eventos:

$$P(B|A_1) = \frac{51}{83} = 0,6150$$

$$P(B|A_2) = \frac{23}{70} = 0,3290$$

$$P(A_1) = \frac{83}{153} = 0,5425$$

$$P(A_2) = \frac{70}{153} = 0,4575$$

$$\begin{aligned} P(A_1|B) &= \frac{P(A_1) \cdot P(B|A_1)}{P(B)} \\ &= \frac{P(A_1) \cdot P(B|A_1)}{P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2)} \\ &= 0,6891 \end{aligned}$$

Es decir, que la probabilidad de tener cáncer de pulmón dado que es un fumador es 0,6891 (aproximadamente 70 de cada 100 fumadores tendrán cáncer de pulmón).

¹Los datos considerados no pertenecen a ningún estudio y menos son datos reales, solamente son usados para expresar el funcionamiento de la metodología de uso del Teorema de Bayes

1.3. ESTIMACIÓN BAYESIANA

La estimación bayesiana permite incorporar información conocida de los parámetros a estimar, por ejemplo, si el objetivo es estimar los coeficientes de un modelo lineal y se conoce la función con la cual se encuentran distribuidos sus valores, se puede incorporar esta información para lograr mejores estimaciones. La característica más importante de la estimación bayesiana es describir la incertidumbre en θ , un parámetro a estimar, con una distribución de probabilidad en el espacio de parámetros Θ . (Knight, 2000)

Por esto, el objetivo de la estadística bayesiana es el uso de información conocida previamente para hacer inferencia, es decir, conocer las propiedades o características de parámetros desconocidos con datos o conocimientos previos. Los métodos frecuentistas consideran no aleatorios a los parámetros desconocidos; en la estimación bayesiana los parámetros de interés se consideran como variables aleatorias con distribuciones a priori que pueden reflejar conocimientos de los valores que podrían tomar dichos parámetros (Athreya y otros, 2003).

La estimación bayesiana fundamenta su proceder sobre el principio de que la información o conocimiento previo sobre algunos rasgos del parámetro son elementos contribuyentes en su estimación (Ntzoufras, 2009). El enfoque bayesiano considera una muestra aleatoria de una población con función de densidad $f_X(x, \theta)$, en la cual el parámetro θ es una variable aleatoria con función de densidad $g(\theta)$, a esta función se la denomina función de densidad a priori. (Mayorga J. 2004)

La función de densidad a posteriori se define de la siguiente manera: sea $Y = (y_1, \dots, y_n)$ variables aleatorias con función de densidad conjunta $h(y, \theta)$ donde θ es una variable aleatoria, además, como se vió en el párrafo anterior, $g(\theta)$ es la función de densidad a priori de θ y describe la incertidumbre en el valor del parámetro θ antes de observar

Y . Se define la función de densidad a posteriori de θ dado $Y = (y_1, \dots, y_n)$ de la siguiente forma:

$$h(\theta|y) = \frac{h(y, \theta) g(\theta)}{\int_{\Theta} h(y, t) g(t) dt} \quad (1.6)$$

Además, Yohai y Boente (2006) mencionan que, una de las ventajas del enfoque bayesiano es que se pueden definir en forma natural estimadores óptimos, sin necesidad de utilizar restricciones como la que se usan en los estimadores insesgados en el enfoque frecuentista. Lehmann y Casella (1998) indican que al momento de buscar un estimador se busca uno que minimice el riesgo promedio ($R(\theta, \delta)$) para una función de pérdida no negativa, los siguientes teoremas son tomados de su trabajo.

TEOREMA 4. *Estimador de Bayes*

Sea θ una variable aleatoria con ley de distribución Λ , sea X una variable aleatoria con función de distribución P_θ . Suponiendo que se desea estimar $g(\theta)$ con una función de pérdida $L(\theta, d)$ no negativa, entonces:

- Existe un estimador δ_0 con riesgo finito
- Para casi todo x , existe un valor δ_Λ que minimiza

$$E\{L[\theta, \delta(x)]|X = x\}$$

Entonces, $\delta_\Lambda(X)$ es el estimador de Bayes

Demostración. Sea δ cualquier estimador de riesgo finito. Entonces $E\{L[\theta, \delta(x)]|X = x\}$ es finita, debido a que L es una función de pérdida no negativa. Por lo tanto:

$$E\{L[\theta, \delta(x)]|X = x\} \geq E\{L[\theta, \delta_\Lambda(x)]|X = x\}$$

$$L[\theta, \delta(x)] \geq L[\theta, \delta_\Lambda(x)]$$

□

TEOREMA 5. *Estimador de Bayes insesgado*

Sea θ una variable aleatoria con ley de distribución Λ , y sea X una variable aleatoria y P_θ la función de distribución condicional de X dado θ . Considerar la estimación de $g(\theta)$ con la función de pérdida cuadrática. Entonces, el estimador de Bayes $\delta(X)$ es insesgado si:

$$E[\delta(X) - g(\theta)]^2 = 0$$

donde la esperanza es tomada respecto a X y θ .

Demostración. Suponiendo que $\delta(X)$ es un estimador de Bayes insesgado para estimar $g(\theta)$. Como $\delta(X)$ es un estimador de Bayes con la función de pérdida cuadrática, entonces

$$\delta(X) = E[g(\theta)|X]$$

con probabilidad igual a 1. Además como $\delta(X)$ es insesgado, entonces

$$E[\delta(X)|\theta] = g(\theta) \quad \forall \theta$$

Condicionando en X , se tiene:

$$E[g(\theta)\delta(X)] = E\{\delta(X)E[g(\theta)|X]\} = E[\delta^2(X)]$$

Condicionando en θ , se tiene:

$$E[g(\theta)\delta(X)] = E\{g(\theta)E[\delta(X)|\theta]\} = E[g^2(\theta)]$$

De lo que se tiene:

$$E[\delta(X) - g(\theta)]^2 = E[\delta^2(X)] + E[g^2(\theta)] - 2E[\delta(X)g(\theta)] = 0$$

□

TEOREMA 6. *Admisibilidad del estimador de Bayes*

Cualquier estimador de Bayes único es admisible.

Demostración. Si δ es un único estimador de Bayes con respecto a la distribución a priori Λ y está dominada por $\tilde{\delta}$, entonces

$$\int R(\theta, \tilde{\delta}) d\lambda(\theta) \leq \int R(\theta, \delta) d\lambda(\theta)$$

lo cual contradice la unicidad □

TEOREMA 7. *Eficiencia asintótica del estimador de Bayes*

Sea X_1, \dots, X_n variables independientes idénticamente distribuidas con función de densidad $f(X_i|\theta)$, donde θ es un valor real del espacio muestral Ω . Si $\pi^*(t|x)$ es la función de densidad a posteriori de $\sqrt{n}(\theta - T_n)$ donde

$$T_n = \theta_0 + \frac{1}{nI(\theta_0)} l'(\theta_0)$$

entonces si:

- Sea $l(\theta)$ la función de máxima verosimilitud, si $\tilde{\theta} = \tilde{\theta}_n$ es cualquier sucesión tal que $\tilde{\theta}_n \xrightarrow{p} \theta$ entonces:

$$l(\theta) = l(\theta_0) + (\theta - \theta_0)l'(\theta_0) - \frac{1}{2}(\theta - \theta_0)^2[nI(\theta_0) + R_n(\theta)] \quad (1.7)$$

donde

$$\frac{1}{n}R_n(\theta) \xrightarrow{p} 0 \quad \text{cuando } n \rightarrow \infty$$

- $\forall \varepsilon > 0$, existe un $\delta > 0$ tal que en (1.7), la probabilidad de un evento

$$\sup \left\{ \left| \frac{1}{n}R_n(\theta) \right| : |\theta - \theta_0| \leq \delta \right\} \geq \varepsilon$$

tiende a cero cuando $n \rightarrow \infty$

- $\forall \delta > 0$, existe $\varepsilon > 0$ tal que la probabilidad del evento

$$\sup \left\{ \frac{1}{n}[l(\theta) - l(\theta_0)] : |\theta - \theta_0| \geq \delta \right\} \leq -\varepsilon$$

tiende a 1 cuando $n \rightarrow \infty$

- La densidad a priori π de θ es continua y positiva para todo θ

se tiene:

$$\int \left| \pi^*(t|x) - \sqrt{I(\theta_0)} \phi \left[t \sqrt{I(\theta_0)} \right] \right| dt \stackrel{p}{\rightarrow} 0$$

Si además la esperanza de θ bajo π existe y viene dado por:

$$\int |\theta| \pi(\theta) d(\theta) < \infty$$

entonces

$$\int (1 + |t|) \left| \pi^*(t|x) - \sqrt{I(\theta_0)} \phi \left[t \sqrt{I(\theta_0)} \right] \right| dt \stackrel{p}{\rightarrow} 0$$

De los teoremas presentados anteriormente se puede extraer que el estimador bayesiano tiene las siguientes propiedades:

- El estimador de Bayes es insesgado, si se cumple $E[\delta(X) - g(\theta)]^2 = 0$
- Si el estimador bayesiano es único, entonces, es admisible
- Los estimadores de Bayes son asintóticamente eficientes

1.4. INFERENCIA BAYESIANA

El proceso de inferencia bayesiana que a continuación se describe, permite concluir acerca de las propiedades de un parámetro considerando datos observados y la distribución a priori. Mediante el cálculo de la distribución a posteriori $f(\theta|y)$, se combina la información previa del parámetro y la muestra observada, esta característica hace a la inferencia bayesiana diferente de la inferencia frecuentista.

La especificación de la distribución a priori es de vital importancia en la inferencia bayesiana, debido su influencia en la forma que tendrá la distribución a posteriori, esta característica motiva la resistencia a la aceptación de este punto de vista por parte de los frecuentistas. Hay que enfatizar en la especificación de la media y la varianza a priori, mediante las cuales se expresan los conocimientos previos:

- La media a priori como una medida de tendencia da información del parámetro de interés.
- Y la varianza a priori expresa la incertidumbre asociada al parámetro, en el caso de distribuciones no informativas, este valor es “grande”.

Se puede dar el caso donde no se tiene información de la distribución a priori, debido a esto es necesario especificar una distribución que no influya en la distribución a posteriori y dejar que los datos “hablen por sí mismos”. Tales distribuciones también se las conoce como *no informativas* o *distribuciones a priori vagas*, estas distribuciones se pueden usar sin ningún problema y la función a posteriori obtenida será válida para realizar la inferencia.

Un problema que se presenta al momento de aplicar la ecuación (1.6) es la intratabilidad analítica de la distribución a posteriori resultante, este problema complica la inferencia, y por ende, la identificación de las características del parámetro, para evitar que se presente este tipo de problema se utilizan distribuciones a priori conjugadas; estas distribuciones tienen la ventaja de influir en la distribución a posteriori resultante logrando que ésta pertenezca a la misma familia de distribuciones.

Definición 3. *Distribución conjugada (Congdon, 2003)*

Sea X_1, \dots, X_n una muestra aleatoria de una población con función de densidad $f_X(x, \theta)$. Una familia D de densidades se dice que es **conjugada** para la función de densidad $f_X(x, \theta)$, o que es **cerrada bajo muestreo** respecto a la función de densidad $f_X(x, \theta)$, si la función de densidad a priori de Θ , $g_\Theta(\theta) \in D$ y si $f_{\Theta|X_1, \dots, X_n}(\theta|x_1, \dots, x_n) \in D$.

Al momento de obtener la distribución a posteriori, sus parámetros están expresados en función de los parámetros de la distribución a priori; se debe poner atención a las distribuciones que pertenecen a la familia exponencial (distribución normal, exponencial, gamma, chi cuadrado, beta, etc).

Ntzoufras (2009), menciona que el proceso de inferencia bayesiana se lo puede dividir en cuatro etapas:

1. **Construcción del modelo:** para la construcción seguimos el siguiente proceso:
 - a) Identificación de la principal variable respuesta Y con sus respectivos datos (muestra observada)
 - b) Búsqueda de una distribución adecuada que describa Y
 - c) Identificación de variables que puedan estar influyendo o puedan ayudar a explicar la variable Y
 - d) Especificar una distribución a priori (puede ser una distribución no informativa la cual incorpore total incertidumbre o una distribución que incorpore información previa, experiencia de fenómenos similares al estudiado, o una distribución conveniente para los cálculos)
2. **Cálculo de la distribución a posteriori:** se debe encontrar el método adecuado de cálculo de la distribución a posteriori (métodos analíticos, métodos asintóticos, técnicas de simulación) y entonces implementar el método seleccionado para lograr la estimación de la distribución a posteriori.
3. **Análisis de la distribución a posteriori:** se procede al cálculo de los momentos, intervalos de confianza, cuantiles y concluir de las propiedades de la distribución a posteriori obtenida.
4. **Inferencia:** se enfoca el análisis en las medidas calculadas anteriormente: medidas descriptivas, intervalos.
5. **Conclusiones:** se concluirá de acuerdo al problema que se esté manejando, y, a los resultados obtenidos.

Para una mejor comprensión de la inferencia bayesiana se expone el siguiente ejemplo ficticio.

1.4.1. EJEMPLO: INFERENCIA BAYESIANA DE LA PROBABILIDAD DE ÉXITO DE UNA VARIABLE BERNOULLI

En cierta investigación social se desea estimar la probabilidad de que un niño ubicado en una determinada condición social se encuentre trabajando en la informalidad. Se procede de la siguiente forma:

1. Construcción del modelo:

- Variable respuesta X :
 el niño ubicado **si está** trabajando actualmente $X_i = 1$
 el niño ubicado **no está** trabajando actualmente $X_i = 0$
- Distribución adecuada de X : para la variable de respuesta existen dos posibles resultados: si está trabajando o no lo está, es decir, que X es una variable aleatoria Bernoulli.

La función de probabilidad de X es:

$$P(X = x) = f(x) = \theta^x (1 - \theta)^{1-x} \quad \text{con } x \in \{0, 1\} \quad (1.8)$$

- Identificación de variables que influyen: Se logra observar que la variable influyente en la función de probabilidad (1.8) es θ , por tanto, debemos buscar información acerca de este parámetro.
- Distribución a priori: sea X_1, X_2, \dots, X_n una muestra (i.i.d.) de una variable aleatoria con distribución Bernoulli con parámetro desconocido θ y se asume que la distribución a priori de este parámetro es:

$$f(\theta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1} \quad \text{con } 0 < \theta < 1$$

donde α y β son parámetros especificados.

Usualmente, se vuelve difícil la obtención de información certera acerca del parámetro θ por cuanto no se conoce o no se tiene acceso a estudios de referencia, debido es esta dificultad, la elección más conveniente es utilizar la distribución

beta con parámetros menores o iguales a $\alpha = \beta = 10^{-3}$, la cual es una distribución no informativa.

La media de la distribución $f(\theta)$ es igual a:

$$\begin{aligned} E[\theta] &= \frac{\alpha}{\alpha + \beta} \\ &= \frac{10^{-3}}{10^{-3} + 10^{-3}} \\ &= \frac{1}{2} \end{aligned}$$

Mientras que la varianza de θ es:

$$\begin{aligned} V[\theta] &= \frac{\alpha \cdot \beta}{(\alpha + \beta + 1) \cdot (\alpha + \beta)^2} \\ &= \frac{10^{-3} \cdot 10^{-3}}{(10^{-3} + 10^{-3} + 1) \cdot (10^{-3} + 10^{-3})^2} \\ &= 0,2495 \end{aligned}$$

2. Cálculo de la distribución a posteriori: para ello se parte del cálculo analítico de la distribución a posteriori cuando la distribución a priori es una beta con parámetros α y β .

Asumiendo $y = \sum_{i=1}^n x_i$

$$\begin{aligned} f(\theta|x_1, \dots, x_n) &= \frac{f(x_1, \dots, x_n; \theta) f(\theta)}{\int_{\Theta} f(x_1, \dots, x_n; t) f(t) dt} \\ &= \frac{\theta^y (1 - \theta)^{n-y} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}}{\int_0^1 \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha+y-1} (1 - t)^{n+\beta-y-1} dt} \\ &= \frac{\theta^{\alpha+y-1} (1 - \theta)^{n+\beta-y-1}}{\int_0^1 t^{\alpha+y-1} (1 - t)^{n+\beta-y-1} dt} \\ &= \frac{1}{\frac{\Gamma(\alpha + y) \Gamma(n + \beta - y)}{\Gamma(\alpha + \beta + n)}} \theta^{\alpha+y-1} (1 - \theta)^{n+\beta-y-1} \\ &= \frac{\Gamma(\alpha + \beta + n)}{\Gamma(\alpha + y) \Gamma(n + \beta - y)} \theta^{\alpha+y-1} (1 - \theta)^{n+\beta-y-1} \end{aligned}$$

Se puede notar, que la distribución a posteriori es también una función beta, es decir, es una distribución conjugada, además sus parámetros están en función de los parámetros de la distribución a priori. Por tanto:

$$\begin{aligned}
 E[\theta|x_1, \dots, x_n] &= \frac{\alpha + y}{(\alpha + y) + (n + \beta - y)} \\
 &= \frac{\alpha + y}{\alpha + \beta + n} \\
 V[\theta|x_1, \dots, x_n] &= \frac{(\alpha + y)(n + \beta - y)}{(\alpha + y + n + \beta - y + 1)(\alpha + y + n + \beta - y)^2} \\
 &= \frac{(\alpha + y)(n + \beta - y)}{(\alpha + \beta + n + 1)(\alpha + \beta + n)^2}
 \end{aligned}$$

los datos del presente ejemplo son: $\alpha = \beta = 10^{-3}$ y $n = 26$ niños.

Reemplazando los datos tenemos:

$$\begin{aligned}
 f(\theta|x_1, \dots, x_n) &= \frac{\Gamma(0,001 + 0,001 + 26)}{\Gamma(0,001 + 17)\Gamma(26 + 0,001 - 17)} \theta^{0,001+17-1} (1 - \theta)^{26+0,001-17-1} \\
 &= \frac{\Gamma(26,002)}{\Gamma(9,001)\Gamma(31,001)} \theta^{16,001} (1 - \theta)^{8,001}
 \end{aligned}$$

3. Análisis de la distribución a posteriori: para el análisis de la distribución a posteriori se recurrirá a las gráficas de las dos distribuciones: a priori y a posteriori.

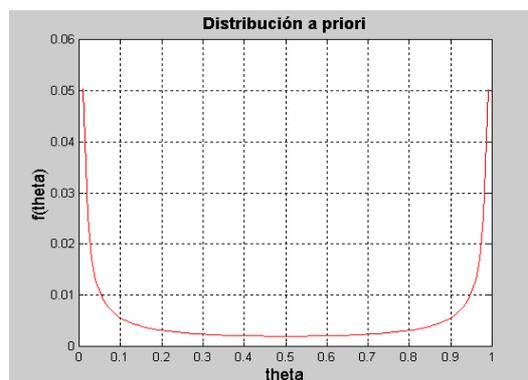


Figura 1.1: Distribución a priori de θ

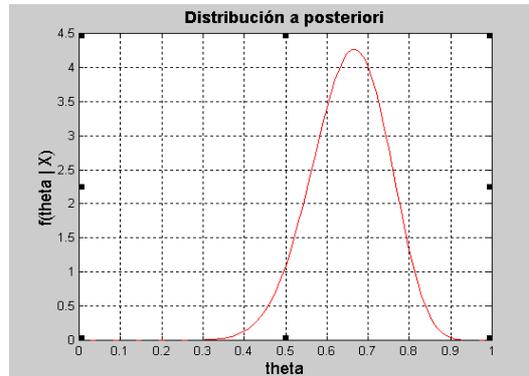


Figura 1.2: Distribución a posteriori de θ

De las figuras 1.1 y 1.2 se puede ver un cambio drástico en la forma de las distribuciones del parámetro θ evidenciando como en la distribución a posteriori se han incluido los datos observados de la variable X , para mayor facilidad de interpretación en la siguiente tabla se encuentran las medidas descriptivas de cada una de las distribuciones.

Cuadro 1.2: Comparación de distribuciones a priori y a posteriori

Parámetros	Distribución a priori	Distribución a posteriori
α	0,001	17,001
β	0,001	9,001
esperanza	0,500	0,654
varianza	0,250	0,008
desviación estándar	0,500	0,092
1 ^{er} cuartil	0	0,593
Mediana (2 ^{do} cuartil)	0,500	0,658
3 ^{er} cuartil	1,000	0,719

4. Inferencia: Se puede ver que la varianza de la distribución a posteriori es menor en comparación a la varianza a priori, indicando así que existe mayor certeza en la estimación del parámetro θ y el valor esperado para este parámetro es 0,654.

Una comparación de los cuartiles entre las dos distribuciones evidencia la traslación que la probabilidad tuvo, dicha probabilidad está concentrándose alrededor de la media a posteriori.

5. Conclusiones: Al observar una muestra de 26 niños se tiene que la probabilidad de que un niño ubicado en una determinada condición social trabaje es 0,654 %, es decir, 6 de cada 10 niños en determinada condición social trabajan.

Capítulo 2

MÉTODOS DE MONTE CARLO VÍA CADENAS DE MARKOV

2.1. INTRODUCCIÓN

En este capítulo se aborda el uso de los métodos para la generación de número aleatorios, y luego, serán incorporados en la inferencia Bayesiana. El capítulo esta enfocado en los Métodos de Monte Carlo vía Cadenas de Markov debido a que éstos son usados ampliamente en la estimación bayesiana. Los métodos de Monte Carlo son no determinísticos; generalmente son usados para aproximar expresiones matemáticas complejas y costosas de evaluar con exactitud.

Los métodos tomaron su nombre en referencia al casino de Monte Carlo por ser “la capital del juego de azar”, al ser la ruleta un generador simple de números aleatorios. El nombre y el desarrollo sistemático de los métodos de Monte Carlo data aproximadamente de 1944 y se mejoraron enormemente con el desarrollo tecnológico de la computadora principalmente en las capacidades de memoria, esto permite realizar gran cantidad de cálculos numéricos. Los métodos de Monte Carlo proporcionan soluciones aproximadas a una gran variedad de problemas matemáticos posibilitando la realización de experimentos con generación de números aleatorios en una computadora.

2.2. MÉTODOS DE MONTE CARLO VÍA CADENAS DE MARKOV

2.2.1. RESEÑA HISTÓRICA

La invención del método de Monte Carlo se asigna a Stanislaw Ulam y a John Von Neumann. Ulam ha explicado como se le ocurrió la idea mientras jugaba una partida de solitario durante una enfermedad que sufrió en el año 1946. Advirtió que resulta mucho más simple tener una idea del resultado general del solitario haciendo pruebas múltiples con las cartas y contando las proporciones de los resultados, que calcular todas las posibilidades de combinación de cartas formalmente. Se le ocurrió que esta misma observación debía aplicarse a su trabajo en el desarrollo de la bomba atómica durante la segunda guerra mundial en el Laboratorio Nacional de Los Álamos en Estados Unidos (Wikipedia, 2012).

La idea consistía en probar con experimentos mentales las posibilidades, y en cada etapa, determinar un número aleatorio distribuido según las probabilidades y totalizar todas las posibilidades para tener una idea de la conducta del proceso físico (Wikipedia, 2012).

Para esa época ya se podían utilizar máquinas de computación, debido a que, comenzaban a estar disponibles y con ellas se podía efectuar las pruebas numéricas y en efecto reemplazar el aparato experimental del físico. Durante una de las visitas de Von Neumann a Los Álamos en 1946, Ulam le mencionó el método. Después de cierto escepticismo inicial, Von Neumann se entusiasmó con la idea y pronto comenzó a desarrollar las posibilidades del método en un procedimiento sistemático. Ulam expresó que *“Monte Carlo comenzó a tener forma concreta y empezó a desarrollarse con todas sus fallas de teoría rudimentaria después de que se lo propuse a Johnny”* (Wikipedia, 2012).

A principios de 1947 Von Neumann envió una carta a Richtmyer a Los Álamos en la que expuso de modo influyente el método de Monte Carlo; este informe puede que sea primero en donde se detalló el método de Monte Carlo. Su carta fue encuadernada junto

con la respuesta de Richtmyer como un informe de Los Álamos y distribuida entre los miembros del laboratorio donde se sugería utilizar el método. Ulam estaba particularmente interesado en el método Monte Carlo para evaluar integrales múltiples. Una de las primeras aplicaciones de este método a un problema determinista fue llevada a cabo en 1948 por Enrico Fermi, Ulam y Von Neumann cuando consideraron los valores singulares de la ecuación de Schrödinger (Wikipedia, 2012).

Hoy en día los Métodos de Monte Carlo son aplicables a diversos problemas, ya sean estocásticos o deterministas. El método de Monte Carlo tiene un error absoluto de la estimación que decrece como $\frac{1}{\sqrt{N}}$ en virtud del teorema del límite central .

2.2.2. CONCEPTOS BÁSICOS Y ALGORITMO BASE

Antes de iniciar la descripción del algoritmo, es necesario definir los siguientes conceptos básicos:

Definición 4. *Proceso estocástico (Ross, 2007)*

Un proceso estocástico $\{X(t), t \in T\}$ es un conjunto de variables aleatorias, tal que $\forall t \in T$, $X(t)$ es una variable aleatoria. El índice t es interpretado como el tiempo, por tanto, $X(t)$ es el estado del proceso estocástico al tiempo t . Si el conjunto T es contable, entonces, el proceso estocástico es de tiempo discreto, si el conjunto T es infinito no numerable, el proceso estocástico es de tiempo continuo (Ross, 2007).

Un ejemplo de proceso estocástico es la cantidad de compradores que se encuentran satisfechos con un producto o servicio al tiempo t .

El espacio de estados de un proceso estocástico es el conjunto de todos los posibles valores que la variable aleatoria $X(t)$ puede tomar. Por lo tanto, el proceso estocástico es una familia de variables aleatorias que describen la evolución a través del tiempo de algún fenómeno.

Definición 5. *Cadena de Markov (Ross, 2007)*

Sea un proceso estocástico $\{X_n, n = 0, 1, 2, \dots\}$ que toma un número finito o contable de posibles valores. Este conjunto de posibles valores se lo denota por un conjunto de enteros no negativos $\{0, 1, 2, \dots\}$. Si $X_n = i$, entonces el proceso está en el estado i en el tiempo n . Si el proceso se encuentra en un estado i , existe una probabilidad fija P_{ij} de que el proceso tome el estado j .

$$P\{X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\} = P\{X_{n+1} = j | X_n = i_n\} = P_{ij}$$

para todos los estados $i_0, i_1, \dots, i_{n-1}, i_n, j$ y para todo $n \geq 0$. El valor P_{ij} representa la probabilidad de que la cadena de Markov estando en el estado i pase al estado j , las probabilidades son no negativas y la cadena de Markov siempre debe encontrarse en algún estado, debido a esto se tiene:

$$\begin{aligned} P_{ij} &\geq 0, & i, j \geq 0 \\ \sum_{j=0}^{\infty} P_{ij} &= 1, & i = 0, 1, \dots \end{aligned}$$

Definición 6. *Probabilidad de transición en n pasos (Ross, 2007)*

En la definición anterior se encuentra la probabilidad de transición a un solo paso, P_{ij} , ahora se define la probabilidad de transición en n pasos como la probabilidad de que una cadena de Markov ubicada en el estado i , después de n pasos llegue al estado j , como:

$$P_{ij}^n = P\{X_{n+k} = j | X_k = i\}, \quad n \geq 0, i, j \geq 0$$

Las ecuaciones de Chapman Kolmogorov proveen un método para el cálculo de las probabilidades de transición en n pasos, estas ecuaciones son:

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m, \quad \forall m, n \geq 0, \forall i, j$$

P_{ij}^{n+m} representa la probabilidad de que la cadena de Markov iniciando en el estado i , el proceso vaya hasta el estado j en $n + m$ transiciones, y, en el camino la cadena de Markov

pasa por el estado k después de n transiciones. Formalmente se tiene:

$$\begin{aligned}
 P_{ij}^{n+m} &= P \{X_{n+m} = j | X_0 = i\} \\
 &= \sum_{k=0}^{\infty} P \{X_{n+m} = j, X_n = k | X_0 = i\} \\
 &= \sum_{k=0}^{\infty} P \{X_{n+m} = j | X_n = k, X_0 = i\} P \{X_n = k | X_0 = i\} \\
 &= \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m
 \end{aligned}$$

Definición 7. Estado accesible (Ross, 2007)

Sean los estados i, j de una cadena de Markov se dice que el estado j es accesible desde el estado i si $P_{ij}^n > 0$ para algún $n \geq 0$.

Definición 8. Clasificación de los estados de cadena de Markov (Ross, 2007)

Sea el estado i , f_i la probabilidad de, comenzando en el estado i , el proceso estocástico vuelva a entrar en el estado i , Es decir, el estado i es recurrente si $f_i = 1$. Si el estado i es recurrente, entonces, la cadena de Markov volverá a entrar infinito número de veces en el estado i . Si $f_i < 1$ el estado es transitorio. Los estados de la cadena de Markov están clasificados de la siguiente forma:

$$\begin{aligned}
 \text{recurrente} & \text{ si } \sum_{n=1}^{\infty} P_{ii}^n = \infty \\
 \text{transitorio} & \text{ si } \sum_{n=1}^{\infty} P_{ii}^n < \infty
 \end{aligned}$$

Definición 9. Cadena de Markov irreducible (Ross, 2007)

Se dice que una cadena de Markov es irreducible si y solamente si todos los estados de la cadena se encuentran comunicados (o son accesibles desde otro estado)

Definición 10. Cadena de Markov aperiódica (Ross, 2007)

El periodo de un estado i se define como:

$$d(i) = \text{mcd}^1 \{n \geq 0 : P_{i,i}^n > 0\}$$

¹mcd: Máximo Común Divisor

Si $d(i) = 1$ entonces i es un estado aperiódico.

Una cadena de Markov se dice aperiódica si todos sus estados son aperiódicos.

Definición 11. Cadena de Markov Recurrente Positiva (Ross, 2007)

El estado de una cadena de Markov se dice recurrente si:

$$\sum_{n=1}^{\infty} P_{ii}^n = \infty$$

Es decir que un estado recurrente es visitado infinito número de veces. Si el estado i es recurrente, entonces se dice que es recurrente positivo si, comenzando en i el tiempo esperado que el proceso retorne al estado i es finito. Si el espacio de estados S es finito, entonces existe al menos un estado recurrente.

TEOREMA 8. Probabilidad invariante (Wu, 2012)

Si P es una matriz estocástica para una cadena de Markov irreducible y aperiódica entonces:

a. existe una única probabilidad invariante tal que $\vec{\pi}P = \vec{\pi}$

b. $\vec{\pi}$ puede ser calculada con

$$\lim_{n \rightarrow \infty} P_{ij}^n = \pi_j$$

y si \vec{v} es cualquier vector de probabilidad inicial

$$\lim_{n \rightarrow \infty} \vec{v} P_{ij}^n = \vec{\pi}$$

La demostración del teorema se expone en el Anexo 2.

Definición 12. Cadena de Markov reversible en el tiempo (Ross, 2007)

Una Cadena de Markov se dice que es reversible en el tiempo si:

$$\pi_i P_{ij} = \pi_j P_{ji} \quad \forall i \neq j$$

Se puede mostrar que bajo esta condición, si un estado se elige de acuerdo con las probabilidades $\{\pi_j\}$ en algún instante de tiempo la sucesión de estados regresará en el tiempo con probabilidades de transición P_{ij}

Ejemplo 2. Se expone el siguiente ejemplo como una aplicación del uso de las definiciones anteriores, el ejemplo es tomado de Ross (2007).

En cualquier día Gary puede estar alegre (A), ni alegre ni triste (NN) o triste (T). Si Gary está alegre, entonces al día siguiente él puede estar A , NN , T con las siguientes probabilidades, 0,5, 0,4, 0,1, respectivamente. Si él se siente ni alegre ni triste, al siguiente día él puede estar A , NN , T con las siguientes probabilidades, 0,3, 0,4, 0,3, respectivamente. Si él está triste, entonces puede estar A , NN , T al siguiente día con las siguientes probabilidades, 0,2, 0,3, 0,5.

Sea X_n , la variable aleatoria que representa el estado de ánimo de Gary en el n –ésimo día, entonces $\{X_n, n \geq 0\}$ es una cadena de Markov con tres estados ($A(0)$, $NN(1)$, $T(2)$) con la matriz de transición

$$\begin{pmatrix} 0,5 & 0,4 & 0,1 \\ 0,3 & 0,4 & 0,3 \\ 0,2 & 0,3 & 0,5 \end{pmatrix}$$

Es necesario conocer ¿qué proporción del tiempo el proceso está en cada uno de los tres estados?.

Para poder aplicar el teorema 4, primero se debe verificar que la cadena de markov es irreducible

- Irreducible: en la matriz de transición se puede ver que todos los estados de la cadena de Markov se encuentran comunicados ($P_{ij} > 0$).
- Aperiodica: se debe extraer el período de cada uno de los estados de la cadena de Markov, para ello es necesario calcular las matrices de transición a diferentes pasos:

$$P^2 = \begin{pmatrix} 0,356 & 0,378 & 0,266 \\ 0,336 & 0,370 & 0,294 \\ 0,322 & 0,364 & 0,314 \end{pmatrix}, P^3 = \begin{pmatrix} 0,345 & 0,373 & 0,282 \\ 0,338 & 0,371 & 0,292 \\ 0,333 & 0,369 & 0,298 \end{pmatrix},$$

$$P^5 = \begin{pmatrix} 0,341 & 0,372 & 0,287 \\ 0,338 & 0,371 & 0,291 \\ 0,337 & 0,370 & 0,293 \end{pmatrix}, P^{10} = \begin{pmatrix} 0,339 & 0,371 & 0,290 \\ 0,339 & 0,371 & 0,290 \\ 0,339 & 0,371 & 0,290 \end{pmatrix},$$

Se puede notar que a medida que aumenta el número de pasos, la matriz de transición tiene sus filas iguales, indicando que basta un solo paso para que la matriz regrese al estado del cual salió, el período de cada estado es 1, por lo tanto la cadena de Markov es aperiódica.

- *Recurrente:* tomando la definición 11, se ve que la matriz es recurrente debido a $\sum_{i=0}^{\infty} P_{ii} = \infty$, esto se cumple para cada estado de la cadena de Markov, por tanto, es positiva.

Las probabilidades límite π_i , $i = 0, 1, 2$ se extrae de la solución del sistema de ecuaciones obtenido de la aplicación del teorema 4, esto debido a que se ha visto que la cadena de Markov es recurrente positiva, aperiódica e irreducible.

$$\pi_0 = 0,5\pi_0 + 0,3\pi_1 + 0,2\pi_3$$

$$\pi_1 = 0,4\pi_0 + 0,4\pi_1 + 0,3\pi_3$$

$$\pi_2 = 0,1\pi_0 + 0,3\pi_1 + 0,5\pi_3$$

$$\pi_0 + \pi_1 + \pi_3 = 1$$

Las probabilidades límite obtenidas son: $\pi_0 = 0,339$, $\pi_1 = 0,371$, $\pi_2 = 0,290$, se puede notar que dichas probabilidades corresponden a los elementos de una de las filas de matriz de transición a 20 pasos. A medida que avanzan los pasos, en este caso, los días, los estados de ánimo que Gary se encuentran distribuidos con la ley

$$P(X_n) = \begin{cases} 0,339 & \text{si } X_n = A \\ 0,371 & \text{si } X_n = NN \\ 0,290 & \text{si } X_n = T \end{cases} \quad \text{cuando; } n \rightarrow \infty \quad (2.1)$$

Ntzoufras (2009), menciona que, el objetivo del uso de los métodos de Monte Carlo es generar una muestra de $f(\theta|y)$, para ello, se debe construir una cadena de Markov con las siguientes propiedades:

- $P(\theta^{(t+1)}|\theta^{(t)})$ debe tener una forma “fácil de generar”
- La distribución de equilibrio de la cadena de Markov debe ser la distribución a posteriori de $f(\theta|y)$

Ntzoufras (2009) expone que: asumiendo que se construye una cadena de Markov con las propiedades anteriores, el algoritmo de los Métodos de Monte Carlo via Cadenas de Markov viene dado de la siguiente forma:

1. Se selecciona un valor inicial $\theta^{(0)}$.
2. Se generan T valores hasta que la distribución de equilibrio sea alcanzada.
3. Medir la convergencia del algoritmo². Si la convergencia falla, entonces se deben generar más observaciones.
4. Eliminar las B primeras observaciones (periodo de inicialización)³.
5. Considerar $\{\theta^{B+1}, \theta^{B+2}, \dots, \theta^T\}$ como una muestra para realizar el análisis posterior.
6. Graficar la distribución a posteriori.
7. Con la distribución a posteriori obtener medidas descriptivas(media, mediana, varianza, desviación estándar, cuantiles, correlaciones).

El resultado de la aplicación del algoritmo anterior es una muestra aleatoria (Ntzoufras, 2009)

²La convergencia de un algoritmo MCMC se refiere a si la Cadena de Markov alcanza a la distribución de equilibrio y puede generar observaciones de la misma

³Son las iteraciones iniciales del algoritmo que van a ser eliminadas para evitar que dichos valores influyan en los resultados.

$$\theta^{(B+1)}, \theta^{(B+2)}, \dots, \theta^{(T)} \quad (2.2)$$

y para cualquier función $\Phi(\theta)$ se puede obtener:

1. Una muestra del parámetro de interés $\Phi(\theta)$ considerando la muestra 2.2

$$\Phi(\theta^{(B+1)}), \Phi(\theta^{(B+2)}), \dots, \Phi(\theta^{(T)})$$

2. Medidas descriptivas de $\Phi(\theta^{(B+1)})$. Como por ejemplo la media muestral a posteriori:

$$\hat{E}(\Phi(\theta)|y) = \frac{1}{T-B} \sum_{t=1}^{T-B} \Phi(\theta^{(t)})$$

2.3. MÉTODOS DE MONTECARLO VÍA CADENAS DE MARKOV: ALGORITMOS

En general, existe una gran complicación al momento de generar vectores aleatorios, cuyas variables aleatorias son dependientes (Ntzoufras, 2009). En esta sección se abordarán los principales algoritmos utilizados para la generación de vectores aleatorios con una función de masa de probabilidad dada como distribución límite.

2.3.1. ALGORITMO HASTING METRÓPOLIS

Ntzoufras (2009), indica que, para el desarrollo del siguiente algoritmo se asumirá la distribución objetivo $f(y)$ de donde se buscará generar una muestra de tamaño T , además, la distribución a posteriori $f(\theta|y)$. El algoritmo de Hasting Metrópolis construye una Cadena de Markov reversible en el tiempo con las deseadas probabilidades límite. El algoritmo tiene los siguientes pasos:

1. Fijar el valor inicial $\theta^{(0)}$
2. Para $t = 1, \dots, T$ realizar los siguientes pasos:

- a) Fijar $\theta = \theta^{(0)}$
- b) Generar θ' desde la distribución propuesta $q(\theta'|\theta)$
- c) Calcular:

$$\alpha = \min \left(1, \frac{f(\theta'|y)q(\theta|\theta')}{f(\theta|y)q(\theta'|\theta)} \right)$$

- d) Actualizar $\theta^{(t)} = \theta'$ con probabilidad α , de lo contrario $\theta^{(t)} = \theta$ (con probabilidad $1 - \alpha$)

2.3.2. EJEMPLO: MUESTRA DE LA DISTRIBUCIÓN BETA UTILIZANDO EL ALGORITMO DE HASTING METRÓPOLIS

Una vez conocido el algoritmo de Hasting Metrópolis, a continuación se desarrolla un ejemplo de aplicación.

Ejemplo 3. *Se desea generar una muestra números aleatorios con distribución beta ($a; b$). Sea la cadena de Markov cuya probabilidad de transición viene dada por la distribución de probabilidad de una variable uniforme ($U_{[p;q]}$). Con esto se tiene:*

$$\begin{aligned} \alpha(x_{n-1}, y) &= \min \left[1 \ ; \ \frac{\frac{1}{q-p} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1}}{\frac{1}{q-p} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x_{n-1}^{a-1}(1-x_{n-1})^{b-1}} \right] \\ &= \min \left[1 \ ; \ \frac{y^{a-1}(1-y)^{b-1}}{x_{n-1}^{a-1}(1-x_{n-1})^{b-1}} \right] \end{aligned}$$

Ahora fijamos los parámetro: $a = 1,5$, $b = 3,25$, utilizando el algoritmo de Hasting Metrópolis⁴ y se obtiene:

⁴En Anexo 2 se encuentra la sintaxis de Matlab utilizada para realizar la simulación

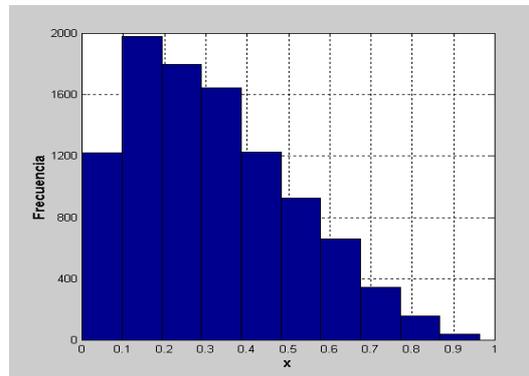


Figura 2.1: Histograma obtenido después de realizar 10000 iteraciones

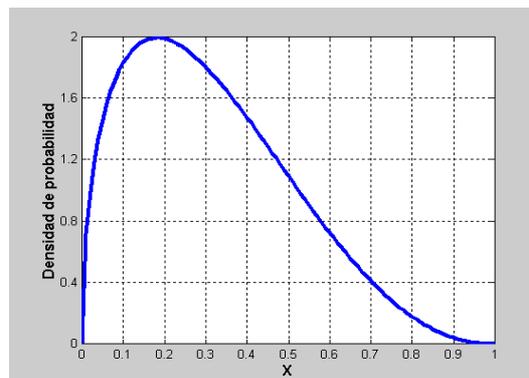


Figura 2.2: Distribución Beta con parámetros (1.5;3.25)

Se puede notar que el histograma obtenido de la simulación se ajusta a la forma de la distribución de probabilidad beta propuesta

Ahora se compara la media y la varianza de la distribución beta (1.5;3.25) con las obtenidas de la simulación.

Cuadro 2.1: Comparación de distribuciones teórica y simulada

	Distribución Beta	Simulación
Media	0,3158	0,3154
Varianza	0,0376	0,0379

Se puede observar que la media y varianza de la distribución beta y la simulación son semejantes, para tener una medida de cuan cercanas son las distribuciones se realizó la siguiente prueba de hipótesis (considerando la prueba χ^2 de bondad de ajuste):

H_0 : La distribución beta y la distribución simulada son iguales

H_1 : La distribución beta y la distribución simulada no son iguales

Se acepta H_0 al nivel de significancia de 5%, debido a $\chi^2 = 1,83326 < \chi_{0,05;7}^2 = 14,069$. Por lo tanto el algoritmo de Hasting Metrópolis es una herramienta válida al momento de generar valores de la distribución beta (1.5;3.25), se espera que así funcione en la generación de valores de otras funciones de distribución de probabilidad.

2.3.3. ALGORITMO DE HASTING METRÓPOLIS COMPONENTE A COMPONENTE

En el algoritmo de Hasting Metrópolis componente a componente, el vector del parámetro θ se encuentra dividido en dos subvectores, los cuales serán generados secuencialmente a medida que el algoritmo avance en sus iteraciones (Ntzoufras, 2009).

En cada paso del algoritmo, el valor θ'_j de la j-ésima posición del vector θ es obtenido de la distribución propuesta $q_j(\theta'_j|\theta^{(t-1)})$. El algoritmo sigue los siguientes pasos:

1. Fijar el valor inicial $\theta^{(0)}$
2. Para $t = 1, \dots, T$ seguir los siguientes pasos:
 - a) Fijar $\theta = \theta^{(t-1)}$
 - b) Para $j = 1, \dots, d$
 - I. Generar θ'_j para la j-ésima componente del vector θ' utilizando la distribución propuesta $q(\theta'_j|\theta)$
 - II. Calcular:

$$\alpha = \min \left(1, \frac{f(\theta'_j|\theta_{\setminus j}, y) q(\theta_j|\theta'_j, \theta_{\setminus j})}{f(\theta_j|\theta_{\setminus j}, y) q(\theta'_j|\theta_j, \theta_{\setminus j})} \right)$$
 donde $\theta_{\setminus j}$ es el vector θ excluyendo la j-ésima componente

III. Actualizar $\theta_j = \theta'_j$ con probabilidad α

c) Actualizar $\theta^{(t)} = \theta$

Este algoritmo genera una observación $\theta^{(t)}$ después de haber obtenido todas las componentes del vector de parámetros (Ntzoufras, 2009).

2.3.4. EJEMPLO: MUESTRA DE LA DISTRIBUCIÓN NORMAL BIVARIADA UTILIZANDO EL ALGORITMO DE HASTING METRÓPOLIS COMPONENTE A COMPONENTE

El algoritmo de Hasting Metrópolis presenta dificultades al momento de generar valores de funciones de distribuciones con varias variables, debido a esto el algoritmo de Hasting Metrópolis componente a componente propone realizar la generación de valores para componentes individuales, uno a la vez. Para una mejor ilustración del algoritmo se desarrolla el siguiente ejemplo.

Ejemplo 4. *Se desea generar la función normal bivalente que sigue la variable aleatoria $\theta = (\theta_1, \theta_2)^T$ con parámetros μ, Σ , donde μ es el vector de medias, el parámetro Σ es la matriz de varianza covarianza. La función considerada en el ejemplo es la siguiente:*

$$(\theta_1, \theta_2)^T \rightsquigarrow N[\mu, \Sigma]$$

con parámetros $\mu = (0, 0)^T$, $\Sigma = \begin{pmatrix} 1 & 0,3 \\ 0,3 & 1 \end{pmatrix}$. Este tipo de algoritmo permite facilitar la generación de valores de distribuciones multivariadas, ya que, si se utilizara el algoritmo de Hasting Metrópolis la tasa de rechazo del valor generado es alta y para obtener un vector de una determinada dimensión tomará más tiempo en generarlo.

Se han generado 5000 observaciones con el algoritmo de Hasting Metrópolis y se han obtenido los siguientes resultados:

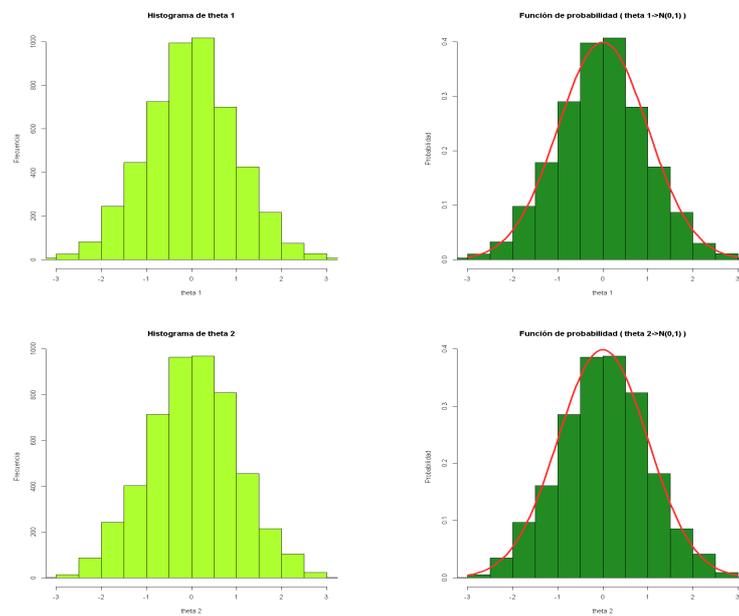


Figura 2.3: Componentes generados de la distribución normal bivariada

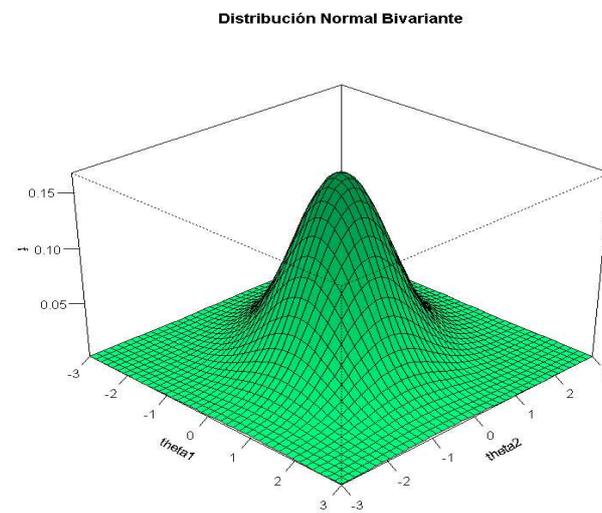


Figura 2.4: Distribución normal bivariada teórica

Se puede notar en los histogramas obtenidos, que las observaciones se ajustan a una

distribución normal de parámetros $(0, 1)$ (esta es la distribución marginal de cada variable simulada). Además mediante las siguientes medidas descriptivas se puede corroborar que las distribuciones generadas se ajustan a la distribución multivariada propuesta.

Medidas descriptivas	θ_1	θ_2
<i>min</i>	-4,412	-3,605
<i>1^{er} cuartil</i>	-0,665	-0,639
<i>mediana</i>	-0,017	0,032
<i>media</i>	-0,030	0,033
<i>3^{er} cuartil</i>	0,613	0,711
<i>max</i>	3,358	3,091
<i>desviación estándar</i>	1,004	0,988
<i>covarianza</i>	0,296	
<i>correlación lineal</i>	0,299	

Mediante los gráficos anteriores se puede decir que la distribución generada es una normal bivariada; a continuación, se usa la prueba de Shapiro Wilk para el contraste de hipótesis:

H_0 : Los datos simulados provienen de una distribución normal

H_1 : Los datos simulados no provienen de una distribución normal

Se ha obtenido el estadístico de Shapiro Wilk correspondiente $W_{\theta_1, \theta_2} = 0,9995$, y el p - valor = 0,2693, a un nivel del 5% se acepta que los datos simulados provienen de una distribución normal bivalente. El algoritmo de Hasting Metrópolis componente a componente ha permitido generar una distribución bivalente con parámetros $\mu = (-0,030, 0,033)^T$, $\Sigma = \begin{pmatrix} 1 & 0,296 \\ 0,296 & 1 \end{pmatrix}$.

2.3.5. MUESTREADOR DE GIBBS

El muestreador de Gibbs es un caso especial del algoritmo de Hasting Metrópolis usando como función de densidad propuesta a $q(\theta'|\theta^{(t)})$ y función de distribución condicional a posteriori $f(\theta_j|\theta_{\setminus j,y})$, donde $\theta_{\setminus j} = (\theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_d)^T$ (Ntzoufras, 2009).

Existen dos principales diferencias entre el Muestreador de Gibbs y el algoritmo de Hasting Metrópolis que son:

1. En el muestreador de Gibbs siempre se acepta un punto candidato
2. Y, es necesario conocer explícitamente las funciones de distribución condicional

Esta última diferencia, hace que el muestreador de Gibbs sea menos aplicable que otros (Ntzoufras, 2009). Por lo general, a este tipo de algoritmo se lo cita como un técnica de simulación diferente debido a su popularidad y ventajas sobre otros algoritmos. Una de las ventajas del muestreador de Gibbs; en cada iteración deben ser generados valores aleatorios de una distribución unidimensional para lo cual, hoy en día se cuenta con innumerables herramientas computacionales para lograr este objetivo, frecuentemente, la forma de la distribución de la cual se desea generar números aleatorios es conocida (Rossi y otros, 2005). Se recalca una desventaja, el muestreador de Gibbs pierde efectividad si los parámetros son “altamente” correlacionados.

El algoritmo utiliza los siguientes pasos:

1. Fijar el valor inicial $\theta^{(0)}$
2. Para $t = 1, \dots, T$ seguir los siguientes pasos:
 - a) Fijar $\theta = \theta^{(t-1)}$
 - b) Para $j = 1, \dots, d$, actualizar θ_j desde $\theta_j \rightsquigarrow f(\theta_j|\theta_{\setminus j,y})$
 - c) Fijar $\theta^{(t)} = \theta$

Dado un estado de la cadena $\theta^{(t)}$, se generan nuevos valores para los parámetros de la siguiente forma:

$$\begin{aligned} \theta_1^{(t)} & \text{ desde } f(\theta_1 | \theta_2^{(t-1)}, \theta_3^{(t-1)}, \dots, \theta_p^{(t-1)}, y) \\ \theta_2^{(t)} & \text{ desde } f(\theta_2 | \theta_1^{(t-1)}, \theta_3^{(t-1)}, \dots, \theta_p^{(t-1)}, y) \\ & \vdots \\ & \vdots \\ \theta_p^{(t)} & \text{ desde } f(\theta_p | \theta_1^{(t-1)}, \theta_2^{(t-1)}, \dots, \theta_{p-1}^{(t-1)}, y) \end{aligned}$$

2.3.6. EJEMPLO: ESTIMACIÓN DE MOMENTOS DE LA FUNCIÓN DE DISTRIBUCIÓN BINOMIAL UTILIZANDO EL MUESTREADOR DE GIBBS

El objetivo es estimar los momentos (media, varianza, etc) de una distribución binomial (para la variable x), se considera la siguiente función de distribución conjunta para las variables x y y :

$$f(x, y) \propto \binom{n}{x} y^{x+\alpha-1} (1-y)^{n-x+\beta-1}$$

Donde $x = 0, 1, \dots, n$, $\alpha, \beta \in \mathbb{R}^+$, además, $0 \leq y \leq 1$.

Se asume: la función de distribución condicional $f(x|y)$ es binomial con parámetros (n, y) , y, la distribución condicional $f(y|x)$ beta con parámetros $x + \alpha$ y $n - x + \beta$.

El problema se enfoca en la estimación de la media y varianza de X , asumiendo que no existe dependencia entre los parámetros α, β, n e ignorando los términos $\alpha - 1, \beta - 1$.

Se fijan los siguientes parámetros: $\alpha = 1,13$, $\beta = 2,36$, $n = 53$, $k = 10000$ iteraciones, $m = 500$ iteraciones eliminadas debido al periodo de inicialización.

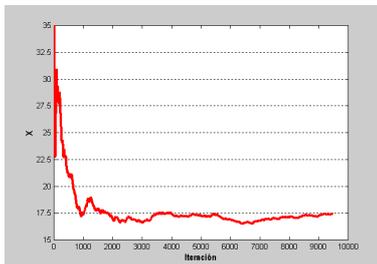


Figura 2.5: Gráfico para la media de X según el número de iteraciones

Para el caso de la variable aleatoria X , tenemos que la media de la variable converge hacia el valor de 17,5 ($\bar{X} = 17,5$)

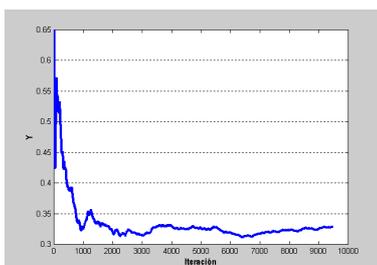


Figura 2.6: Gráfico para la media de Y según número de iteraciones

Ahora para la variable aleatoria Y , tenemos como la media de la variable converge hacia el valor de 0,33 ($\bar{Y} = 0,33$). En el siguiente cuadro se han considerado las siguientes medidas descriptivas

Cuadro 2.2: Medidas descriptivas de las variables simuladas

	X	Y
media	17,40	0,33
mínimo	0	0
máximo	53	0,993
mediana	14	0,272
curtosis	2,42	2,43
sesgo	0,60	0,602

De estas medidas podemos obtener la siguientes conclusiones:

- La distribución de probabilidades de X se encuentra centrada en 17,40, en cambio, la distribución de la variable Y está concentrada en 0,33
- Los valores que la variable X puede tomar están entre 0 y 53, los valores de Y están distribuido entre 0 y 1
- La distribución de X al igual que la de Y señalan que son leptocúrticas, es decir, hay una concentración alta de los valores para cada una de las variables aleatorias alrededor de sus medias
- Las dos distribuciones muestran un sesgo positivo, es decir, los valores de cada una de las variables están concentrados al lado izquierdo de la media

Además de considerar medidas descriptivas, también se hace uso del histograma:

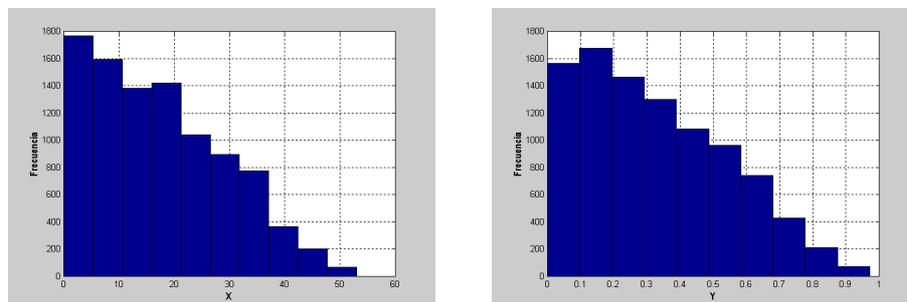


Figura 2.7: Histogramas de las variables X e Y

Todas las conclusiones obtenidas de la interpretación de las medidas descriptivas, se ven confirmadas al momento de evidenciar los histogramas de las variables.

Capítulo 3

ECUACIONES ESTRUCTURALES

3.1. INTRODUCCIÓN

Uno de los objetivos principales al momento de utilizar las técnicas multivariantes es aumentar la capacidad explicativa de las variables consideradas, para esto, la técnica de modelización de ecuaciones estructurales considera variables observables e inobservables; históricamente el análisis factorial es el modelo de ecuaciones estructurales más básico y fue desarrollado por psicometristas (Spearman, Thrustone) para la investigación de las relaciones entre variables; para el caso del presente trabajo, se combina la estimación bayesiana y la modelización de ecuaciones estructurales para de esta forma obtener una herramienta poderosa al momento de analizar temas como el comportamiento de las personas (Byrne, 2009).

En el presente capítulo se abordará la temática de la modelización de ecuaciones estructurales; al inicio se enuncian conceptos básicos junto con la notación que será utilizada, mientras el capítulo avanza se irán desarrollando técnicas, como el análisis factorial exploratorio y el análisis factorial confirmatorio para concluir ampliando los cálculos utilizados para el uso de las técnicas bayesianas en las ecuaciones estructurales específicamente en el análisis factorial confirmatorio.

3.2. ECUACIONES ESTRUCTURALES: CONCEPTOS BÁSICOS

En el campo de la investigación de mercados, especialmente en el trato de clientes, se presentan preguntas como: ¿cuántas variables determinan la satisfacción del cliente?, ¿cómo se encuentran relacionadas las variables de satisfacción del cliente y la imagen de una determinada empresa?, ¿cuál es la variable que más aporta a la satisfacción del cliente?, ¿la satisfacción del cliente está relacionada directamente con la frecuencia de compra?, las cuales son de difícil tratamiento si tan solo se consideran técnicas multivariantes como la regresión múltiple, el análisis discriminante u otras técnicas, debido a que dichas técnicas estudian tan solo una relación al mismo tiempo y para este tipo de problemas lo que se desea es examinar las interacciones entre variables; debido a estos problemas fue desarrollada la técnica de modelización de ecuaciones estructurales como una extensión de las demás técnicas multivariantes. Esta técnica considera la siguiente ecuación lineal $x = \Lambda\xi + \epsilon$, bajo ciertas especificidades, la modelización de ecuaciones estructurales permite la realización del análisis factorial exploratorio y el confirmatorio, el desarrollo de estos modelos se encuentra en las siguientes secciones. Byrne (2009) indica que la técnica de modelización de ecuaciones estructurales tiene las siguientes características:

- Estima relaciones de dependencias múltiples
- Tiene la capacidad de representar conceptos no observados o variables latentes para la estimación de relaciones de dependencia.

La modelización de ecuaciones estructurales estima una serie de regresiones múltiples distintas y a la vez relacionadas mediante la especificación de un modelo estructural.

Esta metodología utiliza los siguientes conceptos:

La *variable latente* es un concepto supuesto, no observado y difícilmente medible directamente que solo puede ser aproximado mediante variables medibles u observables. Este tipo de variables son de gran utilidad al momento de estudiar conceptos teóricos (Byrne, 2009).

El *modelo estructural* es donde se relacionan variables independientes y dependientes:

el conocimiento previo, la experiencia, la teoría, etc. son las directrices que indican la asociación entre variables (Byrne, 2009).

El *modelo de medida* permite usar varias variables con el fin de explicar una única variable dependiente o independiente. Por ejemplo, la variable dependiente puede ser un concepto representado por una escala (por ejemplo escala de Likert). En el modelo de medida el investigador puede evaluar la contribución de cada valor de la escala. Las variables observadas que se obtienen al usar una herramienta como la encuesta se las conoce como indicadores; para representar una variable no observable surge la pregunta ¿cuántos indicadores se necesitan para lograr la mejor aproximación de una variable latente?, su respuesta, como número mínimo se puede usar un solo indicador, sin embargo, esto obliga a que el indicador considerado necesita tener un alta fiabilidad lo cual no siempre es posible asegurar; dada la experiencia el número de indicadores debería ser al menos tres por cada variable latente a ser estimada, hay que aclarar que, al momento de realizar la modelización uno o varios indicadores puede ser utilizados varias veces para aproximar algunas variables latentes (Byrne, 2009).

Los *grados de libertad* de un modelo de ecuaciones estructurales indica si este es identificado¹, para esto se debe tomar en cuenta el tamaño de la matriz de covarianzas o de correlaciones y la cantidad de coeficientes a ser estimados. La diferencia entre el número de correlaciones y el número de coeficientes del modelo estimado se denomina grados de libertad (Byrne, 2009). Estos grados de libertad tienen igual sentido que en la regresión común. Un grado de libertad es un elemento no restringido de la matriz de datos. El número de grados de libertad de un modelo de ecuaciones estructurales se calcula de la siguiente forma:

$$gl = \frac{1}{2} [(p + q)(p + q + 1)] - t \quad (3.1)$$

¹Un modelo se dice que es identificado si sus grados de libertad son 0, bajo esta condición se puede realizar la estimación de los parámetros del modelo.

Donde:

p : número de indicadores endógenos.

q : número de indicadores exógenos.

t : número de coeficientes estimados en el modelo.

Hair (2005) menciona que los grados de libertad deben ser mayores o iguales a cero, para que se pueda estimar el modelo, si se cumple esta condición el modelo está *identificado* o *sobreidentificado*. En base al número de grados de libertad existe la siguiente clasificación:

- $gl = 0$: el modelo está **identificado**. El modelo ofrece un ajuste perfecto pero su estimación no puede ser generalizada.
- $gl > 0$: el modelo está **sobreidentificado**. Es decir que tiene más información en la matriz de datos que el número de parámetros a estimar, este modelo puede ser generalizado.
- $gl < 0$: el modelo está **infraidentificado o subidentificado**. Este modelo no permite hacer estimaciones hasta que se fije o restrinja algún o algunos parámetros en el modelo.

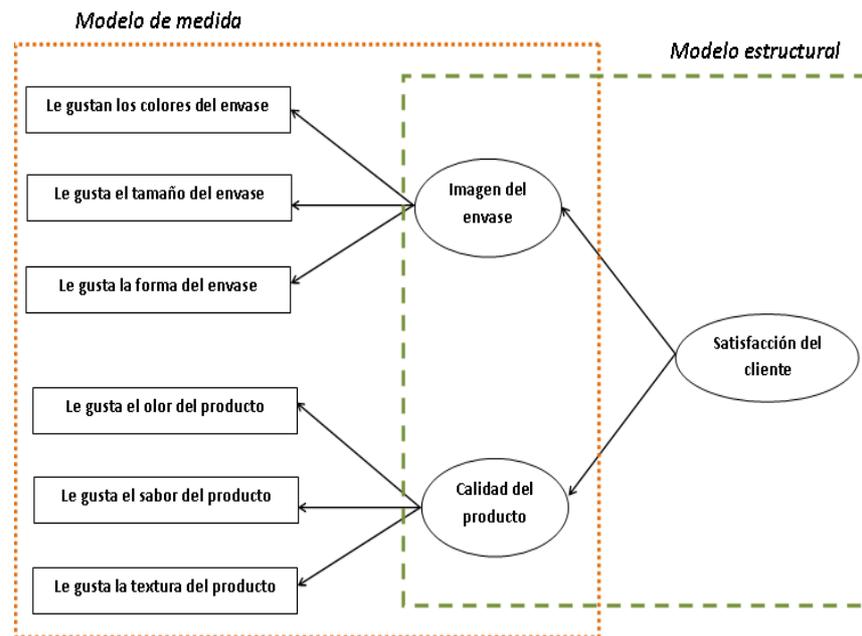


Figura 3.1: Representación gráfica de un modelo de ecuaciones estructurales

El análisis factorial es la forma más básica de la aplicación de la modelización de ecuaciones estructurales, el objetivo es aproximar de la mejor manera todas las variables en un número reducido de factores; el análisis factorial también es utilizado como una técnica exploratoria porque muy a menudo no existen previamente restricciones sobre las cargas de los indicadores. De esto surge el análisis factorial exploratorio, el cual, para un conjunto de indicadores busca la menor cantidad posible de variables latentes no correlacionadas para ser consideradas en la modelización para lograr una mejor comprensión de los datos. Jöreskog desarrolla algunos años después el análisis factorial confirmatorio, en donde se permite tener variables latentes correlacionadas, debido a esta ventaja, se puede fijar cierto número de cargas para los indicadores, correlaciones y varianzas para variables latentes; este tipo de técnica como su nombre lo indica es de carácter confirmatorio, es decir, ya se tiene un conocimiento previo del comportamiento del fenómeno de estudio y puede ser utilizado para confirmar los resultados de la aplicación del análisis factorial exploratorio (Hair y otros, 2005).

3.2.1. ANÁLISIS FACTORIAL EXPLORATORIO

Conocido con las siglas EFA debido a nombre en inglés (*Exploratory Factor Analysis*) es un modelo básico de la técnica de ecuaciones estructurales y es donde se exploran las relaciones entre variables, es un paso previo al análisis factorial confirmatorio, aquí es donde se evidencia la redundancia de variables y correlaciones entre variables latentes; recibió mucha atención durante algunos años hasta el advenimiento del análisis factorial confirmatorio (Hair y otro, 2005).

Definición del modelo

Lee (2007) indica que el análisis factorial exploratorio satisface la siguiente ecuación lineal:

$$x = \Lambda\xi + \epsilon \quad (3.2)$$

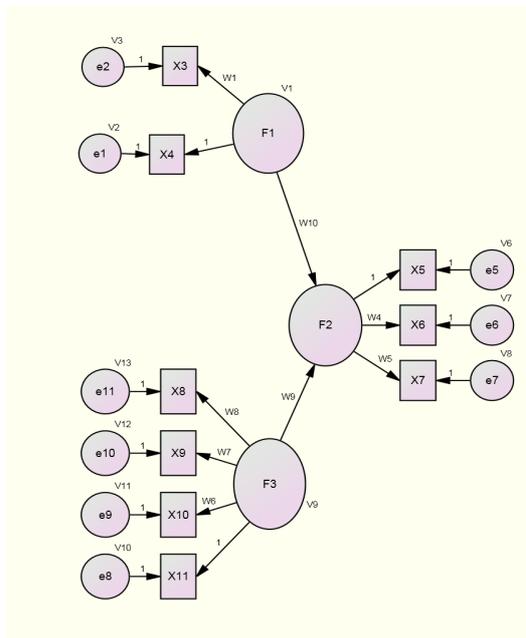
Donde: x : vector aleatorio de dimensión $(p \times 1)$, Λ : es la matriz de cargas de los indicadores y es de dimensión $(p \times q)$, ξ es el vector de variables latentes de dimensión $(q \times 1)$ y finalmente tenemos a ϵ el cual es el vector de errores o residuos y es de dimensión $(p \times 1)$.

Dentro de este tipo de análisis se asume que ξ tiene distribución normal $N(0, I)$ y ϵ tiene distribución normal $N(0, \Psi_\epsilon)$ donde Ψ_ϵ es una matriz diagonal y ξ no está correlacionada con ϵ .

El vector de variables observables x está distribuido con la ley $N(0, \Sigma)$, donde:

$$\Sigma = \Lambda\Lambda^T + \Psi_\epsilon \quad (3.3)$$

El análisis factorial exploratorio busca explicar la mayor proporción de la varianza contenida en las variables dependientes utilizando el número más pequeño posible de factores o indicadores, es decir que, este tipo de análisis busca agrupar variables observables que estén correlacionadas y permitan explicar la misma variable latente (Lee, 2007).



Elaboración: Autor

Figura 3.2: Representación gráfica del análisis factorial exploratorio

En el gráfico se representan: las variables observables (X_i) con sus respectivos términos de error (e_i), las inobservables (F_i), las cargas factoriales (W_i) y varianzas de las variables inobservables (V_i), estos son los elementos del modelo de ecuaciones estructurales; este tipo de representación se la realiza como punto de partida del análisis factorial exploratorio ya que aquí, es donde se incorporan las relaciones y variables que interactuarán en el modelo; se debe tomar en cuenta que en el gráfico no se representan correlaciones (flechas de doble punta) entre variables debido a que dentro del modelo no son consideradas, y las relaciones de influencia son las que dominan la representación gráfica, esta es una de las características de este análisis.

3.2.2. ANÁLISIS FACTORIAL CONFIRMATORIO

El análisis factorial confirmatorio (CFA) es una extensión de análisis factorial exploratorio, dentro de la modelización se tiene el control para especificar todas las interacciones

entre las variables, esto es: las correlaciones entre variables no observables, correlaciones entre indicadores o factores, influencia fija de un factor sobre una variable inobservable (Hair y otros, 2005).

Definición del modelo

Lee (2007) dice que el análisis factorial confirmatorio satisface la siguiente ecuación lineal:

$$x = \Lambda\xi + \epsilon \quad (3.4)$$

Tenemos que las hipótesis y definiciones para x , Λ , ξ y ϵ son las mismas que se consideraron en el análisis factorial exploratorio. En el análisis factorial confirmatorio, las variables latentes pueden ser correlacionadas, de esta forma, ξ sigue la ley normal $N[0, \Phi]$ donde Φ es una matriz de covarianza la cual debe ser definida positiva, bajo la hipótesis que ξ es independiente de ϵ la matriz de covarianza de x viene dada por:

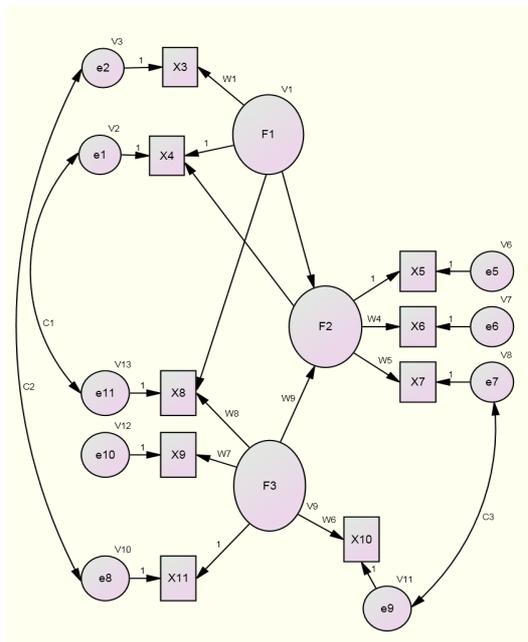
$$\Sigma = \Lambda \Phi \Lambda^T + \Psi_\epsilon \quad (3.5)$$

Una característica muy importante del análisis factorial confirmatorio es que: los elementos de las matrices Λ , Φ , Ψ_ϵ pueden ser fijados, es decir, utilizando conocimiento previo del comportamiento del fenómeno o para contrastar hipótesis puede dar lugar a una asignación de mayor o menor carga (según el criterio del experto) a un determinado indicador para la estimación de una variable latente (Lee, 2007).

Por ejemplo tenemos:

1. Si $\lambda_{kh} = 0$ tenemos que el h –ésimo indicador no influye en la k –ésima variable latente.
2. Si ϕ_{jh} tenemos que el j –ésimo y el h –ésimo indicador no están correlacionados.

En la práctica, al momento de reproducir el análisis factorial confirmatorio se fijan parámetros utilizando conocimientos previos para lograr que el modelo esté identificado.



Elaboración: Autor

Figura 3.3: Representación gráfica del análisis factorial confirmatorio

En la representación gráfica se encuentran: variables observables (X_i), variables latentes (F_i), términos de error (e_i), cargas factoriales (W_i), varianzas de variables latentes (V_i), y covarianzas (C_i), estos últimos elementos diferencian al análisis factorial confirmatorio del factorial exploratorio, todos estos elementos interactúan entre ellos, principalmente las covarianzas permiten que el análisis factorial confirmatorio logre mejores ajustes en comparación al factorial exploratorio.

3.2.3. ANÁLISIS FACTORIAL DE ORDEN SUPERIOR

Debido a que en la práctica se pueden encontrar modelos en donde existen variables latentes generando a otras de la misma naturaleza, es decir, se tienen variables latentes exógenas y endógenas, por esto el análisis factorial confirmatorio se lo ha generalizado para obtener un análisis factorial de orden superior en donde se consideren variables latentes endógenas y exógenas con sus respectivos factores (Lee, 2007).

Definición del modelo

Lee (2007) señala que el primero se construye el análisis factorial de segundo orden para luego realizar la generalización, el análisis factorial de segundo orden viene dado por la siguiente ecuación:

$$x = B(\Lambda \xi + \epsilon) + \delta = B \Lambda \xi + B \epsilon + \delta \quad (3.6)$$

donde: B y Λ son matrices de cargas de los indicadores considerados, ξ es el vector de variables latentes y ϵ y δ son los residuos. Tenemos que: $\xi \rightsquigarrow N[0, \Phi]$, $\epsilon \rightsquigarrow N[0, \Psi_\epsilon]$ y finalmente $\delta \rightsquigarrow N[0, \Psi_\delta]$ donde las matrices Ψ_ϵ y Ψ_δ son diagonales. Además la matriz de covarianza viene dada por:

$$\Sigma(\theta) = B(\Lambda \Phi \Lambda^T + \Psi_\epsilon) B^T + \Psi_\delta \quad (3.7)$$

Los elementos del vector θ pueden ser de tres tipos:

1. Parámetros fijos previamente asignados en base a conocimientos o experiencias anteriores.
2. Parámetros restringidos que son desconocidos parcialmente y tienen algún tipo de relación con parámetros de otros indicadores.
3. Parámetros totalmente libres.

Generalizando se tiene el análisis factorial de orden- n como sigue:

$$x = \Lambda_{(1)}\Lambda_{(2)} \dots \Lambda_{(k)}\xi_k + \dots + \Lambda_{(1)}\Lambda_{(2)}\xi_2 + \Lambda_{(1)}\xi_1 \quad (3.8)$$

donde: $\Lambda_{(1)}, \Lambda_{(2)}, \dots, \Lambda_{(k)}$ son las matrices de cargas para los indicadores, además ξ_1, \dots, ξ_k son vectores aleatorios independientes idénticamente distribuidos. La matriz de covarianza de x viene dadas por:

$$\Sigma = \Lambda_{(1)} \dots \Lambda_{(k)} \Phi \Lambda_{(1)}^T \dots \Lambda_{(k)}^T + \dots + \Lambda_{(1)} \Phi \Lambda_{(1)}^T \quad (3.9)$$

donde Φ_1, \dots, Φ_k son matrices de covarianza de ξ_1, \dots, ξ_k respectivamente. Los elementos de las matrices $\Lambda_{(1)}, \dots, \Lambda_{(k)}$ y Φ_1, \dots, Φ_k pueden ser fijados a cualquier valor.



Fuente: *The American Customer Satisfaction Index Organization*

Elaboración: Autor

Figura 3.4: Representación gráfica del modelo estructural para la elaboración de un análisis factorial confirmatorio de orden superior

El objetivo es construir un índice de satisfacción del cliente. Este es el modelo ACSI (*American Customer Satisfaction Index*) aplicado en Estados Unidos para la estratificación de empresas de diferentes tipos en base al índice calculado. Este gráfico fue obtenido del siguiente enlace http://www.theacsi.org/images/stories/images/about/model_lg.gif. Este tipo de modelo se ha vuelto común debido a que el índice de satisfacción obtenido se ha convertido en un estándar de comparación para empresas e incluso entidades gubernamentales en Estados Unidos, este índice es calculado anualmente en donde se utiliza una muestra de consumidores según la empresa que se esté evaluando y luego se publican los resultados. En el capítulo siguiente se hará un abordaje completo al modelo ACSI (*The ACSI*, 2012).

3.3. ESTIMACIÓN BAYESIANA DE ECUACIONES ESTRUCTURALES

En el enfoque bayesiano un parámetro es tomado como una variable aleatoria y la inferencia respecto a sus posibles valores se obtiene aplicando el cálculo de probabilidades para determinar la distribución del parámetro condicionada a la información disponible. Este proceso se lo realizará varias veces durante la estimación de parámetros de ecuaciones estructurales, en esta sección se presenta a la estimación bayesiana aplicada a un modelo de análisis factorial confirmatorio. Existen casos en donde no es posible utilizar la distribución a priori o en donde el cálculo de una distribución a posteriori simplemente no se lo puede obtener directamente debido a su complejidad, para esto se utilizan los métodos de Monte Carlo con Cadenas de Markov, para los cuales si se realiza una simulación con un número suficientemente grande de observaciones para una función de densidad $p(\theta|Y)$ se pueden aproximar: la media, la varianza, la mediana, etc.

Para la mayoría de modelos de ecuaciones estructurales obtener una distribución a posteriori es sumamente complicado, por consiguiente realizar una simulación de esta distribución se torna una tarea complicada por no decir imposible, para lograr obtener observaciones de la distribución a posteriori se utiliza el aumento de datos, es decir, a las variables no observables se las trata como si fuesen datos perdidos, para así aumentar o agregar datos observados, de esta forma el análisis de la distribución a posteriori se facilita.

La estrategia de aumentar datos da una aproximación muy útil, para resolver el problema debido a la inclusión de variables latentes en un modelo, lo que hace esta estrategia es que en vez de utilizar una función de densidad a posteriori intratable $p(\theta|Y)$ se utiliza la función $p(\theta, \Omega|Y)$ donde Ω es el conjunto de variables latentes consideradas en el modelo. Es común que $p(\theta, \Omega|Y)$ presente dificultad para tratarla directamente, es aquí en donde los métodos de Monte Carlo con Cadenas de Markov intervienen para realizar la simulación con el fin de obtener observaciones de $p(\theta, \Omega|Y)$ y así conseguir mediante un proceso iterativo observaciones de las funciones de densidad condicionales $p(\theta|\Omega, Y)$ y $p(\Omega|\theta, Y)$.

Para un modelo M , sea $\theta = (\theta_1, \dots, \theta_a)$ un vector de parámetros, sea $\Omega = (\Omega_1, \dots, \Omega_b)$ matriz de variables latentes, el algoritmo del muestreador de Gibbs se mueve a través de las componentes de θ y Ω generando nuevas observaciones condicionadas a la información a priori. En la j - ésima iteración se tiene los siguientes valores: $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_a^{(j)})$, $\Omega^{(j)} = (\Omega_1^{(j)}, \dots, \Omega_a^{(j)})$ y en la $j + 1$ - ésima iteración se tiene:

$$\begin{array}{l}
\theta_1^{(j+1)} \quad \text{desde} \quad p(\theta_1 | \theta_2^{(j)}, \dots, \theta_a^{(j)}, \Omega^{(j)}, Y) \\
\theta_2^{(j+1)} \quad \text{desde} \quad p(\theta_2 | \theta_1^{(j)}, \dots, \theta_a^{(j)}, \Omega^{(j)}, Y) \\
\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\
\theta_a^{(j+1)} \quad \text{desde} \quad p(\theta_a | \theta_1^{(j)}, \dots, \theta_{a-1}^{(j)}, \Omega^{(j)}, Y) \\
\Omega_1^{(j+1)} \quad \text{desde} \quad p(\Omega_1 | \theta^{(j+1)}, \Omega_2^{(j)}, \dots, \Omega_b^{(j)}, Y) \\
\Omega_2^{(j+1)} \quad \text{desde} \quad p(\Omega_2 | \theta^{(j+1)}, \Omega_1^{(j)}, \dots, \Omega_b^{(j)}, Y) \\
\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\
\Omega_b^{(j+1)} \quad \text{desde} \quad p(\Omega_b | \theta^{(j+1)}, \Omega_1^{(j)}, \dots, \Omega_{b-1}^{(j)}, Y)
\end{array}$$

Existen $a + b$ pasos en la $j + 1$ - ésima iteración del muestreador de Gibbs. En cada paso las componentes de θ y Ω son calculadas y actualizadas considerando las demás componentes que ya han sido actualizadas. En la simulación se puede optar primero por simular las componentes de θ y luego las de Ω , o viceversa.

3.3.1. ANÁLISIS FACTORIAL CONFIRMATORIO USANDO ESTIMACIÓN BAYESIANA

Sea el modelo de análisis factorial confirmatorio dado por la ecuación:

Para $i = 1, \dots, n$

$$x_i = \Lambda \xi_i + \epsilon_i \quad (3.10)$$

la cual es análoga a la ecuación 3.4. Donde x_i es el vector aleatorio de variables observadas de dimensión $p \times 1$, Λ es la matriz de factores de dimensión $p \times q$, ξ_i es el vector de cargas

factoriales de dimensión $q \times 1$ y ϵ_i es el vector de errores, además, se supone ξ_i y ϵ_i no están correlacionados, $\epsilon_i \rightsquigarrow N[0, \Psi_\epsilon]$, Ψ_ϵ es una matriz diagonal, $\xi_i \rightsquigarrow N[0, \Phi]$, y Φ es una matriz definida positiva.

Sea $X = [x_1, \dots, x_n]$ el vector de datos observados, $\Omega = (\xi_1, \dots, \xi_n)$ el vector de coeficientes (cargas factoriales) y θ es el vector de parámetros estructurales en donde están los elementos desconocidos de Λ , Φ , Ψ_ϵ del modelo. Se asume que el modelo de ecuaciones estructurales está identificado. A las cargas factoriales de las variables latentes que se encuentran en Ω se las tratará como datos perdidos, y aumentar el conjunto de datos observado X con Ω en el análisis a posteriori.

Una muestra suficientemente “grande” de (θ, Ω) es obtenida mediante la simulación de la distribución a posteriori $[\theta, \Omega | X]$ utilizando el muestreador de Gibbs. En la $j + 1$ –ésima iteración del algoritmo y con valores $\Omega^{(j)}$, $\Psi_\epsilon^{(j)}$, $\Lambda^{(j)}$, $\Phi^{(j)}$:

- (I) Generar $\Omega^{(j+1)}$ utilizando $p(\Omega | \Psi_\epsilon^{(j)}, \Lambda^{(j)}, \Phi^{(j)}, X)$
- (II) Generar $\Psi_\epsilon^{(j+1)}$ utilizando $p(\Psi_\epsilon | \Omega^{(j)}, \Lambda^{(j)}, \Phi^{(j)}, X)$
- (III) Generar $\Lambda^{(j+1)}$ utilizando $p(\Lambda | \Psi_\epsilon^{(j)}, \Omega^{(j)}, \Phi^{(j)}, X)$
- (IV) Generar $\Phi^{(j+1)}$ utilizando $p(\Phi | \Psi_\epsilon^{(j)}, \Lambda^{(j)}, \Omega^{(j)}, X)$

Cada una de estas distribuciones deben estar implementadas en el muestreador de Gibbs de tal forma que después de realizar la simulación y haber eliminado iteraciones de calentamiento se tenga todos los parámetros requeridos. Debido a que el proceso es simulado, es necesario considerar una medida de convergencia que permita decidir si las estimaciones que se van obteniendo son buenas o se las debe rechazar, para esto se define el siguiente estadístico.

Definición 13. *Estadístico de Convergencia*

Conocido como C.S. por sus siglas en inglés, el estadístico de convergencia se basa en la

idea de que para un punto después de un número grande de observaciones simuladas el error estándar asociado el método de Monte Carlo utilizado es cercano a cero. El estadístico de convergencia es una medida de cuanto se va reduciendo la incertidumbre en la estimación obtenida, para un número infinito de observaciones el estadístico de convergencia deberá ser cercano a 1. El valor del estadístico de convergencia dependerá del fenómeno de estudio, para algunos fenómenos en los cuales no es necesario una alta precisión un estadístico asociado de 1,1 sería adecuado, es decir, la convergencia dependerá del fenómeno, de los datos incorporados, del número de parámetros a estimar y de la precisión requerida por el estudio.

$$C.S = \sqrt{1 + \frac{SE^2}{SD^2}} \quad (3.11)$$

SE: estimación del error estándar, es una medida de la variabilidad de esperanza del parámetro a estimar.

SD: estimación de la desviación estándar después de haber eliminado las iteraciones consideradas en el periodo de calentamiento.

El estadístico de convergencia global será el máximo valor de los estadísticos de convergencia de cada una de las estimaciones consideradas en el modelo.

Capítulo 4

APLICACIÓN: CONSTRUCCIÓN DE UN ÍNDICE DE SATISFACCIÓN DEL CLIENTE CONSIDERANDO EL MODELO ACSI

4.1. INTRODUCCIÓN

En la vida cotidiana, una de las preocupaciones de una empresa puede ser el nivel de satisfacción que tienen sus clientes, por ejemplo Papa John's ® es una cadena de restaurantes la cual busca lograr calificaciones altas en cuanto a satisfacción de sus compradores, debido a la naturaleza de su negocio (pizzería) un nivel alto de satisfacción de sus comensales le reporta un aumento en sus ventas. Como esta empresa existen muchas, que tienen como principal meta mantener niveles altos de satisfacción de sus clientes, pero, ¿Qué es la satisfacción del cliente?, ¿Cómo se mide la satisfacción del cliente?, ¿Qué se debe hacer para aumentar la satisfacción del cliente? son algunas de las preguntas que se resolverán a lo largo del presente capítulo.

Se abordará el modelo ACSI (*American Customer Satisfaction Index*) el cual es usado internacionalmente como un estándar de comparación entre empresas de un mismo segmento de mercado (restaurantes, entidades gubernamentales, etc.), además se planteará el modelo para una empresa ecuatoriana para finalmente concluir acerca del ajuste del modelo y principalmente dar a notar las cualidades que son más influyentes a la hora de lograr satisfacer a los clientes (*The ACSI*, 2012).

4.2. FILOSOFÍA DE LA EMPRESA: EL ENFOQUE MARKETING

La gestión del marketing empresarial está basada en la constante búsqueda de intercambio con un público definido. Debido a esto, surgen las siguientes preguntas: ¿Cuán importante es la satisfacción de los intereses de la organización y de los clientes?, ¿Qué enfoque debe guiar los esfuerzos de la empresa? (Kotler, 2001).

Los intereses de la empresa y de los clientes muchas veces no se encuentran en un punto de equilibrio originando una pugna entre ellos, por esta razón, la empresa debe desarrollar las actividades del marketing bajo una dirección o enfoque. Existen cinco enfoques que pueden direccionar las actividades de la empresa:

1. **Enfoque producción:** sostiene que los consumidores favorecerán aquellos productos que estén: muy disponibles y sean de bajo costo. Los directores de organizaciones con enfoque producción concentran sus esfuerzos en alcanzar economías de escala¹ y amplia distribución (Kotler, 2001).
2. **Enfoque producto:** sostiene que los consumidores favorecerán aquellos productos que ofrezcan mayor calidad o mejores resultados. Los directivos de las empresas con enfoque producto centrarán sus esfuerzos en hacer buenos productos y mejorarlos a lo largo del tiempo (Kotler, 2001).
3. **Enfoque ventas:** indica que si a los consumidores no se les “empuja”, no comprarán suficientes productos de la empresa. Por lo tanto, la organización debe llevar políticas agresivas de venta y promoción.
4. **Enfoque marketing:** sostiene que la clave para alcanzar los objetivos de la organización consiste en identificar las necesidades y deseos del público objetivo y entregar los bienes de consumo que cubran las necesidades del mercado objetivo, de una forma más efectiva y eficiente que la competencia (Kotler, 2001).

¹Economía de escala: Cualquier situación de producción, incluso la prestación de servicios financieros, en la que el coste por unidad producida disminuye a medida que aumenta el número de unidades producidas.

5. **Enfoque marketing social:** supone que la tarea de las organizaciones es identificar las necesidades, deseos e intereses de sus públicos objetivos, satisfaciéndolos de manera más efectiva que la competencia y de forma que se preserve o realce el bienestar a largo plazo de los consumidores y de la sociedad (Kotler, 2001).

El enfoque marketing y marketing social se caracterizan por dar gran importancia a la identificación de necesidades y deseos del público, el presente trabajo se encuentra orientado en base al enfoque marketing donde la organización está direccionada a la búsqueda de satisfacción del cliente intercambiando ya sean productos o servicios (Kotler, 2001).

4.2.1. EL ENFOQUE MARKETING

Está basado en cuatro pilares:

Definición del mercado: las empresas tienen que definir su público objetivo cuidadosamente; éste es el mercado. Ninguna compañía puede operar en todos los mercados y satisfacer todas las necesidades. Tampoco puede hacer un buen trabajo en un mercado demasiado extenso (Kotler, 2001).

Orientación al cliente: requiere que la empresa defina las necesidades del cliente desde el punto de vista de este último y no desde el de la empresa. Cada producto presenta múltiples aspectos que la dirección no puede conocer sin investigar al consumidor potencial. El objetivo es intercambiar (vender, alquilar, concesionar, etc.) a través de la satisfacción de las necesidades del cliente (Kotler, 2001).

Coordinación del marketing: desafortunadamente no todos los empleados están entrenados o motivados para trabajar con el cliente. Kotler (2001) menciona que la coordinación del marketing significa dos cosas:

- a. Coordinación entre las áreas de la empresa
- b. Inclusión; los empleados deben conocer del efecto de sus actuaciones sobre la satisfacción del cliente.

Rentabilidad: el propósito del enfoque marketing es ayudar a las organizaciones a alcanzar sus objetivos, que en el caso de empresas privadas es maximizar el beneficio, o en el caso de organizaciones no lucrativas es sobrevivir y atraer fondos para desarrollar sus actividades. Así pues, el objetivo no es el beneficio por el beneficio, sino alcanzarlo como una consecuencia de un buen trabajo realizado (Kotler, 2001).

Las empresas hacen dinero satisfaciendo las necesidades de los consumidores mejor que la competencia. El objetivo es encontrar el camino de obtener rentabilidad a través de la satisfacción de un determinado grupo de personas.

4.3. SATISFACCIÓN DEL CLIENTE: DEFINICIÓN Y BENEFICIOS DE LOGRARLA

Existen indicadores como por ejemplo los financieros, los cuales ayudan a tener una imagen de la posición en la que está el mercado, sin embargo, estas no son suficientes, se necesitan medidas cualitativas que alerten a tiempo de la situación en la que sobrevive una organización, las compañías buscan medir la satisfacción estableciendo sistemas de control de actitudes tanto de sus clientes y como de miembros de la organización involucrados (Kotler, 2001).

Para lograr identificar los cambios en las preferencias y satisfacción de los consumidores principalmente se utilizan los siguientes sistemas:

Sistema de quejas y sugerencias: mediante este sistema se recoge, analiza y responde a las quejas de los clientes. Como un ejemplo, en entidades gubernamentales como el Hospital Carlos Andrade Marín se cuenta con buzones y papeletas disponibles para que puedan quejarse los afiliados, este buzón es únicamente abierto por personal autorizado que tratará de solucionar los problemas manejables con el fin de evitar nuevos reclamos.

Paneles de consumidores: algunas compañías utilizan esta herramienta como un medio más directo y representativo de obtención de información, el sistema consiste en orga-

nizar un grupo de consumidores previamente puestos de acuerdo para compartir sus actitudes y experiencias con un determinado producto o servicio (Promonegocios, 2012).

Encuestas a clientes: consiste en obtener información de una muestra representativa de clientes mediante la contestación de un formulario previamente diseñado donde se valoren los atributos de interés para la organización. Este tipo de instrumento recolecta información mediante preguntas que usan escalas como por ejemplo las escalas de cinco puntos (muy satisfecho, satisfecho, ni satisfecho ni insatisfecho, insatisfecho, muy insatisfecho).

Compradores “Espías”: consiste en usar personal eventual pagado por la propia empresa para que se involucre desde: la compra, el uso del producto o servicio, exposición a promociones y trato del personal de la organización para luego expresar sus experiencias vividas durante el proceso. Las ventajas de este método son: su costo bajo y la información de cada área a la que se expuso directamente a este tipo de persona (Promonegocios, 2012).

Análisis de clientes perdidos: es un método muy eficiente debido a que este tipo de cliente conoce el problema que lo llevó a cambiar de organización. Es de gran utilidad para identificar los puntos más débiles de la empresa o producto o servicio. Conocer el punto de vista de estos clientes y aplicar los correctivos necesarios evitará más pérdidas de clientes por insatisfacción (Promonegocios, 2012).

4.3.1. ¿QUÉ ES LA SATISFACCIÓN DEL CLIENTE?

El tema de la satisfacción ha tomado una gran importancia en las empresas de hoy en día principalmente en las privadas; no existe una definición precisa de lo que significa “satisfacción del cliente”; para el desarrollo del presente trabajo se tomará la siguiente definición:

La satisfacción del cliente es el nivel del estado de ánimo de una persona que resulta de comparar el rendimiento percibido de un producto o servicio con sus expectativas.

Philip Kotler.

El objetivo de tener clientes complacidos o plenamente satisfechos es uno de los factores clave para alcanzar el éxito en los negocios debido a que un cliente satisfecho regresará a comprar (puede ser otro tipo de intercambio) aumentando así la fidelidad y lealtad del cliente hacia la empresa (Kotler, 2001).

Ejemplo 5. *El secreto de la rentabilidad de L. L. Bean (Kotler, 2001)*

Una de las más exitosas empresas de venta por correo es la empresa L. L. Bean, que se especializa en vestidos. Ha coordinado su marketing interno y externo. A sus clientes les ofrece lo siguiente:

100 % de GARANTÍA

Todos los productos están garantizados para dar un 100 % de satisfacción en cualquier forma que se usen; devuélvanos cualquier compra si no está de acuerdo con ella y se la reemplazaremos, le devolveremos su dinero o su crédito, como prefiera. No queremos que tenga nada de la empresa Bean que no sea completamente satisfactorio.

Para motivar a sus empleados en la satisfacción adecuada de sus clientes, la empresa Bean expone el siguiente cartel en sus oficinas:

¿Qué es un cliente?

Un cliente es la persona más importante en esta oficina... en persona o en correo...

Un cliente no depende de nosotros... nosotros dependemos de él.

Un cliente no es una interrupción en nuestro trabajo... es el propósito de nuestro trabajo.

No le hacemos un favor al servirle... él nos hace un favor al darnos la oportunidad de servirle.

Un cliente no es una persona con quién discutir. Nadie debe ganar una discusión con un cliente.

Un cliente es una persona que nos aporta sus deseos. Es nuestro trabajo satisfacerlos rentablemente para él y para nosotros.

En el ejemplo se puede notar que la dirección de la empresa L.L. Bean gira en torno a la satisfacción del cliente; todo dependerá de los objetivos de la misma. El cliente es visto como la principal fuente de dinero, por tal motivo se le da una importancia elevada a su satisfacción plena.

4.3.2. BENEFICIOS DE LOGRAR ALTOS NIVELES DE SATISFACCIÓN EN EL CLIENTE

Los beneficios obtenidos al lograr niveles altos de satisfacción en clientes se clasifican de la siguiente forma:

Buenas referencias de la empresa: un nivel alto de satisfacción en un cliente ayuda a que este haga buenas referencias hacia otros posibles clientes ayudando a que la organización crezca (Kotler, 2001).

Lealtad: un cliente satisfecho vuelve a comprar en la empresa generando así la posibilidad de conocer y adquirir nuevos productos o servicios (Kotler, 2001).

Fidelidad: necesidades satisfechas generan comodidad en un cliente logrando que este se

sienta cómodo y se resista al cambio de empresa o proveedor; el cliente guarda fidelidad (Kotler, 2001).

Algunas formas de medir cuán satisfechos se encuentran los clientes son: sumar las escalas positivas para una pregunta donde se evalúe la satisfacción obtenida por un determinado atributo para luego obtener un promedio, utilizar el cuestionario SERVQUAL² para medir la satisfacción obtenida por atributos de un determinado servicio, utilizar el modelo *ACSI* el cual es un modelo de ecuaciones estructurales donde se miden diversos atributos según la naturaleza del negocio a evaluar (*The ACSI*, 2012).

4.4. MODELO *ACSI*: AMERICAN CUSTOMER SATISFACTION INDEX

El Índice de Satisfacción del Consumidor Americano es un indicador que mide la satisfacción de los consumidores especialmente en el mercado estado unidense, el *ACSI* utiliza alrededor de ochenta mil entrevistas anuales a consumidores para obtener información acerca de los bienes y servicios utilizados como fuente de datos para la aplicación de un modelo econométrico desarrollado por la Escuela de Negocios de la Universidad de Michigan. En el modelo *ACSI* se busca representar la causa y efecto mediante el uso de índices (*The ACSI*, 2012). El modelo *ACSI* surge como una adaptación del modelo *SCSB*³ por sus siglas en inglés que fue implementado en la economía sueca para el año 1989. Los modelos *SCSB* y *ACSI* fueron desarrollados por Claes Fornell⁴ (*The ACSI*, 2012).

El *ACSI* utiliza dos métodos relacionados para medir la satisfacción del cliente: la entrevista al consumidor y la modelación econométrica:

²SERVQUAL: es un cuestionario donde el cliente da información acerca del nivel de satisfacción brindado por atributos de un servicio, fue elaborado en Estados Unidos por el *Marketing Science Institute* para obtener información con un alto nivel de fiabilidad y validez acerca de un servicio

³SCSB: Swedish Customer Satisfaction Barometer

⁴Claes Fornell: es considerado uno de los expertos mundiales en cuanto a la satisfacción del consumidor. Ha fundado empresas como *ACSI*, *CFI Group*, *ForSee Results*, etc. Las cuales se dedican a la investigación de mercados.

- a. La entrevista es realizada ya sea vía correo electrónico, personal o telefónicamente a consumidores o clientes previamente seleccionados en una muestra aleatoria utilizando el método *Random Digit Dialing* el cual genera números de teléfono aleatorios asegurando una representatividad de la población a ser estudiada (*The ACSI*, 2012).
- b. Todos los datos recolectados son analizados utilizando un modelo de ecuaciones estructurales para extraer puntajes para las variables latentes consideradas y correlaciones para las variables observables.

El resultado obtenido del uso de esta metodología es un índice, el cual está en una escala entre 0 a 100, este índice es de gran utilidad para alrededor de 225 compañías, 45 industrias estadounidenses, debido a que la utilizan para medir la satisfacción de los usuarios convirtiéndola en un estándar al momento de compararse diferentes empresas de un mismo sector, por ejemplo, en el sector de las compañías dedicadas a la elaboración de *software* se compara a Microsoft ® con el resto de empresas (*The ACSI*, 2012).

4.4.1. MODELOS DE ECUACIONES ESTRUCTURALES PARA LA ESTIMACIÓN DE LA SATISFACCIÓN DEL CLIENTE

Uno de los objetivos del presente trabajo es la construcción de una variable donde se pueda identificar el nivel de satisfacción de los clientes de una empresa ecuatoriana, para lo cual se utilizará el modelo *ACSI*. El estudio fue hecho por un medio de comunicación nacional que desea medir la satisfacción de sus informados, para esto es necesaria la realización de un análisis factorial exploratorio previo para considerar cuales son los indicadores necesarios para estimar los factores latentes y finalmente la satisfacción; al final se concluirá el trabajo con un análisis factorial confirmatorio donde se usará la información obtenida previamente y se obtendrán niveles de satisfacción.

Los dos tipos de análisis serán realizados considerando cada uno de ellos una muestra de 210 clientes que respondieron a un cuestionario⁵ que contiene preguntas de los siguientes

⁵En el Anexo 4 se puede observar el cuestionario considerado para obtener información de la satisfacción de los informados

atributos: credibilidad, cantidad de información, promociones, identificación del cliente con el producto, diseño, colores, variedad de información, confiabilidad, precio, etc.

Para mayor comprensión del modelo planteado por Claes Fornell se procede a dar una explicación breve de cada uno de los factores considerados para la medición de la satisfacción del cliente.

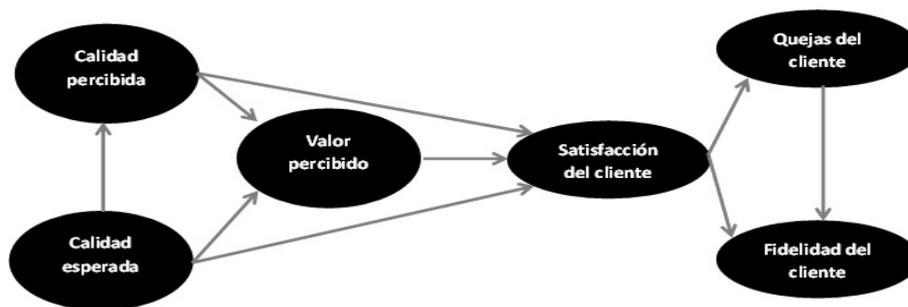


Figura 4.1: Representación gráfica del modelo ACSI.

- 1) **Calidad percibida:** es una medida que representa de la calidad obtenida de los atributos del producto o servicio por parte del cliente. Esta medida considera un uso reciente el producto o servicio en evaluación. La calidad del producto o servicio está relacionada con la satisfacción de las necesidades del cliente y la disponibilidad del producto o servicio (*The ACSI, 2012*).
- 2) **Calidad esperada:** es una medida de la calidad que el cliente piensa que va a obtener al adquirir un producto o servicio determinado. Esta medida representa la experiencia previa con el bien a evaluar (si no se ha tenido experiencia en cuanto al uso del bien también es representada en esta medida), referencias de cualquier tipo (comentarios a favor o en contra del bien, experiencias de otras personas, etc) acerca del bien (*The ACSI, 2012*).
- 3) **Valor percibido:** es una medida que relaciona la calidad del bien percibida por el cliente y el precio pagado por adquirirlo. El precio es muy importante si el cliente va a realizar la primera compra (*The ACSI, 2012*).

- 4) **Quejas del cliente:** es el porcentaje de personas que han realizado una queja directamente a la empresa o proveedor de bien en un intervalo del tiempo (e.g. un mes). Si el porcentaje de personas con quejas aumenta la satisfacción del cliente disminuirá (*The ACSI*, 2012).
- 5) **Fidelidad del cliente:** es una medida que toma en cuenta la proporción de personas que señalan que no cambiarían: de proveedor, de bien para reemplazarlo con otro, disposición a pagar más por el mismo bien. La fidelidad es un factor de importancia debido a que está relacionado directamente con la rentabilidad del bien o servicio provisto (*The ACSI*, 2012).

No será considerado el impacto de quejas del cliente debido a que carece de sentido evaluar un servicio que no se provee al consumidor directo, además, el factor de fidelidad no puede medirse únicamente estimando los efectos de la satisfacción del cliente debido a que se estaría limitando la medición haciendo que esta no sea objetiva y refleje el mismo comportamiento de la variable de satisfacción del cliente, por lo tanto se considera un modelo reducido en estos dos factores, en la siguiente figura se representa gráficamente el modelo reducido que se utilizará para la estimación de la satisfacción del cliente.

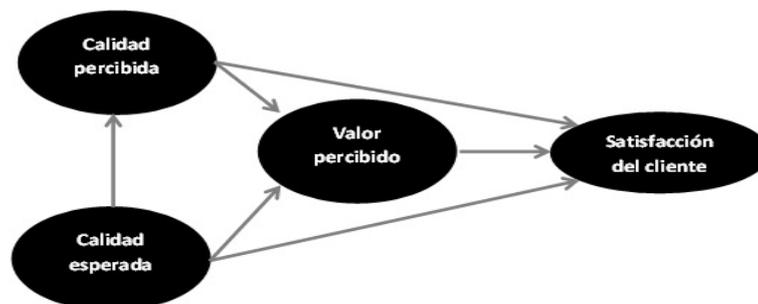


Figura 4.2: Representación gráfica del modelo planteado utilizando la metodología ACSI.

Para la realización del análisis factorial confirmatorio se necesita considerar un modelo ya definido y ajustado con datos semejantes a los que serán utilizados para este análisis, para ello es necesario plantear el análisis factorial exploratorio y obtener el modelo ajustado para después realizar su confirmación con otros datos.

ANÁLISIS FACTORIAL EXPLORATORIO

Debido a la naturaleza del Análisis Factorial Exploratorio (AFE), el objetivo principal de esta sección es la búsqueda de la menor cantidad de indicadores posibles que permitan la identificación del modelo, hay que recalcar que este modelo es la base del Análisis Factorial Confirmatorio (AFC), y por lo tanto, el éxito de la modelación final en el AFC solamente se dará cuando el AFE considere un modelo adecuado y con la representación de correlaciones en el AFC se obtenga un modelo aceptable (bajo las medidas de ajuste).

Para establecer una hoja de ruta, se definen los pasos realizados en la modelización:

- 1) Obtener una base de datos en donde se encuentren las variables de estudio totalmente depuradas, es decir, que no existan valores perdidos, valores atípicos, etc.
- 2) Realizar un análisis descriptivo de las variables depuradas, con la finalidad de entender el comportamiento de cada una de ellas.
- 3) Modelizar y utilizar el método de máxima verisimilitud para establecer un punto de partida para iniciar el proceso de simulación considerando los Métodos de Monte Carlo.
- 4) Simular el modelo obtenido hasta lograr estimaciones de ecuaciones estructurales fiables (aquí se considera el estadístico de convergencia).
- 5) Concluir acerca de la validez del modelo y las estimaciones obtenidas.

Desarrollo del análisis factorial exploratorio:

1) Base de datos depurada

La información fue recolectada mediante un cuestionario a una muestra aleatoria simple de personas que han recibido información mediante el medio de comunicación, previamente se han definido 13 indicadores que forman el modelo de ecuaciones estructurales, aquí es donde, se combina la experiencia y la información en la estructura del cuestionario utilizado, se ha considerado lo siguiente:

- Recolección 210⁶ casos (encuestas)
- 13 variables observables (indicadores)
- 16 variables inobservables correspondientes a los términos de error
- 4 variables latentes (factores)
 - 1 variable latente exógena (calidad esperada)
 - 2 variables latentes exógenas y endógenas a la vez (calidad percibida y valor percibido)
 - 1 variable latente endógena (satisfacción)

Para dar una mejor idea del modelo planteado inicialmente considerando la teoría y la experiencia, a continuación, se expone una representación gráfica

⁶Se realizó un muestreo aleatorio simple, en el Anexo 4, se encuentran los cálculos previos para la determinación del tamaño muestral

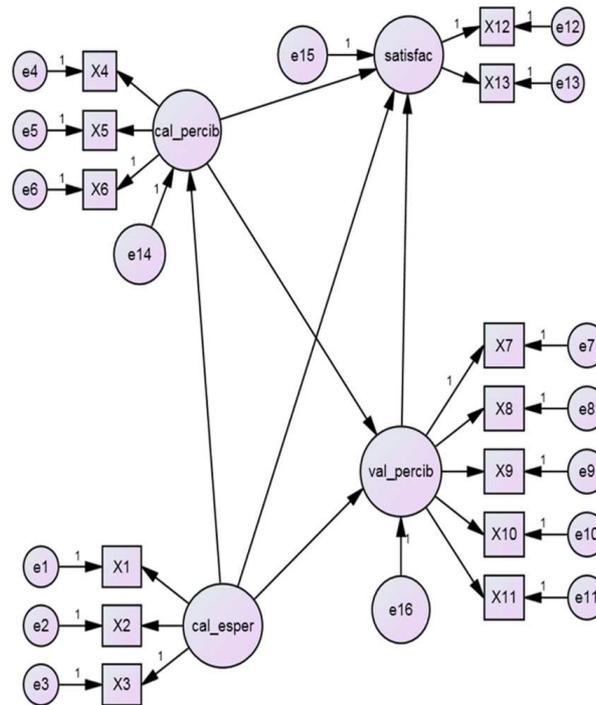


Figura 4.3: Representación gráfica del modelo utilizado para la realización del análisis factorial exploratorio.

Como se puede notar, se encuentran las variables observables asociadas a cada variable latente o factor a estimar, además, se consideran las variables de error para cada una de las variables observables y las no observables endógenas, esto se debe a que no se puede asegurar, en este caso, que las mediciones obtenidas son totalmente fiables y por ende se encuentra implícito el término de error. Dentro de la modelización se aplica la restricción de igualdad a cero ($E(e_i) = 0 \forall i = 1, \dots, 16$) para cada término de error, esto debido a su naturaleza de error de medida.

2) Análisis Descriptivo

Las variables consideradas en el levantamiento de información son escalas de Likert, y para efecto de la modelación se asumirá que son continuas a pesar de ser variables ordinales discretas, este criterio es debido a la técnica utilizada y se ha recodificado las

variables para asegurar que todas estas se encuentren en escala creciente sus atributos con relación a la satisfacción (a mayor valor de X_i mayor valor en *satisfacción*).

Cuadro 4.1: Estadísticos Descriptivos

Variable	Mínimo	Máximo	Media	Mediana	Desv. típ.	C.A*	Curtosis
X1	1	7	5,943	6	1,278	-1,212	0,877
X2	1	7	5,586	6	1,371	-1,131	0,971
X3	1	7	5,514	6	1,623	-0,985	0,133
X4	1	7	5,943	6	1,232	-1,052	0,525
X5	1	3	1,867	2	0,580	0,015	-0,135
X6	1	4	3,005	3	0,761	-0,271	-0,554
X7	1	7	5,838	6	1,516	-1,485	1,835
X8	1	7	5,519	6	1,664	-1,057	0,358
X9	1	7	4,262	4	2,251	-0,186	-1,417
X10	1	7	5,424	6	1,956	-1,070	-0,157
X11	1	7	5,057	6	2,129	-0,781	-0,811
X12	1	3	2,167	2	0,660	-0,194	-0,731
X13	1	3	2,110	2	0,636	-0,095	-0,539

*:coeficiente de asimetría

- Se debe tomar muy en cuenta al valor de la esperanza matemática de cada una de las variables, esto debido a que dentro del modelo seleccionado como adecuado después de cada análisis factorial (exploratorio y confirmatorio) realizado, las medias son los interceptos en cada una de las regresiones consideradas en el modelo de medida.
- Existe concentración de observaciones en los valores altos de cada una de las variables, a excepción de la variable X_5 , estas afirmaciones se encuentran sustentadas con el valor del coeficiente de asimetría de cada variable, la variable X_5 , es la única que tiene una asimetría positivo. Este tipo de comportamiento en estas variables es coherente, debido a que los clientes tienden a dar calificaciones altas a los atributos evaluados.
- Las desviaciones típicas muestran que la variabilidad es reducida, la variable X_9 es la que tiene más alta variabilidad (desviación típica=2,251) y la variable X_5 es la de menor variabilidad (desviación típica=0,580).

Una vez realizado el análisis descriptivo, se puede continuar con la modelización considerando la estimación de máxima verosimilitud.

3) Modelización inicial

La estrategia de modelización consiste en tomar como punto de partida al modelo representado en la figura 4.3, y mediante la evaluación de los resultados determinar un modelo final, el cual será el punto de partida para la simulación. El modelo obtenido finalmente difiere en las relaciones planteadas en el modelo inicial. La representación gráfica del modelo final obtenido es el siguiente.

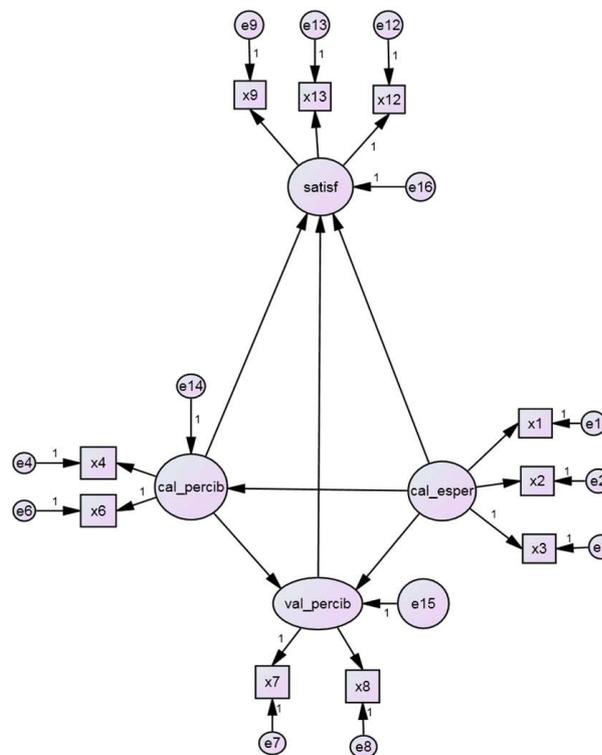


Figura 4.4: Representación gráfica de resultado obtenido en la aplicación del AFE

Se puede notar que las variables eliminadas son: *X5*, *X10*, *X11*, esto debido a su bajo aporte a cada variable inobservable asociada, además inicialmente la variable *X9* se encontraba aportando a la variable *val_percib*, con los datos considerados en la modelización.

zación esta variable pasó a aportar información a la variable *satisfacción*. Finalmente después de un cambio de nombres⁷ en las variables el modelo obtenido es el siguiente:

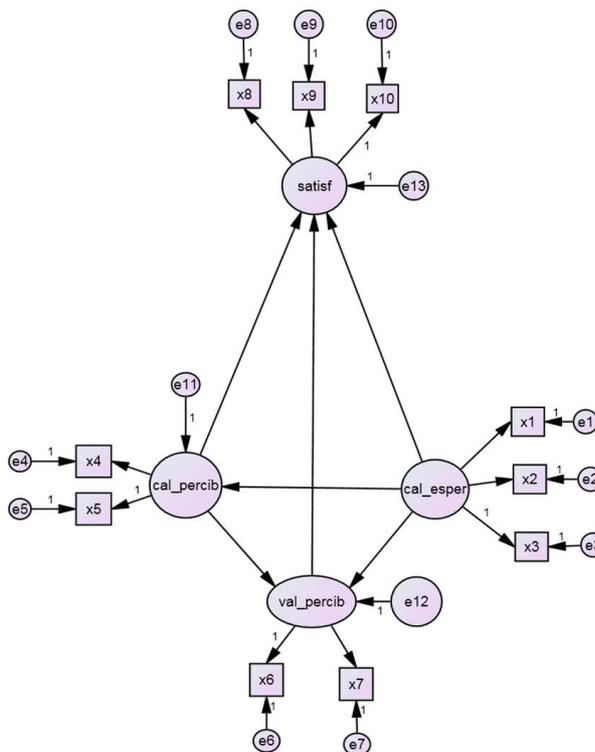


Figura 4.5: Representación gráfica del modelo resultante después de realizar un cambio de nombres en las variables

Del modelo obtenido se tiene: 10 indicadores endógenos respecto a las variables inobservables, además, se debe estimar: 12 regresores (los regresores se encuentran representados por las flechas de una sola punta) (6 del modelo de medida y 6 del modelo estructural), 14 varianzas (13 varianzas correspondientes a los términos de error y 1 varianza correspondiente a la variable *cal_esper*) aplicando la ecuación 3.1 se obtiene el número de grados de libertad

$$gl = \frac{1}{2} [(10 + 0) (10 + 0 + 1)] - (6 + 6 + 14) = 29$$

⁷En el anexo 4, se puede encontrar el cambio de nombres realizado para que la estructura del modelo no cause confusión.

Como se había indicado en el capítulo 3, los grados de libertad del modelo obtenido son mayores a 0 ($gl = 29$), entonces se puede decir que dicho modelo es sobreidentificado, y permite una adecuada estimación de sus parámetros.

Cuadro 4.2: Estimaciones de máxima verosimilitud para el Análisis Factorial Exploratorio

			Estimación	Error estándar	p-valor
<i>cal_percib</i>	< -	<i>cal_esper</i>	0,477	0,120	***
<i>val_percib</i>	< -	<i>cal_esper</i>	1,002	0,387	0,010
<i>val_percib</i>	< -	<i>cal_percib</i>	0,662	0,589	0,261
<i>satisfac</i>	< -	<i>cal_percib</i>	1,133	0,681	0,096
<i>satisfac</i>	< -	<i>cal_esper</i>	0,266	0,427	0,534
<i>satisfac</i>	< -	<i>val_percib</i>	-0,406	0,376	0,280
X1	< -	<i>cal_esper</i>	1,346	0,274	***
X2	< -	<i>cal_esper</i>	1,368	0,281	***
X3	< -	<i>cal_esper</i>	1		
X4	< -	<i>cal_percib</i>	1,320	0,281	***
X5	< -	<i>cal_percib</i>	1		
X6	< -	<i>val_percib</i>	1		
X7	< -	<i>val_percib</i>	0,906	0,154	***
X8	< -	<i>satisfac</i>	0,591	0,350	0,092
X9	< -	<i>satisfac</i>	0,740	0,172	***
X10	< -	<i>satisfac</i>	1		

***: estimación significativa al 0,01

Se puede notar que existen relaciones que tienen el valor de 1 en su estimación, esto se debe que en la modelización de ecuaciones estructurales se exige fijar indicadores para dar una escala de medida a la variable inobservable y además permitir la estimación del modelo (debido a los grados de libertad). Existen 7 estimaciones que resultan ser significativas al 0,05, esto indica que se rechaza la igualdad a cero de cada estimación, el error estándar muestra que la variabilidad obtenida en las estimaciones es baja, además existen relaciones que tienen estimaciones no significativas ($p - valor > 0,100$), a estas relaciones se las pondrá énfasis al aplicar la estimación bayesiana.

La medida fundamental considerada en este análisis es el estadístico de verosimilitud χ^2 , en el cuál se mide el ajuste global que tiene el modelo planteado, el objetivo es obtener valores bajos para este estadístico, ya que este valor mide la distancia entre la matriz de varianza-covarianza estimada y la observada, en el caso del modelo planteado se tiene:

$\chi^2 = 60,739$ con 29 grados de libertad, esto nos da un nivel de significancia 0,000499372, este es el valor más bajo obtenido (en comparación a otros modelos planteados), mediante la aplicación de la estimación de máxima verosimilitud no se puede asegurar que las matrices de varianza-covarianza estimada y observada sean estadísticamente iguales, por tanto, con la estimación bayesiana se busca mejorar este ajuste.

4) Simulación (Estimación Bayesiana)

El proceso de estimación bayesiana toma como punto de partida a las estimaciones de máxima verosimilitud, además se debe fijar las distribuciones a priori que se usará el algoritmo para obtener las estimaciones. Debido a que no se tiene información relacionada o estudios similares al planteado en este trabajo, se usará la distribución uniforme ($U [-3 \cdot 10^{40}, 3 \cdot 10^{40}]$) como distribución a priori para todas las relaciones a estimar, el rango escogido para las distribución uniformes es el más grande posible (debido al software utilizado), esto es debido a que se quiere usar una distribución no informativa para que “los datos hablen por si mismos” y las estimaciones obtenidas sean producto de la evaluaciones de los compradores en general. Un ejemplo de las distribuciones consideradas se puede observar en el siguiente gráfico.

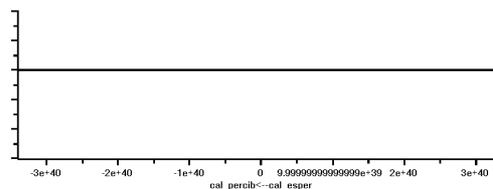


Figura 4.6: Ejemplo de distribuciones a priori utilizadas en la estimación

Este tipo de distribución deja libertad a que el valor de la estimación se encuentre entre el intervalo $[-3 \cdot 10^{40}, 3 \cdot 10^{40}]$. Una vez establecidas las distribuciones a priori y el punto inicial, el cual está dado por la estimación de máxima verosimilitud, puede iniciar la simulación para obtener la estimación bayesiana. El proceso de simulación se encuentra dividido en las siguientes etapas:

- a. Simulación inicial: simular el proceso considerando un número de iteraciones adecuado (debería ser al menos 100 iteraciones), esto permitirá obtener las primeras estimaciones de las relaciones y, principalmente, calcular la autocorrelación en la serie de iteraciones con el fin de determinar el periodo de calentamiento.
- b. Calcular el periodo de calentamiento: mediante la representación gráfica de la autocorrelación, se busca identificar el periodo de calentamiento, es decir, el número de iteraciones que serán desechadas con la finalidad de eliminar influencias entre estimaciones, el gráfico de la autocorrelación considera las 100 primeras estimaciones, como se notará, no necesariamente la autocorrelación disminuirá con una velocidad lineal (por lo general, disminuye a una velocidad exponencial), pero por practicidad se asumirá velocidad lineal para calcular el número de iteraciones en las cuales la autocorrelación disminuye a 0.
- c. Simulación final: simular el proceso hasta obtener la mayor convergencia posible en cada una de las estimaciones

Desarrollo

- a. Simulación inicial: se han considerado 10000 iteraciones con el fin de obtener los gráficos de las autocorrelaciones; se ha electo tal número de iteraciones debido a que en esta primera etapa no se conoce cual debería ser el periodo de calentamiento, de ser el caso, si se hubiera electo pocas iteraciones a simular y el número de iteraciones a eliminar es mayor, entonces, se debe simular nuevamente un número de iteraciones mayor al período de calentamiento y ahí si eliminarlas, de ninguna forma se debe eliminar menos iteraciones a las consideradas en el periodo de calentamiento, ya que, esto implica que ingrese información correlacionada en las estimaciones afectando los resultados finales. Hay que mencionar que en esta etapa no se ha considerado un criterio de convergencia debido a que se desconoce el grado de autocorrelación de las estimaciones.
- b. Calcular el periodo de calentamiento: una vez simulado el proceso, se obtienen los

gráficos de las autocorrelaciones, el uso de estos diagramas permite identificar el número mínimo de iteraciones para el cual las estimaciones del proceso simulado pierden autocorrelación. Los autocorrelogramas obtenidos son:

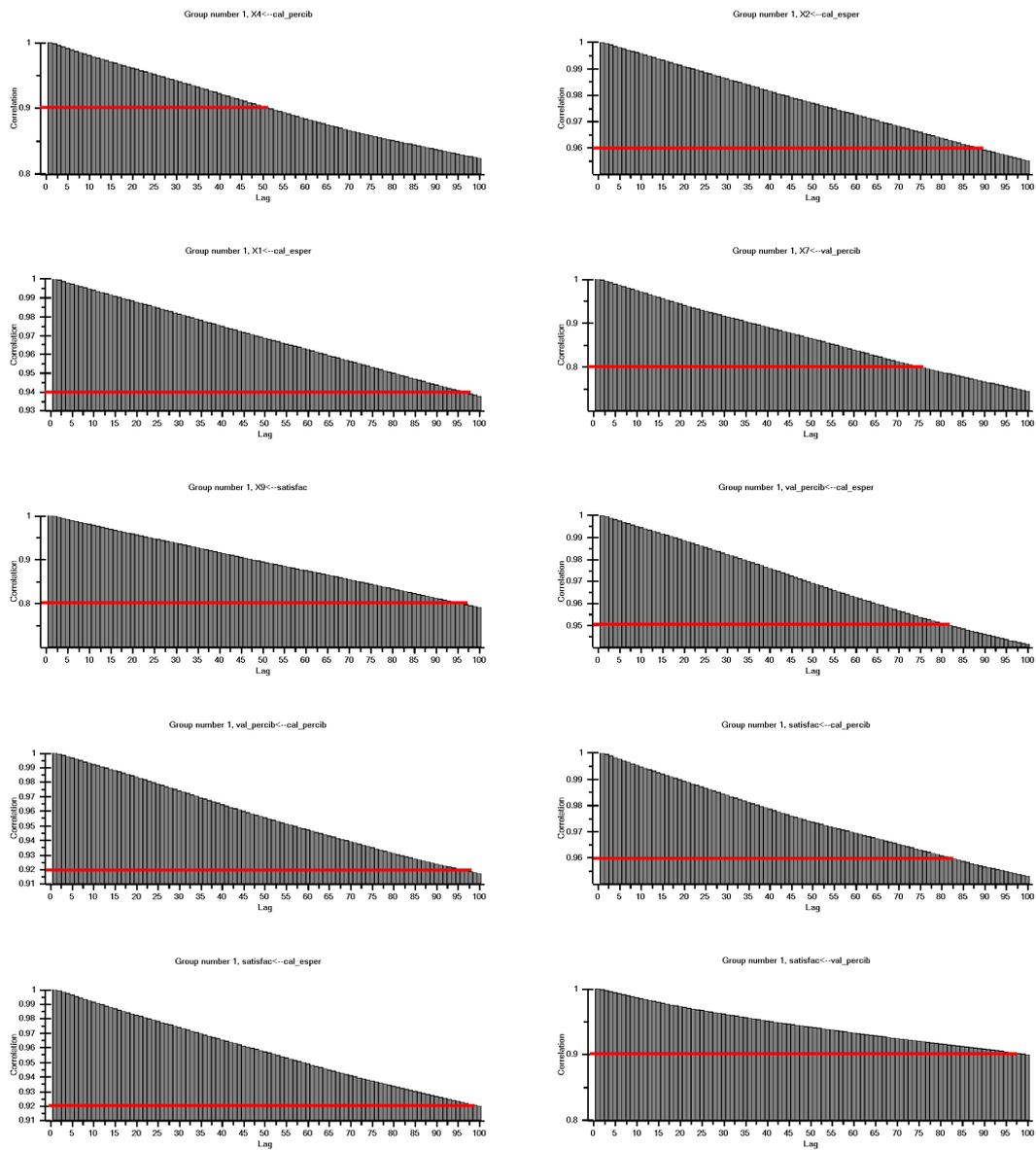


Figura 4.7: Autocorrelogramas de las estimaciones bayesianas del Análisis Factorial Exploratorio considerando 10000 iteraciones

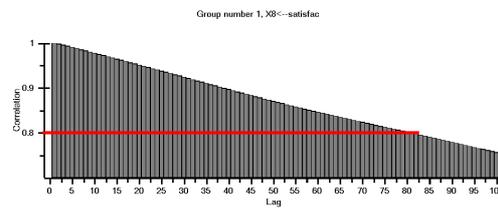


Figura 4.8: Autocorrelogramas de las estimaciones bayesianas del Análisis Factorial Exploratorio considerando 10000 iteraciones

En el siguiente cuadro se marcan los puntos considerados en cada uno de los autocorrelogramas y en la cuarta columna se calcula el número de iteraciones necesarias para eliminar la autocorrelación.

Cuadro 4.3: Cálculo del período de calentamiento

Relación	Rezago (A)	Autocorrelación (B)	Iteraciones para reducir autocorrelación (A/(1-B))
$X4 < -cal_percib$	50	0,90	500
$X2 < -cal_esper$	90	0,96	2250
$X1 < -cal_esper$	98	0,94	1634
$X7 < -val_percib$	77	0,80	385
$X9 < -satisfac$	97	0,80	485
$val_percib < -cal_esper$	82	0,95	1640
$val_percib < -cal_percib$	98	0,92	1225
$satisfac < -cal_percib$	83	0,96	2075
$satisfac < -cal_esper$	99	0,92	1238
$satisfac < -val_percib$	93	0,90	930
$X8 < -satisfac$	83	0,80	415

Como se puede observar, la relación $X2 < -cal_esper$, es la que mayor número de iteraciones necesita (2250) para eliminar la autocorrelación en su estimación, para asegurarnos que la autocorrelación se elimine de las estimaciones totalmente se escoge como período de calentamiento 2500 iteraciones, las iteraciones siguientes a este número se tomarán como válidas e ingresarán al cálculo de la estimación bayesiana.

- c. Simulación final: en esta etapa la simulación busca el número de iteraciones hasta que el criterio de convergencia haya sido satisfecho. El proceso de simulación se realizó

en diferentes etapas (40000, 65000, 80000, 100000 y 1664000 iteraciones) para poder observar la convergencia en diferentes puntos y como finalidad se debe realizar la simulación hasta identificar el número de iteraciones para el cual el estadístico de convergencia es lo más cercano a 1, y con ello, aceptar la validez de las estimaciones obtenidas, en este caso, la convergencia más cercana a 1 es 1,004 en la relación $X2 < -cal_esper$ al realizar 1664000 iteraciones.

Cuadro 4.4: Estimaciones bayesiana y estadísticos de convergencia para el Análisis Factorial Exploratorio

Relación	Número de iteraciones simuladas									
	40000+2500		65000+2500		80000+2500		100000+2500		1664000+2500	
	Media	E.C.	Media	E.C.	Media	E.C.	Media	E.C.	Media	E.C.
$X4 < -cal_percib$	1,459	1,010	1,405	1,007	1,419	1,008	1,390	1,004	1,390	1,000
$X2 < -cal_esper$	2,135	1,022	1,956	1,017	1,891	1,018	1,897	1,015	1,798	1,004
$X1 < -cal_esper$	1,917	1,022	1,772	1,018	1,725	1,018	1,757	1,015	1,696	1,003
$X7 < -val_percib$	0,863	1,010	0,897	1,006	0,899	1,005	0,896	1,005	0,897	1,000
$X9 < -satisfac$	0,739	1,015	0,741	1,008	0,724	1,010	0,727	1,005	0,728	1,000
$cal_percib < -cal_esper$	0,707	1,018	0,650	1,015	0,636	1,015	0,637	1,012	0,643	1,003
$val_percib < -cal_esper$	1,603	1,015	1,402	1,014	1,425	1,011	1,462	1,011	1,305	1,002
$val_percib < -cal_percib$	0,545	1,018	0,589	1,013	0,491	1,010	0,490	1,010	0,630	1,002
$satisfac < -cal_percib$	1,843	1,022	1,674	1,018	1,691	1,016	1,608	1,014	1,873	1,004
$satisfac < -cal_esper$	-0,091	1,021	-0,011	1,018	-0,045	1,016	0,018	1,014	-0,001	1,004
$satisfac < -val_percib$	-0,445	1,014	-0,432	1,008	-0,415	1,007	-0,412	1,007	-0,496	1,003
$X8 < -satisfac$	0,623	1,008	0,591	1,004	0,600	1,003	0,605	1,002	0,565	1,000

E.C. : estadístico de convergencia

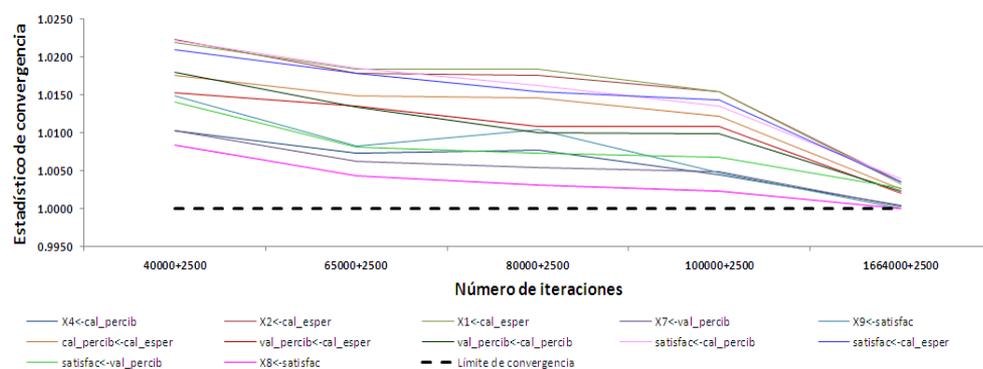


Figura 4.9: Convergencia de las estimaciones en el Análisis Factorial Exploratorio

Con la representación gráfica de los estadísticos de convergencia de cada una de las relaciones en las etapas realizadas, se puede notar que a medida que avanzan las iteraciones las estimaciones convergen aproximándose a 1, en 1664000 iteraciones, además, el estadístico de convergencia alcanza su mínimo (1,004) y si se considera un número mayor de iteraciones la convergencia disminuirá levemente. Para continuar con el análisis se realiza una comparación entre las 100000 primeras iteraciones y las 50000 últimas iteraciones.

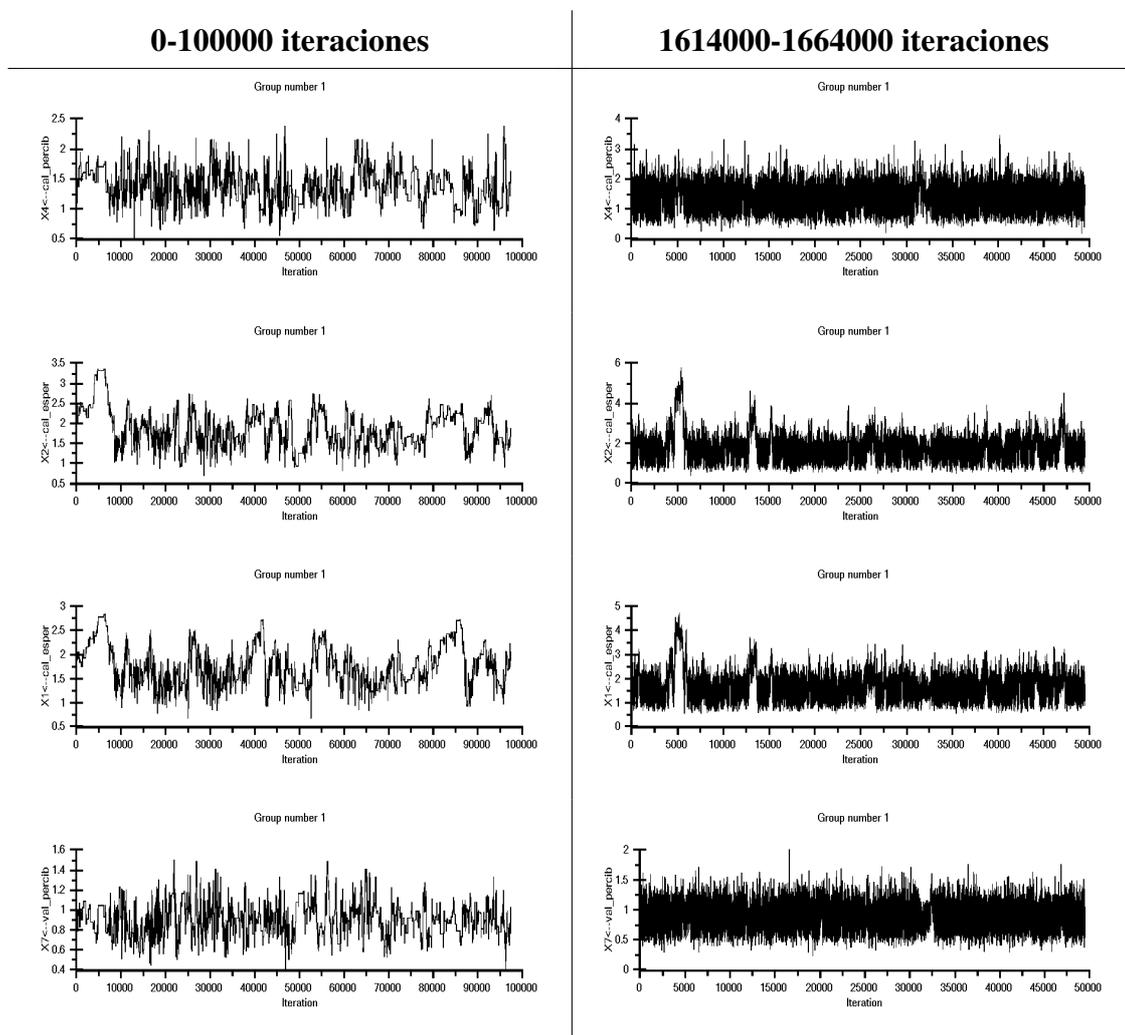


Figura 4.10: Comparación de series de iteraciones de estimaciones bayesianas en el análisis factorial exploratorio

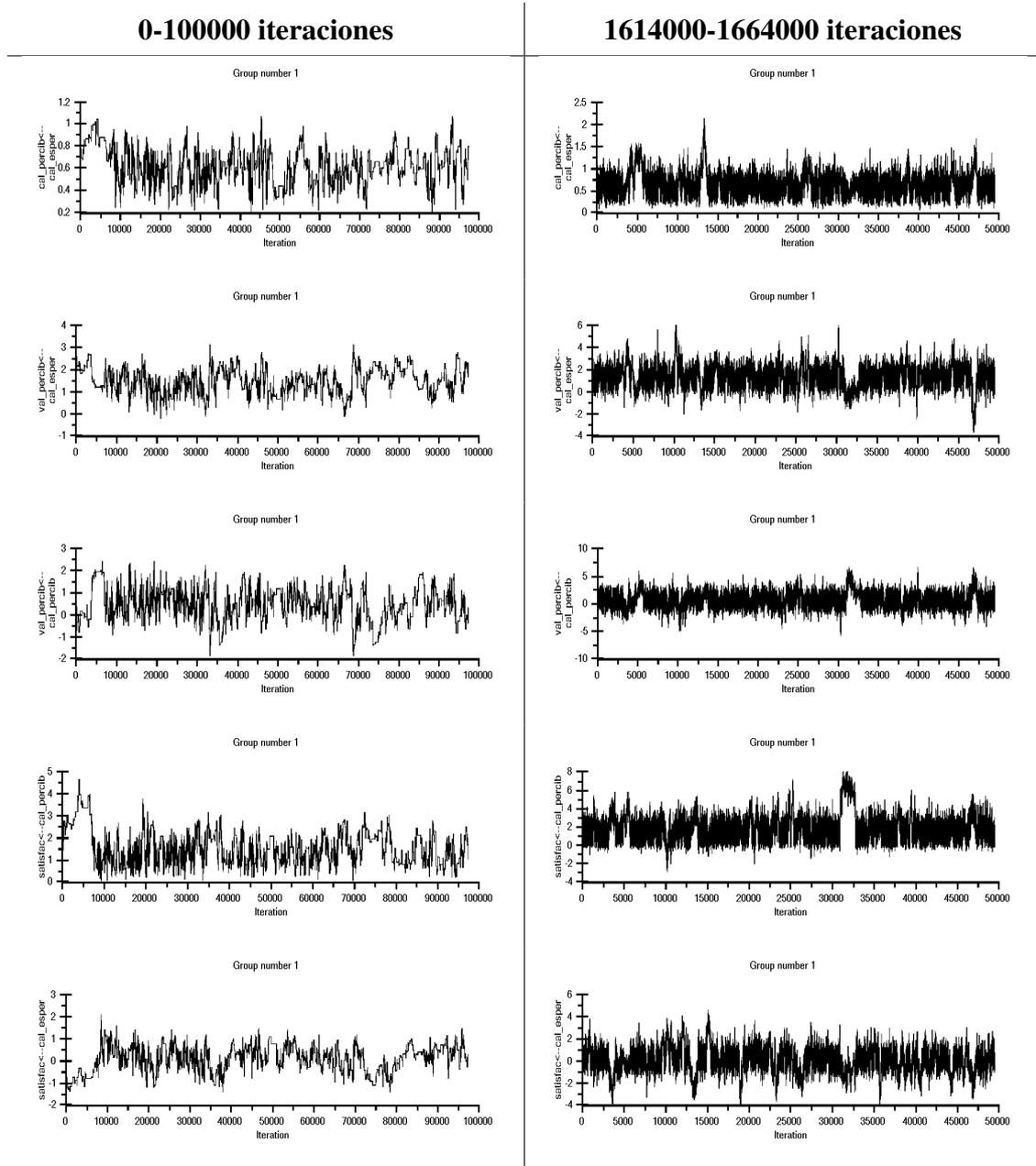


Figura 4.11: Comparación de series de iteraciones de estimaciones bayesianas en el análisis factorial exploratorio

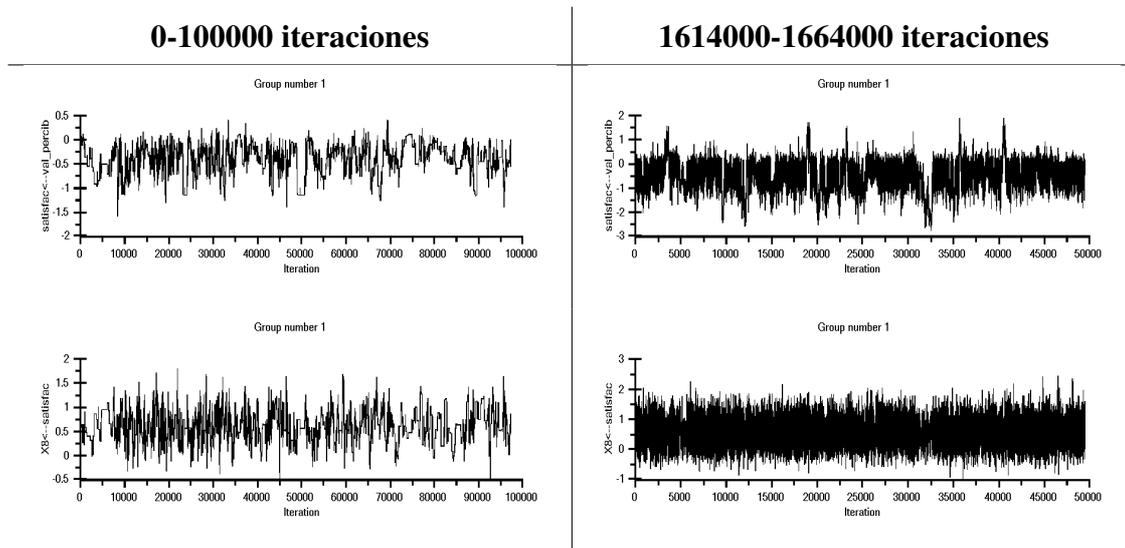


Figura 4.12: Comparación de series de iteraciones de estimaciones bayesianas en el análisis factorial exploratorio (continuación)

Para realizar esta comparación, ya se ha eliminado 2500 iteraciones, considerando las 50000 primeras iteraciones y las 50000 últimas iteraciones, se puede observar en los gráficos de la izquierda, las primera iteraciones tienen mayor variabilidad en comparación a las del lado derecho, además, en las últimas 50000 iteraciones se evidencia mayor convergencia de la estimación, en cambio, las primeras 50000 iteraciones muestran que el algoritmo no encuentra estimaciones estables (es decir, con poco variabilidad), las primeras 100000 iteraciones consideradas ya muestran valores cercanos a 1 en sus estadísticos de convegnencia ($EC_{X7 < -val_percib} = 1,005$, $EC_{X8 < -satisfac} = 1,003$), esto debido a que existen algunas estimaciones bayesianas que convergen con mayor rapidez, y otras como $X1 < -cal_percib$ convergen con lentitud, es por esto que en 1664000 iteraciones recién se alcanza la convergencia máxima en todas las estimaciones bayesianas (1,004). En el anexo 4, se encuentra la comparación de los histogramas de las primeras 100000 iteraciones y las últimas 50000 iteraciones. Para terminar esta etapa del análisis, se enuncia los resultados obtenidos para el análisis factorial exploratorio simulado 1664000 veces.

Cuadro 4.5: Resultados (AFE)

Relación	Media	E.E	D.E.	E.C.	95 % L.I.	95 % L.S.	Sesgo	Curtosis
$X4 < -cal_percib$	1,390	0,010	0,330	1,000	0,831	2,127	0,694	1,242
$X2 < -cal_esper$	1,798	0,052	0,608	1,004	1,068	3,464	2,066	6,179
$X1 < -cal_esper$	1,696	0,042	0,519	1,003	1,036	3,033	1,763	4,684
$X7 < -val_percib$	0,897	0,004	0,165	1,000	0,603	1,259	0,461	0,682
$X9 < -satisfac$	0,728	0,003	0,178	1,000	0,457	1,144	0,830	1,071
$cal_percib < -cal_esper$	0,643	0,016	0,224	1,003	0,335	1,228	1,412	3,195
$val_percib < -cal_esper$	1,305	0,060	0,922	1,002	-0,514	3,223	-0,036	2,951
$val_percib < -cal_percib$	0,630	0,084	1,213	1,002	-1,722	3,593	0,330	2,061
$satisfac < -cal_percib$	1,873	0,118	1,320	1,004	0,174	5,788	1,365	2,579
$satisfac < -cal_esper$	-0,001	0,090	1,075	1,004	-2,329	2,278	-0,022	0,799
$satisfac < -val_percib$	-0,495	0,041	0,555	1,003	-1,902	0,558	-0,582	1,952
$X8 < -satisfac$	0,565	0,005	0,346	1,000	-0,088	1,282	0,217	0,312

E.E. : error estándar

D.E. : desviación estándar

E.C. : estadístico de convergencia

L.I. : límite inferior

L.S. : límite superior

Lo que primero llama la atención, es la estimación de la relación $satisfac < -cal_esper$, la cual tiene un valor cercano a 0 (-0,001), estadísticamente la calidad esperada no influye en la satisfacción, al contrario de lo que sucede en la relación $cal_percib < -cal_esper$, en donde, la calidad que el comprador espera es estadísticamente significativa sobre la calidad que el percibe. Los intervalos de confianza al 95 % indican que las relaciones: $val_percib < -cal_esper$, $val_percib < -cal_percib$, $satisfac < -cal_esper$, $satisfac < -val_percib$, $X8 < -satisfac$; no estadísticamente distintas de cero, hay que recordar que este modelo obtenido, es el mejor posible, y en el análisis factorial confirmatorio se mejorará este modelo con correlaciones y fijando distribuciones a posteriori informativas.

5) Conclusiones del análisis factorial exploratorio:

Esta es la sección final del análisis factorial exploratorio, de todos los resultados obtenidos se extraen conclusiones sobre la validez del análisis realizado, y se terminará con el

análisis factorial exploratorio estableciendo un modelo final que será el punto de partida para el análisis factorial confirmatorio:

- La simulación ha considerado 1664000 iteraciones, en las cuales se ha podido obtener estimaciones bayesianas estables y fiables, un modelo con una convergencia global de 1,004 (E.C.).
- Se estableció un periodo de calentamiento de 2500 iteraciones para eliminar las observaciones de cada una de las estimaciones bayesianas para eliminar en su totalidad la autocorrelación, esto ayuda a garantizar que los resultados obtenidos no tengan influencia de la inicialización del proceso de simulación.
- El proceso de simulación se ha realizado en diferentes etapas para permitir el análisis del comportamiento de las estimaciones bayesianas obtenidas, se puede notar que a medida que aumentan el número de iteraciones la convergencia también lo hace, alcanzando su máxima convergencia en 1664000 iteraciones.
- Mediante la representación gráfica de la figura 4.12 se puede notar que los estadísticos de convergencia de cada una de las estimaciones decae en el último periodo de simulación, además se puede notar que la convergencia en la relación $satisfac < -cal_esper$ disminuye (E.C. = 1,021) en las 40000 iteraciones, para luego decaer E.C. = 1,004 en las iteraciones finales. Este comportamiento es normal debido a que las estimaciones bayesianas son obtenidas mediante simulación.
- En las primeras 100000 iteraciones, los gráficos de las estimaciones obtenidas en función de las iteraciones realizadas indican que existe baja convergencia, que la variabilidad se encuentra presente en todas las estimaciones bayesianas y sería demasiado prematuro y erróneo tomar los valores de las estimaciones bayesianas como válidos, con el fin de eliminar al máximo la variabilidad de las estimaciones bayesianas se simuló el proceso durante 1664000 iteraciones, los resultados son evidentes al comparar las estimaciones obtenidas inicialmente (100000 iteraciones) y las últimas estimaciones obtenidas; la variabilidad disminuye, es decir, las estimaciones bayesianas obtenidas en las últimas iteraciones varían muy poco entre sí, es por esta

razón que su estadístico de convergencia global es 1,004, además se puede notar que existen estimaciones bayesianas con mayor convergencia que otras, por ejemplo, la relación $X4 < -cal_percib$ alcanza una convergencia de 1,00001, indicando que la estimación obtenida es muy buena, y que los valores simulados varían en el orden de las cien milésimas.

- Se puede decir que las estimaciones bayesianas obtenidas son buenas (en función de su estadístico de convergencia), esto no es todo, hay que recordar que el fin del análisis factorial exploratorio es obtener el menor número de indicadores que puedan representar el fenómeno, en este caso, la satisfacción del cliente y cada uno de sus componentes (siguiendo la metodología ACSI), para ello, se puede observar en el cuadro 4.5, en los límites existen relaciones para las cuales el valor de 0 se encuentra contenido en sus intervalos, esto indica que estadísticamente no están influyendo la una sobre la otra, esto puede corregirse al eliminar alguna de las relaciones, el caso más notorio es para la relación $satisf < -cal_esper$, para el cual la estimación bayesiana es cercana a 0, esta relación no se tomará en cuenta para el modelo final.

Después de haber analizado los resultados obtenidos, y de considerar la eliminación de la relación $satisf < -cal_esper$ debido a su aporte nulo (estadísticamente igual a cero), además de la eliminación de la variable $X8$ y de sus componentes (término de error $e8$, relación $satisf < -X8$), se ha establecido el modelo resultante del análisis factorial exploratorio, este modelo es el punto de partida para la realización del análisis factorial confirmatorio.

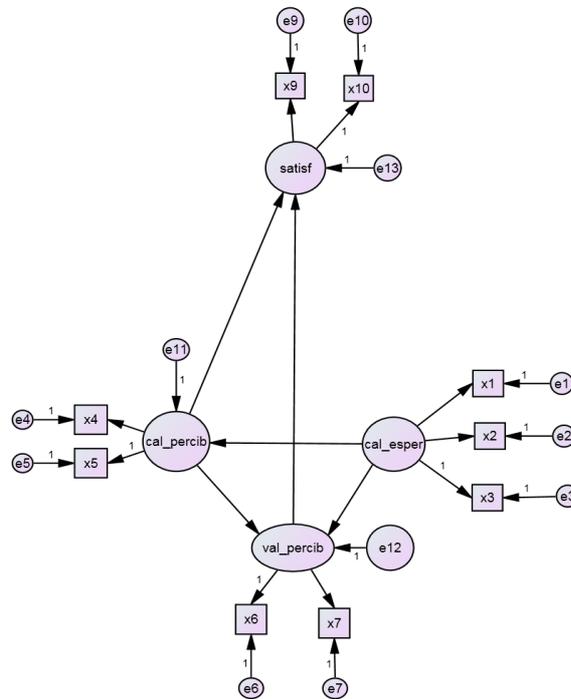


Figura 4.13: Modelo resultante después de haber realizado el análisis factorial exploratorio utilizando estimación bayesiana

ANÁLISIS FACTORIAL CONFIRMATORIO

Una vez realizado el Análisis Factorial Exploratorio (AFE), se han podido obtener el conjunto de indicadores que permitirán la realización del Análisis Factorial Confirmatorio (AFC), el objetivo principal es establecer un modelo de ecuaciones estructurales que permita la estimación de la satisfacción del cliente de un determinado producto; para la consecución del objetivo, en este tipo de análisis se puede realizar la representación las covarianzas entre variables, esta representación es de gran utilidad debido a que en la práctica, no ha sido posible obtener variables totalmente independientes, además, esta herramienta permite dar mayor representatividad de la realidad al modelo. Para la realización del AFC se fija un conjunto de pasos a seguir, los pasos de la modelización constan de:

- 1) **Punto de partida Análisis Factorial Confirmatorio:** se utilizará el modelo resultante del AFE y mediante el método de máxima verisimilitud se obtienen estimaciones para las relaciones planteadas, la información obtenida será el punto de partida para la simulación.
- 2) **Simulación del modelo:** una vez obtenidos los resultados de las estimaciones de máxima verisimilitud, se deben establecer las distribuciones a priori, esto como un elemento adicional para la estimación bayesiana. En esta etapa se están representadas las covarianzas en el modelo; las covarianzas serán calculadas en función de los indicadores y estos resultados entrarán a la simulación para obtener una estimación bayesiana de las covarianzas.
- 3) **Análisis de los resultados:** una vez realizada la simulación, se debe analizar los resultados obtenidos, para concluir sobre su validez, hay que recordar que esta es la etapa final de estimación de cargas factoriales y de esta dependerá la estimación de las variables inobservables, en especial, de la satisfacción del cliente.

Desarrollo del análisis factorial confirmatorio:

1) Punto de partida Análisis Factorial Confirmatorio

El modelo considerado, es el resultante del AFE (figura 4.17), para establecer el punto de partida de la simulación es necesario obtener las estimaciones de máxima verosimilitud de cada una de las relaciones, debido a la naturaleza del AFC es necesario realizar previamente mostrar las matrices de varianzas y covarianzas, y de correlaciones entre indicadores. Se puede notar que las varianzas son bajas, es decir, la variabilidad en los indicadores es reducida, por cuanto los consumidores tienden a tener un mismo criterio para la evaluación del producto (en base al indicador que se desea analizar), el indicador X_9 tiene la menor varianza ($Var(X_9) = 0,402$) y el indicador X_7 tiene la mayor varianza ($Var(X_7) = 2,754$), la información proporcionada por la matriz de varianzas-covarianzas indica que existe homogeneidad en las respuestas de los consumidores.

Para describir las covarianzas no se toma en cuenta si existe algún tipo de relación entre indicadores que buscan explicar una misma variable inobservable, ya que, si se incorpora

Cuadro 4.6: Matriz de varianzas-covarianzas

VARIABLES INOBSERVABLES	INDICADORES	X1	X2	X3	X4	X5	X6	X7	X9	X10
Calidad esperada	X1	1,625								
	X2	0,781	1,871							
	X3	0,724	0,608	2,621						
Calidad percibida	X4	0,191	0,629	0,120	1,511					
	X5	0,271	0,330	0,212	0,257	0,576				
Valor percibido	X6	0,943	0,761	0,407	0,438	0,282	2,288			
	X7	0,625	0,710	0,533	0,482	0,393	0,960	2,754		
Satisfacción	X9	0,078	0,145	0,229	0,120	0,104	0,008	0,110	0,402	
	X10	0,162	0,198	0,038	0,195	0,137	0,160	0,113	0,210	0,434

Cuadro 4.7: Matriz de correlaciones

VARIABLES INOBSERVABLES	INDICADORES	X1	X2	X3	X4	X5	X6	X7	X9	X10
Calidad esperada	X1	1								
	X2	0,448	1							
	X3	0,351	0,275	1						
Calidad percibida	X4	0,122	0,373	0,060	1					
	X5	0,281	0,318	0,172	0,276	1				
Valor percibido	X6	0,489	0,368	0,166	0,236	0,245	1			
	X7	0,295	0,312	0,198	0,236	0,312	0,382	1		
Satisfacción	X9	0,096	0,168	0,223	0,155	0,216	0,008	0,104	1	
	X10	0,193	0,219	0,036	0,241	0,274	0,161	0,104	0,503	1

la relación entre un par de indicadores de una variable latente se estará asumiendo que existe multicolinealidad, de esta forma, se introducirá información errada a la estimación del modelo, el análisis de las covarianzas se centra entre indicadores de distintas variables inobservables. De la observación la matriz de varianzas-covarianzas se puede concluir:

- Considerando las variables *calidad percibida* y *calidad esperada*, la mayor covarianza se encuentra entre los indicadores X2 y X4, se debe representar esta relación en el modelo.
- Analizando las variables *valor percibido* y *calidad esperada*, los indicadores X6 y X1 presentan la mayor covarianza, también se la debe tomar en cuenta.
- Para las variables *satisfacción* y *calidad esperada*, se encuentran los indicadores X9 y X3 con la mayor covarianza, este relación se debe incorporar al modelo.

De la matriz de correlación se puede concluir:

- Los indicadores se encuentran mínimamente correlacionados linealmente, debido a esto, la correlación máxima está dado por los indicadores $X9$ y $X10$ ($corr(X9, X10) = 0,503$), esta correlación no presenta ningún problema al momento de realizar la simulación, por esta razón las cargas factoriales de cada uno de estos indicadores es significativo tanto en las estimaciones de máxima verosimilitud como en las bayesianas.
- Existen pares de indicadores para los cuales la correlación es cercana a cero, por ejemplo $X3$ y $X4$ ($corr = 0,060$), $X6$ y $X9$ ($corr = 0,008$), indicando que existen atributos del producto que no se encuentran relacionados entre si, es decir, si el productor desea dar impulso (por ejemplo, promociones, campañas de publicidad, etc.) a estos factores, no estará afectando a los atributos con los que no están correlacionados.
- Se debe notar que los valores de las correlaciones son todos mayores a cero, esto se debe al trabajo previo realizado con la base de datos, para que, a medida que aumente la calificación de un atributo también aumente la satisfacción del cliente.

Para una mejor comprensión, se representa gráficamente el modelo de ecuaciones estructurales usado para el análisis factorial confirmatorio.

Cuadro 4.8: Ajuste global del modelo para combinaciones de covarianzas representadas

Covarianzas representadas	X1 - X6	X9 - X3	X2 - X4	X1 - X6	X1 - X6	X2 - X4	X1 - X6
				X2 - X4	X9 - X3	X9 - X3	X2 - X4
gl	38,3	37,7	38	28,4	27,2	25,3	17,3
χ^2_{gl}	21	21	21	20	20	20	19
<i>p</i> - valor	0,012	0,014	0,013	0,100	0,130	0,190	0,570
<i>NFI</i>	0,899	0,901	0,900	0,925	0,928	0,933	0,954

hipótesis nula de igualdad de matrices de varianza covarianza estimada y observada:

$$H_0 : S = \Sigma$$

$$H_1 : S \neq \Sigma$$

Además el índice de bondad de ajuste normado muestra una tendencia creciente en cuanto al número de relaciones de covarianza representadas, afirmando que el modelo con tres covarianzas representadas es el mejor para continuar con el análisis factorial confirmatorio.

Una vez determinado el modelo a utilizarse en el análisis factorial confirmatorio, es necesario obtener las estimaciones de máxima verosimilitud, tanto para las relaciones entre variables como para las covarianzas.

Cuadro 4.9: Estimaciones de máxima verosimilitud para el Análisis Factorial Confirmatorio

			Estimación	Error estándar	p-valor
<i>cal_percib</i>	< -	<i>cal_esper</i>	0,428	0,102	***
<i>val_percib</i>	< -	<i>cal_percib</i>	1,144	0,600	0,057
<i>val_percib</i>	< -	<i>cal_esper</i>	0,466	0,292	0,110
<i>satisf</i>	< -	<i>cal_percib</i>	1,305	0,837	0,119
<i>satisf</i>	< -	<i>val_percib</i>	-0,347	0,349	0,320
X1	< -	<i>cal_esper</i>	1,199	0,241	***
X2	< -	<i>cal_esper</i>	1,298	0,260	***
X3	< -	<i>cal_esper</i>	1		
X4	< -	<i>cal_percib</i>	1,188	0,270	***
X5	< -	<i>cal_percib</i>	1		
X6	< -	<i>val_percib</i>	1		
X7	< -	<i>val_percib</i>	1,099	0,211	***
X9	< -	<i>satisf</i>	0,720	0,177	***
X10	< -	<i>satisf</i>	1		
e1	< - >	e6	0,335	0,119	***
e2	< - >	e4	0,307	0,100	***
e9	< - >	e3	0,189	0,059	***

***: estimación significativa al 0.01

En el cuadro anterior se puede ver que existen algunas relaciones que no son significativas, principalmente las relaciones consideradas en el modelo estructural, es aquí, en donde entra la estimación bayesiana para mejorar la significancia de las relaciones.

2) Simulación del modelo

Con las estimaciones de máxima verosimilitud, y las distribuciones a priori, se puede iniciar el proceso de simulación. Debido a que inicialmente en el análisis factorial exploratorio no se incorporaron distribuciones a priori informativas, en esta etapa del proceso se utilizará las mismas usadas inicialmente, ya que este proceso es confirmatorio, por lo tanto se usará la distribución uniforme ($U [-3 \cdot 10^{40}, 3 \cdot 10^{40}]$) como distribución a priori para todas las relaciones a estimar; un ejemplo de la distribución a priori se encuentra en la figura 4.8. La simulación se la realiza en las siguientes 2 etapas:

- a. Calcular el periodo de calentamiento: el proceso es simulado 10000 iteraciones, con el fin de encontrar el periodo de calentamiento óptimo para eliminar la autocorrelación en las estimaciones; se asumirá velocidad lineal en la disminución de la autocorrelación para calcular el número de iteraciones en las cuales la autocorrelación disminuye a 0.
- b. Simulación final: simular el proceso hasta que converja en cada una de las estimaciones bayesianas, este proceso tomará menos iteraciones en comparación al AFE, esto debido a que ya se ha eliminado información que no aporta significativamente al modelo.

Desarrollo

- a. Cálculo del periodo de calentamiento: se ha considerado 10000 iteraciones como la primera etapa de simulación, con los autocorrelogramas obtenidos se procede al cálculo del periodo de calentamiento. Los autocorrelogramas obtenidos son:

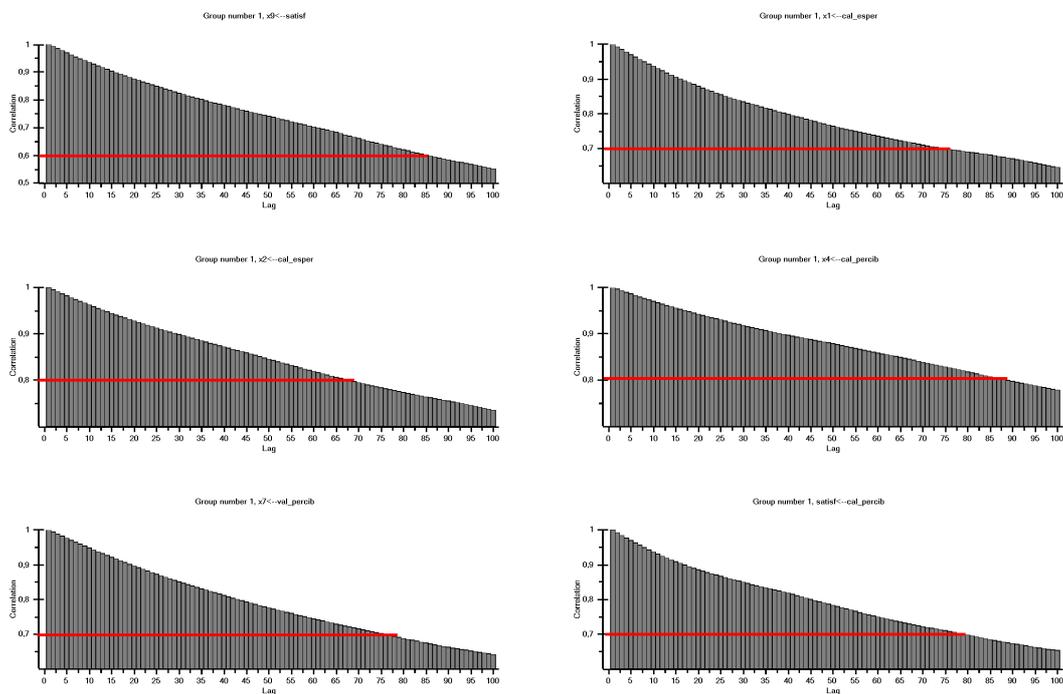


Figura 4.15: Autocorrelogramas de las estimaciones bayesianas del Análisis Factorial Confirmatorio considerando 10000 iteraciones

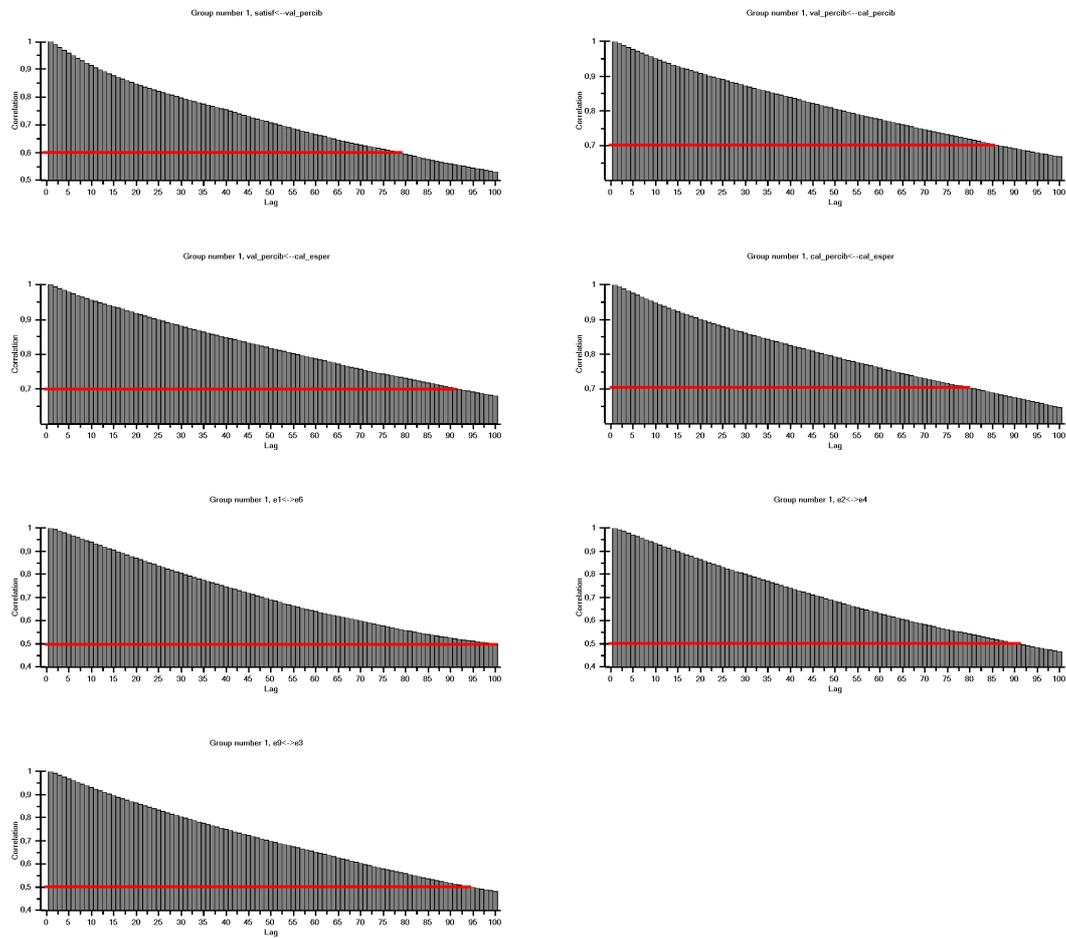


Figura 4.16: Autocorrelogramas de las estimaciones bayesianas del Análisis Factorial Confirmatorio considerando 10000 iteraciones

Con la ayuda del siguiente cuadro, se calcula el número de iteraciones necesarias para eliminar la autocorrelación en cada estimación bayesiana.

Cuadro 4.10: Cálculo del período de calentamiento para el Análisis Factorial Confirmatorio

Relación	Rezago (A)	Autocorrelación (B)	Iteraciones para reducir autocorrelación (A/(1-B))
$X9 < -satisfac$	85	0,6	213
$X1 < -cal_esper$	77	0,7	257
$X2 < -cal_esper$	68	0,8	340
$X4 < -cal_percib$	88	0,8	440
$X7 < -val_percib$	77	0,7	257
$satisf < -cal_percib$	80	0,7	267
$satisf < -val_percib$	80	0,6	200
$val_percib < -cal_percib$	85	0,7	284
$val_percib < -cal_esper$	93	0,7	310
$cal_percib < -cal_esper$	80	0,7	267
$e1 < - > e6$	100	0,5	200
$e2 < - > e4$	93	0,5	186
$e9 < - > e3$	95	0,5	190

Se puede notar que la relación $X4 < -cal_percib$ necesita mayor número de iteraciones (440) para eliminar la autocorrelación, solamente considerando 500 iteraciones como periodo de calentamiento, ya se han eliminado las estimaciones relacionadas; el número de iteraciones necesarias para la eliminación de la autocorrelación en el AFE fue 2500, para el AFC solamente se necesitan 500, esto indica que el modelo considerado para el AFC converge con mayor rapidez por tanto se eliminan menor número de iteraciones. Por cuestiones practicas, se ha elegido 1000 iteraciones como periodo de calentamiento.

- b. Simulación final: la simulación del modelo ha sido desarrollada en varias etapas para permitir la observación de la convergencia, en correspondencia al AFE; las etapas consideradas son: 40000, 65000, 80000, 100000 y 360000 iteraciones. El proceso simulado alcanza su mayor convergencia en 360000 iteraciones con un estadístico de convergencia máximo de 1.0015, esta cantidad de iteraciones es mucho menor a las realizadas en el AFE, en dicho análisis se realizaron 1664000 iteraciones alcanzando un estadístico de convergencia máximo de 1,0040, por lo tanto, en el Análisis Factorial Confirmatorio se necesitan menor número de iteraciones para obtener una mejor

convergencia.

Cuadro 4.11: Estimaciones bayesiana y estadísticos de convergencia para el Análisis Factorial Confirmatorio

Relación	Número de iteraciones simuladas									
	40000+1000		65000+1000		80000+1000		100000+1000		360000+1000	
	Media	E.C.	Media	E.C.	Media	E.C.	Media	E.C.	Media	E.C.
$x_9 < -satisf$	0,671	1,003	0,671	1,002	0,638	1,004	0,623	1,005	0,633	1,0015
$x_1 < -cal_esper$	1,335	1,005	1,390	1,010	1,421	1,005	1,414	1,002	1,418	1,0005
$x_2 < -cal_esper$	1,476	1,007	1,555	1,014	1,518	1,004	1,516	1,002	1,540	1,0009
$x_4 < -cal_percib$	1,210	1,002	1,227	1,002	1,196	1,001	1,193	1,001	1,217	1,0006
$x_7 < -val_percib$	1,079	1,003	1,088	1,002	1,119	1,003	1,106	1,002	1,096	1,0007
$satisf < -cal_percib$	1,358	1,011	1,533	1,005	1,349	1,003	1,362	1,004	1,390	1,0011
$satisf < -val_percib$	-0,331	1,007	-0,401	1,004	-0,336	1,002	-0,357	1,005	-0,361	1,0015
$val_percib < -cal_percib$	1,186	1,006	1,210	1,002	1,138	1,002	1,159	1,001	1,194	1,0009
$val_percib < -cal_esper$	0,558	1,007	0,548	1,003	0,599	1,003	0,581	1,002	0,572	1,0008
$cal_percib < -cal_esper$	0,463	1,007	0,500	1,013	0,491	1,004	0,493	1,001	0,495	1,0009

E.C. : estadístico de convergencia

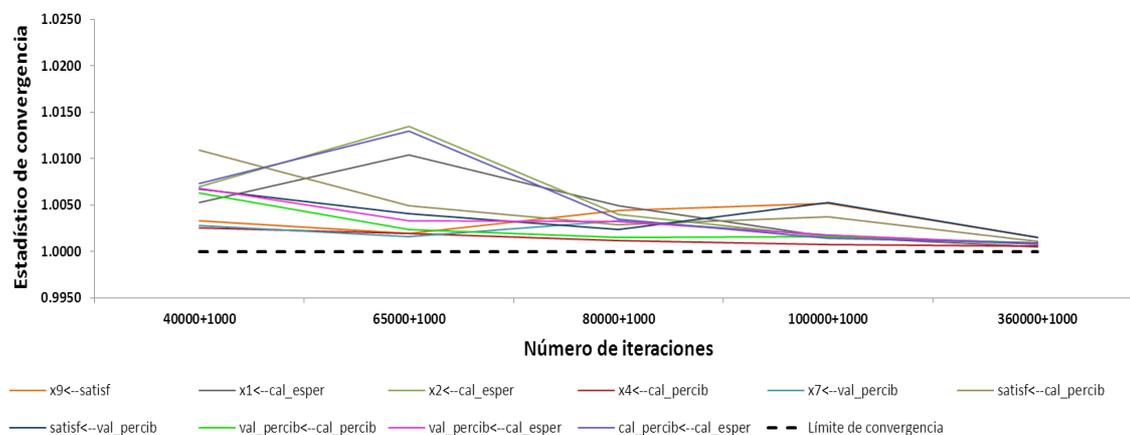


Figura 4.17: Convergencia de las estimaciones en el Análisis Factorial Confirmatorio

Mediante la anterior representación gráfica, se puede notar que los estadísticos de convergencia de cada una de las estimaciones converge a un valor cercano a 1 con mayor velocidad que en el AFE, además, hay que observar que los valores de los estadísticos de convergencia después de haber realizado 360000 es también menor que en el AFE,

esto se debe a que en el Análisis Factorial Confirmatorio se han incorporados las covarianzas entre variables dando como resultado un mejor ajuste global del modelo. Siguiendo la misma estructura de desarrollo del AFE, se procede a realizar las comparaciones entre las estimaciones obtenidas en las primeras 50000 y últimas 90000 iteraciones.

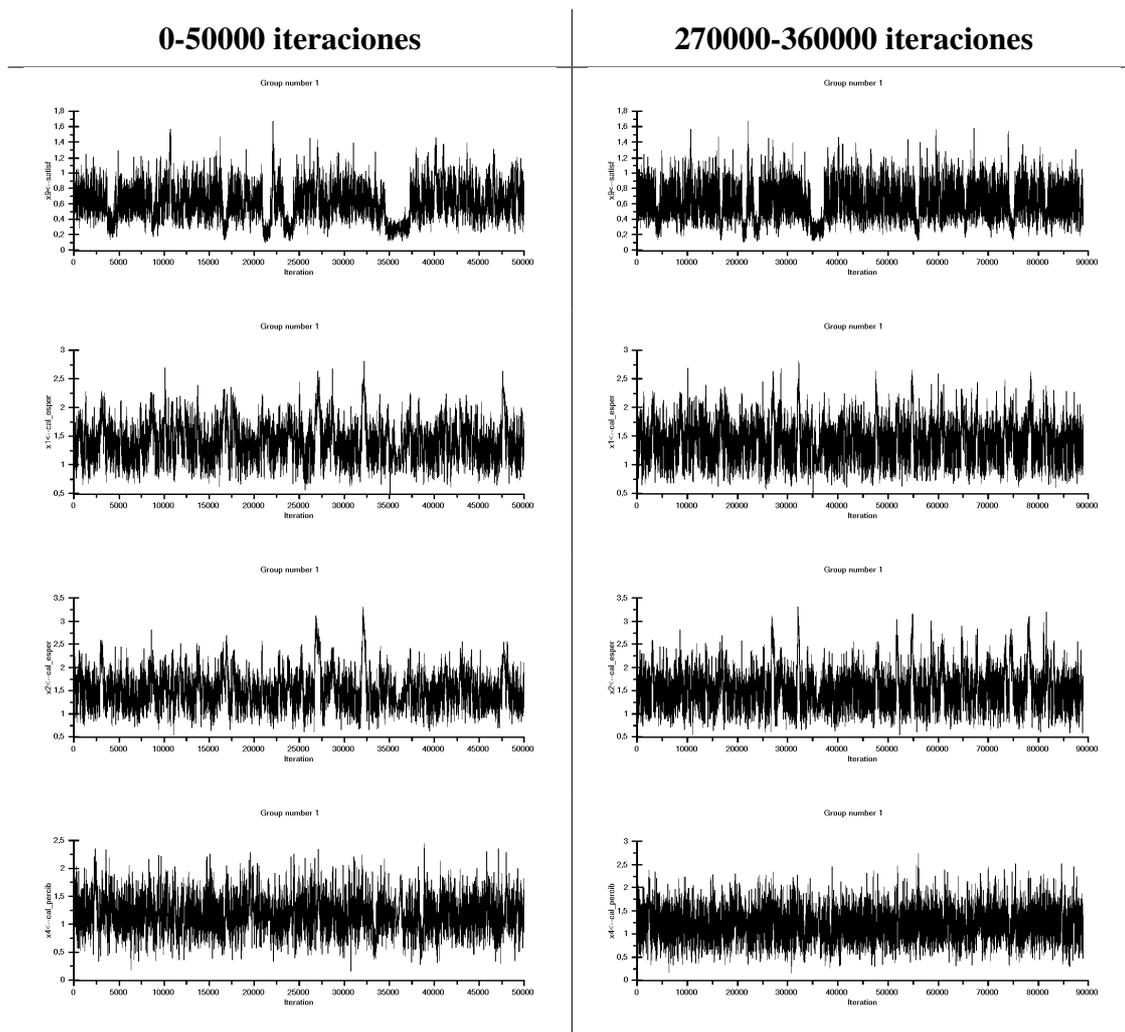


Figura 4.18: Comparación de series de iteraciones estimaciones bayesianas en el análisis factorial confirmatorio

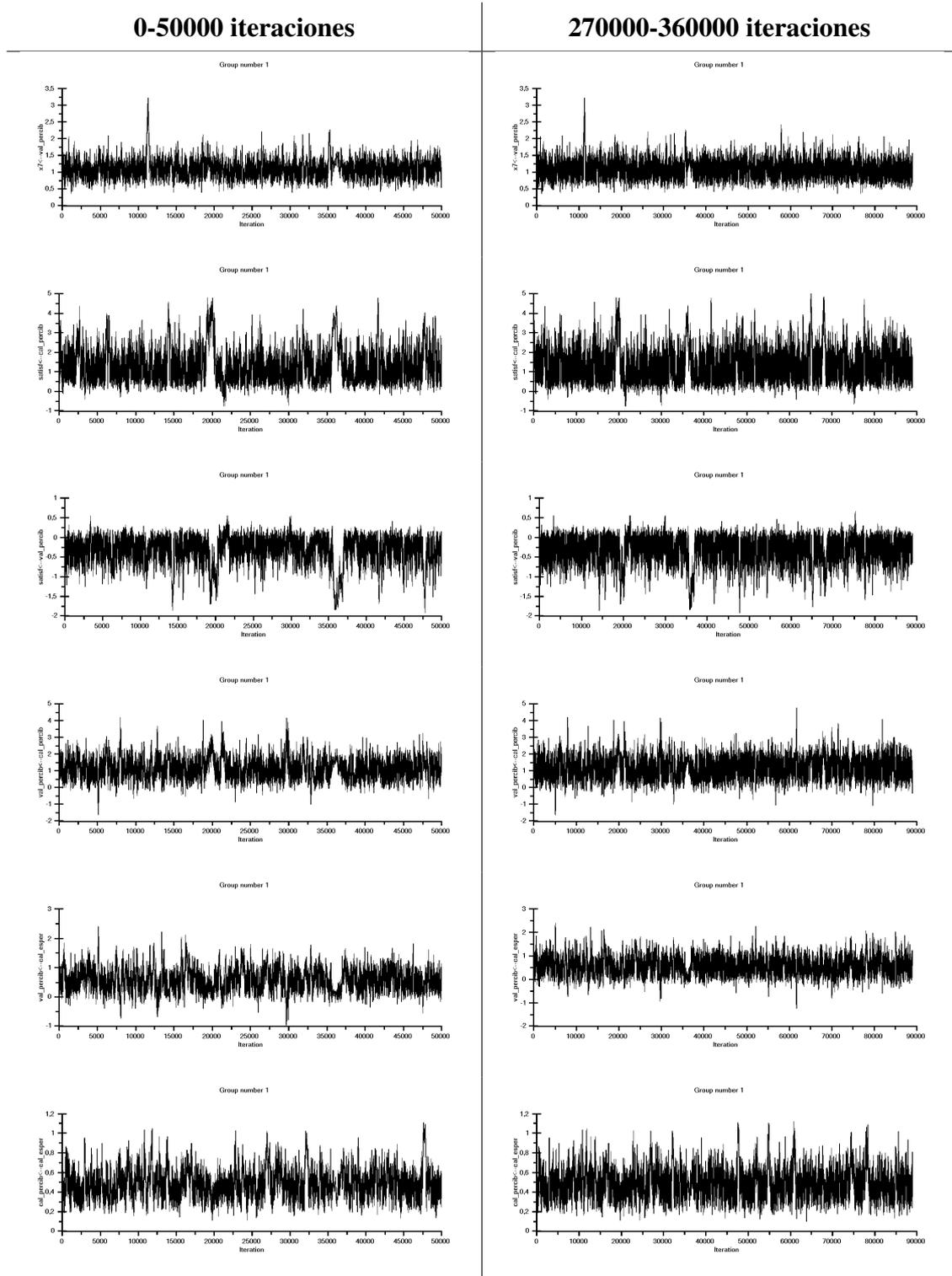


Figura 4.19: Comparación de series de iteraciones estimaciones bayesianas en el análisis factorial confirmatorio (continuación)

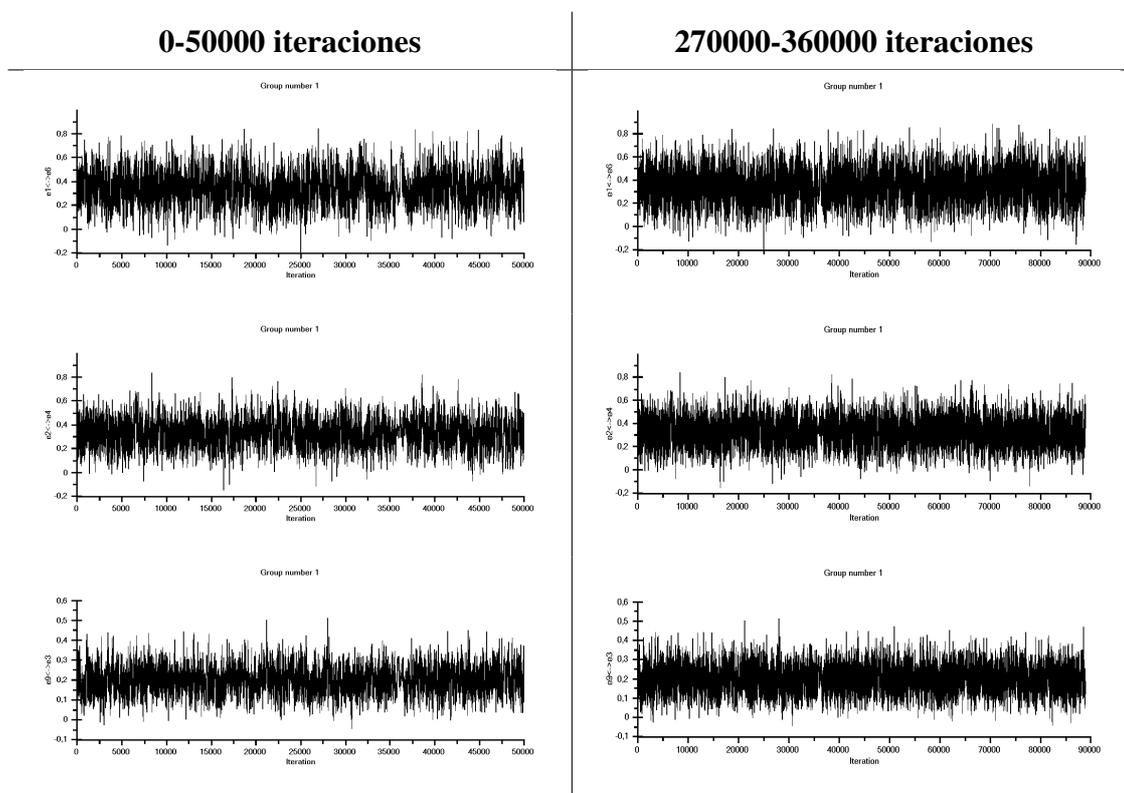


Figura 4.20: Comparación de series de iteraciones estimaciones bayesianas en el análisis factorial confirmatorio (continuación)

Para la comparación de las estimaciones, ya se ha eliminado 1000 iteraciones del periodo de calentamiento, considerando las primeras y las últimas 50000 iteraciones, se puede notar que ya en las primeras 100000 iteraciones la varianza de las estimaciones es reducida, es por esta razón que el estadístico de convergencia máximo alcanza el valor de 1,005, para este número de iteraciones ya se podía decir que el modelo alcanza una buena convergencia, hay que recordar que en AFE, la convergencia máxima fue 1,004, sin embargo por la naturaleza del AFC, se espera lograr mejores resultados debido a la incorporación de las covarianzas, se realizaron 360000 iteraciones, en donde la simulación del procesó alcanzó su convergencia máxima llegando a tomar el valor de 1,0015. En el anexo 4 se ha incluido los histogramas de las estimaciones bayesianas obtenidas, realizando la misma comparación que la hecha para series de iteraciones. Para continuar con el análisis, se presenta el siguiente cuadro con los

resultados obtenidos después de 360000 iteraciones.

Cuadro 4.12: Resultados (AFC)

Relación	Media	E.E	D.E.	E.C.	95 % L.I.	95 % L.S.	Sesgo	Curtosis
$x9 < -satisf$	0,633	0,011	0,207	1,0015	0,256	1,071	0,394	0,413
$x1 < -cal_esper$	1,418	0,010	0,323	1,0005	0,904	2,160	0,628	0,207
$x2 < -cal_esper$	1,540	0,017	0,391	1,0009	0,951	2,499	0,907	0,976
$x4 < -cal_percib$	1,217	0,010	0,306	1,0006	0,672	1,899	0,449	0,420
$x7 < -val_percib$	1,096	0,009	0,245	1,0007	0,711	1,631	1,257	4,665
$satisf < -cal_percib$	1,390	0,039	0,825	1,0011	0,313	3,599	1,192	1,443
$satisf < -val_percib$	-0,361	0,018	0,340	1,0015	-1,265	0,087	-1,2946	1,982
$val_percib < -cal_percib$	1,197	0,025	0,587	1,0009	0,187	2,443	0,460	0,901
$val_percib < -cal_esper$	0,572	0,015	0,370	1,0008	-0,043	1,386	0,408	0,735
$cal_percib < -cal_esper$	0,495	0,006	0,146	1,0009	0,268	0,847	0,840	0,942
$e1 < - > e6$	0,346	0,003	0,131	1,0003	0,095	0,611	0,090	0,058
$e2 < - > e4$	0,325	0,003	0,109	1,0004	0,119	0,546	0,147	0,145
$e9 < - > e3$	0,200	0,001	0,063	1,0001	0,082	0,328	0,194	0,133

E.E. : error estándar

D.E. : desviación estándar

E.C. : estadístico de convergencia

L.I. : límite inferior

L.S. : límite superior

Con el punto de partida de las estimaciones de máxima verosimilitud, se pudo ver que existían estimaciones que eran representativas, eran significativamente iguales a cero, esto, dentro del modelo estructural, después de haber simulado 360000 iteraciones, se han obtenido estimaciones que son significativas que permiten estimar principalmente la variable satisfacción, hay que notar también que las estimaciones bayesianas de las relaciones $satisf < -val_percib$, $cal_percib < -cal_esper$, no son significativas, esto evidencia la dificultad de la estimación de estas relaciones. Para las demás estimaciones bayesianas, se puede notar que son significativas permitiendo establecer ya el modelo que será utilizado para la estimación de la satisfacción del cliente.

3) Conclusiones del Análisis Factorial Confirmatorio

Con esta etapa se concluye el AFC, de todos los resultados obtenidos durante la realización del análisis, se extraen conclusiones sobre la validez del modelo, como principal

producto de este análisis se extraen las ecuaciones que permitan estimar las variables: satisfacción, calidad esperada, calidad percibida, y, valor percibido.

- La simulación fue realizada por 360000 iteraciones, para las cuales alcanza su mayor convergencia global (1,0015), para valores menores o mayores de las 360000 iteraciones, la convergencia disminuye, por lo tanto se ha tomado el mejor modelo posible para el AFC.
- El periodo de calentamiento se ha establecido en 1000 iteraciones, como se vió durante el calculo de las iteraciones necesarias para eliminar la autocorrelación, solamente se es necesario retirar las primeras 500 estimaciones, es valor es mucho menor al necesario en el AFE, avidenciando el éxito del modelo obtenido en el AFE.
- A medida que se llega a las 360000 iteraciones, los estadísticos de convergencia van acercandose a 1, si se observa la figura 4.22, ya para las primeras 80000 iteraciones el ajuste global del modelo alcanza el valor de 1,005, dicho valor ya es considerado bueno, y finalmente a las 360000 iteraciones se alcanzó una convergencia global de 1,0015. Estos resultados apoyan la validez del modelo, indicando que la velocidad de convergencia de las estimaciones bayesianas del AFC es mayor en comparación al AFE.
- En el cuadro 4.12, se puede observar que las desviación típicas son menores que 1, por lo tanto, la varianza es más reducida, indicando que existe poca variabilidad en las estimaciones bayesianas, es decir, ratificando la confiabilidad del modelo.
- Refiriendonos al mismo cuadro anlaizado anteriormente, se puede obsrvar que las estimaciones de las covarianzas son muy buenas, su desviación estándar es menor que 0,250, indicado que su varianza es más pequeña, además, el error estándar es cercano a cero, mostrando que las estimaciones bayesianas varían muy poco a medida que aumenta el número de iteraciones simuladas, los intervalos de confianza para estas relaciones no contienen al valor 0, confirmando su significancia estadística.

- En general las estimaciones bayesianas obtenidas después de realizar son muy buenas, en función de los resultados obtenidos, tanto para las desviaciones estándar, los estadísticos de convergencia y los intervalos de confianza. Este modelo simulado ha alcanzado un p-valor de 0,510 para la prueba de hipótesis de igualdad de matrices de varianzas estimada y observada, inicialmente se había obtenido un valor de 0,570, lamentablemente, en este modelo existían estimaciones no significativas y no permitía inferir sobre la satisfacción del cliente.
- La calidad percibida tiene un efecto positivo sobre la satisfacción del cliente, es decir, a medida que aumenta la calidad percibida, la satisfacción aumenta, además, la calidad esperada tiene un efecto positivo sobre la calidad percibida, en cambio la calidad esperada no tiene efecto sobre la satisfacción, es decir, en función del modelo, la imagen del producto que un cliente tiene no es suficiente para que este se encuentre satisfecho, los consumidores contrastan las sensaciones percibidas con la imagen que tienen del producto y esto es lo que satisface al cliente. Una estrategia de mercado debe estar basada en fortalecer las características evaluadas en los indicadores $X4$ para *calidad percibida* y en $X1$ y $X2$ para *calidad percibida*, ya que estos indicadores son los que más aportan positivamente a las variables latentes.
- Como conclusión final, el modelo obtenido después de la realización del AFC es totalmente válido para la estimación de la satisfacción del cliente debido a los resultados obtenidos.

Finalmente se establecen las ecuaciones que permiten estimar a las variables inobservables: *satisfacción*, *calidad esperada*, *calidad percibida* y *valor percibido*, las ecuaciones que se representan el modelo de medida.

$$X1 = 5,943 + 1,418 \cdot \widehat{cal_esper}$$

$$X2 = 5,584 + 1,540 \cdot \widehat{cal_esper}$$

$$X3 = 5,514 + \widehat{cal_esper}$$

$$X4 = 5,943 + 1,217 \cdot \widehat{cal_percib}$$

$$\begin{aligned}
X5 &= 3,005 + \widehat{cal_percib} \\
X6 &= 5,839 + \widehat{val_percib} \\
X7 &= 5,519 + 1,096 \cdot \widehat{val_percib} \\
X9 &= 2,110 + 0,633 \cdot \widehat{satisfac} \\
X10 &= 2,167 + \widehat{satisfac}
\end{aligned}$$

Mediante la combinación de las ecuaciones de cada variable inobservable y desjenuado las mismas se tiene el siguiente conjunto de ecuaciones:

$$\widehat{cal_esper} = \frac{X1 + X2 + X3 - 17,041}{3,958} \quad (4.1)$$

$$\widehat{cal_percib} = \frac{X4 + X5 - 8,948}{2,217} \quad (4.2)$$

$$\widehat{val_percib} = \frac{X6 + X7 - 11,358}{2,096} \quad (4.3)$$

$$\widehat{satisfac} = \frac{X9 + X10 - 4,277}{1,633} \quad (4.4)$$

Este conjunto de ecuaciones es el que permite la estimación de las variables inobservables.

SATISFACCIÓN DEL CLIENTE: RESULTADOS

Después de haber aplicado la metodología de modelización de estimación bayesiana de ecuaciones estructurales, el presente trabajo concluye enunciando los resultados obtenidos, previamente se hace una breve introducción a ciertos conceptos necesarios para lograr un mejor entendimiento.

Primero se debe responder la siguiente pregunta, ¿qué es necesario para satisfacer al cliente?, se debe considerar las características del producto que determinan el nivel de satisfacción del cliente, se debe tener en cuenta que las mismas no solo incluyen las cualidades o calidad del bien en cuestión, sino también las características de los demás servicios que las rodean. El hecho de brindar un valor agregado al bien producido añadiéndole características

propias de la empresa, puede lograr un aumento de la satisfacción y hacer que el cliente lo prefiera frente a otros productos del mismo tipo.

Por tanto, la satisfacción del cliente estimada tendrá los siguiente usos:



Figura 4.21: Usos de la satisfacción del cliente estimada

Para ayudar al lector, se presentan los elementos del modelo que permitirán concluir sobre la satisfacción del cliente:

- Nivel del factor: es el puntaje promedio alcanzado por la muestra considerada, se encuentra en una escala entre 0 y 100.
- Impacto o carga factorial: ayuda a predecir el incremento (o decremento) en la satisfacción del cliente, y es el resultado principal de la técnica de modelización utilizada, se encuentra en una escala entre 0 y 100.

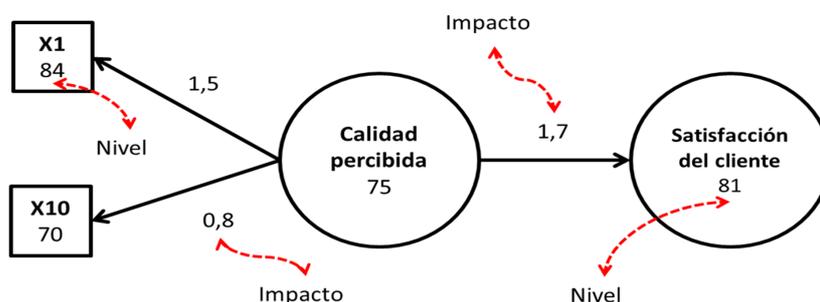


Figura 4.22: Ejemplo de lectura del modelo de ecuaciones estructurales

Los elementos presentados previamente son los cuales permiten la obtención de interpretaciones; además, para ayudar con la asociación de conceptos medidos se presenta el siguiente cuadro:

Cuadro 4.13: Asociación de conceptos

Concepto	Variable
Calidad esperada	
Credibilidad	X1
Cantidad de información	X2
Calificación del precio	X3
Calidad percibida	
Gusto por el diario	X4
Cercanía a sus gastos y necesidades	X5
Valor percibido	
Resistencia al cambio	X6
Innovación	X7
Satisfacción del cliente	
Frecuencia de lectura	X9
Frecuencia de compra	X10

Con la presentación del modelo resultante, se concluye con la exposición de los elementos necesarios para evaluar los resultados obtenidos:

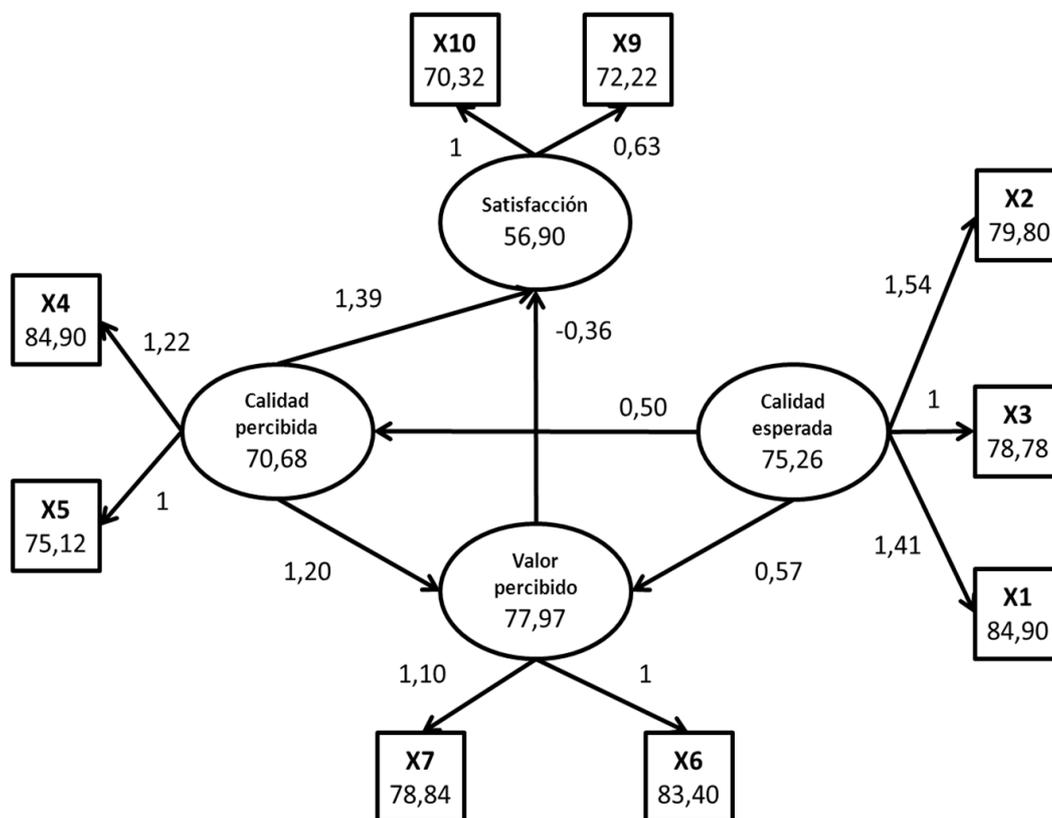


Figura 4.23: Modelo resultante

■ Calidad esperada:

- En promedio, la calidad esperada por lector alcanza un nivel de 75,26, indicando que todavía hace falta trabajo para que el lector tenga una mejor imagen del diario.
- Se puede notar que la credibilidad es el concepto mejor evaluado en promedio (84,90), esto refleja que el diario tiene credibilidad entre sus lectores, es decir, que lector espera obtener información veraz cuando compra el diario.
- Entre los conceptos evaluados, la cantidad de información es la de mayor impacto

(1,54), el lector compra el diario esperando obtener una “buena” cantidad de información.

- La cantidad de información (79,80) y el precio del diario (78,78) tienen similares calificaciones, indicando que los lectores perciben adecuado el precio del diario para la cantidad de información que este provee.

■ **Calidad percibida:**

- El nivel de calidad que el lector percibe alcanza el 70,68, indicando que este factor es el segundo menor evaluado, sugiriendo que hace una estrategia de mercado que ayude a mejorar el nivel de calidad percibida.
- El gusto por el diario es el concepto de mayor calificación (84,90) en cuanto a calidad percibida, es decir, al lector le gusta el diario y eso genera una percepción de mejor calidad en el usuario, esta aseveración se ve sustentada por un impacto de 1,22.
- La cercanía e identificación del lector con el diario tiene el menor peso en cuanto a considerar al bien de mejor calidad, además, un nivel de 75,12 indica que puede mejorar todavía esta sensación de cercanía.
- La calidad esperada aporta positivamente a la calidad percibida (0,50), esto indica que la imagen previa del diario que el lector tiene influye en el nivel de calidad que este percibe, no se puede esperar que el lector tenga una buena sensación de calidad si se deja descuidada la imagen que el diario proyecta a la población.

■ **Valor percibido:**

- El lector del diario, percibe un valor de 77,97, este factor es el de mejor nivel en comparación a los demás, el lector percibe que el bien que lee es muy valioso.
- La resistencia al cambio llega a un nivel de 83,40, esto debido que el lector menciona que prefiere quedarse con el diario que le presentan, debido a que lo percibe de un alto valor.

- La innovación del diario, es decir la adaptabilidad del diario a la actualidad es el factor de mayor impacto (1,10), el lector percibe que el diario se encuentra actualizado en cuanto a innovaciones presentadas.
 - La calidad percibida alcanza un impacto de 1,20 frente a un impacto de 0,57 de la calidad esperada, es decir, el lector contrasta sus expectativas del diario con el bien que tiene “en sus manos”, y este contraste es lo que influye o impacta en mayor cantidad a que el lector perciba un mayor valor del diario.
- **Satisfacción del cliente:**
- Este es el factor de menor evaluación, alcanzando un nivel de 56,90, esto indica que la satisfacción del lector es baja, y esto es de cuidado debido a que el mercado de los medios de comunicación se encuentra tanto periódicos gratuitos como pagados, este detalle se debe tomar en cuenta para mejorar la satisfacción del cliente.
 - La frecuencia de lectura y la frecuencia de compra alcanzan una evaluación de 72,22 y 70,32, respectivamente; se puede ver que la frecuencia de lectura tiene mejor evaluación, es decir que, el usuario lee más que antes y lo valora (al diario) más que antes.
 - El impacto de la frecuencia de compra (1) es mayor en cuanto a la satisfacción del cliente, un cliente satisfecho compra el diario frecuentemente, esto debido a que percibe alto nivel de calidad, alto valor percibido y tiene una buena imagen del bien.
 - El valor percibido no es estadísticamente significativo⁸ frente a la satisfacción del cliente, no se puede tomar como un valor influyente el impacto de $-0,36$, al contrario, se ve que el valor que el cliente percibe no actúa en la satisfacción del cliente.

⁸Este resultado se puede observar en el cuadro 4.12

Se ha podido identificar:

Fortalezas :

- Los lectores creen en la información presentada por el diario
- Los leyentes se resisten a cambiar de fuente de información
- A los usuarios les gusta el diario, en su forma y contenido
- Los clientes leen más el diario que antes,
- Los usuarios perciben un alto valor en el diario

Debilidades :

- Baja satisfacción del cliente
- La frecuencia de compra es la menor evaluada
- No existe relación significativa entre el valor percibido y satisfacción del cliente
- El lector no necesariamente se siente identificado con el diario

Oportunidades :

- La innovación tiene un fuerte impacto en la generación de valor
- La cantidad de información genera expectativa de calidad en el cliente
- La percepción de calidad del lector influye positivamente en la satisfacción

Para concluir, los resultados del modelo obtenido sugieren la implementación de una estrategia para aumentar la satisfacción del cliente, la cual esté enfocada en los siguientes aspectos:



Figura 4.24: Mecanismo de generación de satisfacción del cliente

Se debe aumentar la calidad esperada por el cliente fortaleciendo los aspectos de cantidad de información y precio acorde a los servicios proporcionados, no hay que olvidar que en el mercado se encuentran periódicos gratuitos y estos son de gran rotación en los lectores debido a su condición, la credibilidad es una fortaleza que tiene el diario, este es el pilar del medio de comunicación y debe ser cuidadosamente manejado, conjuntamente con las mejoras en la cercanía del diario hacia el lector (identificación del lector con el diario) aportarán a un aumento en la calidad percibida, un cliente que perciba mejor calidad en su diario de frecuente lectura hará que el cliente esté satisfecho, y por ende, la frecuencia de compra aumente.

Capítulo 5

CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

- El Teorema de Thomas Bayes es una herramienta muy útil en estos últimos tiempos debido al gran desarrollo de la computación, principalmente de los Métodos de Monte Carlo, permitiendo así, mejorar las estimaciones como por ejemplo, del método clásico de máxima verosimilitud.
- Los Métodos de Monte Carlo permiten la generación de muestras de vectores aleatorios con distribuciones estacionarias, estos métodos son de gran beneficio al momento de generar vectores aleatorios cuyas variables aleatorias sean dependientes entre si.
- Los Métodos de Monte Carlo están fundamentados en la constucción de cadenas de Markov reversibles en tiempo con distribución probabilidades límite, esta es la distribución a posteriori con la cual se realiza la inferencia bayesiana.
- Los modelos de ecuaciones estructurales posibilitan la estimación de variables inobservables aumentando la capaciad explicativa de los indicadores utilizados, esta característica da un instrumento apropiado al momento de analizar variables de difícil medición dentro del comportamiento humano.
- El Análisis Factorial Exploratorio y el Análisis Factorial Confirmatorio forman parte del mecanismo de estimación de variables latentes, inicialmente el AFE busca identificar las relaciones entre indicadores y variables latentes, a la vez busca desechar relaciones

que no aportan información mientras que el AFC busca establecer un modelo válido a partir del resultante del AFE haciendo uso de las covarianzas entre indicadores a fin de representar la realidad del fenómeno de estudio.

- La combinación de la estimación de máxima verosimilitud junto a la estimación bayesiana permiten la estimación de modelos de ecuaciones estructurales mejores en cuanto a significancia estadística y ajuste del modelo.
- La metodología *ACSI* es válida para la estimación de la satisfacción del cliente mediante el uso de la técnica de ecuaciones estructurales, además, se ha podido representar el fenómeno de la generación de satisfacción mediante otros factores inobservables como: la calidad percibida, la calidad esperada y el valor percibido.
- La satisfacción el cliente tiene una fuerte influencia de la calidad percibida, es decir, el consumidor es esceptico a la imagen de un producto y espera contrastar sus expectativas con las características que en verdad posee dicho producto.
- La combinación de una buena imagen y sobretodo un buen contenido del producto permitirá que los consumidores esten más satisfechos.
- Un cliente satisfecho es la mejor inversión, ya que este: se vuelve leal hacia la empresa productora, fiel al producto generado por ella y hace buenas referencias del mismo a otros posibles consumidores.
- Los modelos obtenidos tanto en el Análisis Factorial Exploratorio como en el Análisis Factorial Confirmatorio se ajustan bien a los datos debido a los buenos resultados en cuanto al ajuste y convergencia.

5.2. RECOMENDACIONES

- La modelización de ecuaciones estructurales es la única herramienta que permite incorporar variables inobservables, es aconsejable que se utilice esta herramienta al momento de investigar las relaciones entre conjuntos de variables observables y no observables, ya que esta técnica describirá las relaciones causales dentro del modelo.
- La combinación de la técnica de ecuaciones estructurales y la estimación bayesiana forman una herramienta poderosa que debería ser utilizada en el estudio de las variables relacionadas al comportamiento humano, en la actualidad, esta combinación ha tomado fuerza debido a las inmensas aplicaciones, principalmente comerciales, que se pueden realizar.
- Es necesario ampliar el estudio de la técnica de ecuaciones estructurales hacia áreas mucho más específicas en las cuales se traten, por ejemplo, variables dicotómicas, variables nominales, ecuaciones estructurales no lineales, y por qué no, la combinación de los modelos multinivel con modelos de ecuaciones estructurales; todo esto como un aporte científico al estudio en el ámbito social.
- Se recomienda realizar varios tipos de modelos de ecuaciones estructurales con el uso de varias metodologías, como la *ACSI*, para plantear el Análisis Factorial Confirmatorio con la finalidad de poder contrastarlas y seleccionar la mejor opción en función de los índices de ajuste.
- Es conveniente realizar un modelo de ecuaciones estructurales donde se incorpore factores como la lealtad y la fidelidad con sus respectivos indicadores, ya que estos factores ayudarían a representar ampliamente la realidad del consumidor.
- Para hoy en día existen diferentes variaciones del Algoritmo de Hasting Metrópolis, es oportuno realizar un modelo de ecuaciones estructurales y comparar las estimaciones bayesianas obtenidas con diferentes métodos de Monte Carlo en especial las variaciones del Algoritmo de Hasting Metrópolis.

- Es recomendable construir un cuestionario en donde se incorpore una batería de preguntas por cada factor latente, y separados por variables de control, con el objetivo de poder medir de mejor manera estos factores y minimizar la correlación entre ellos. El valor percibido es un factor de difícil estimación por tanto se debería incorporar mayor número de indicadores para lograr una mejor estimación.
- Después del estudio de metodologías como la *ACSI*, es apropiada la combinación de estas metodologías para la construcción del Índice de Satisfacción del Consumidor Ecuatoriano incorporando información de la Encuesta Nacional de Ingresos y Gastos de Hogares Urbanos y Rurales (ENIGHUR).

Bibliografía

- [1] ATHREYA K. y DELAMPADY Mohan y KRISHNAN T., “Markov Chain Montecarlo Methods”, 2003.
- [2] BERMUDEZ José, BEAMONTE Eduardo, “Análisis Bayesiano de Datos de Supervivencia Gamma utilizando muestreo de Gibbs”, Estadística Española, Volumen 35, Número 134, 1993.
- [3] BYRNE Barbara, “Structural Equations Modeling with AMOS: basic concepts, applications, and programming”, Multivariate applications series, New York, 2009.
- [4] BRADLEY Carlin, THOMAS Louis, “Bayes and Empirical Bayes Methods for Data Analysis”, Chapman & Hall/CRC, New York, 1996.
- [5] CAPA Holger, “Simulación MCMC”, 2009.
- [6] COBA Marielisa, “Modelización de ecuaciones estructurales”, Proyecto de titulación previo a la obtención del título de Ingeniería Matemática, Facultad de Ciencias, Escuela Politécnica Nacional, 2006.
- [7] CONGDON Peter, “Applied Bayesian Modelling”, John Wiley & Sons, Chichester, 2003.
- [8] CONGDON Peter, “Bayesian Statistical Modelling”, John Wiley & Sons, Chichester, 2006.
- [9] DEPARTAMENTO DE ESTADÍSTICA, “Montecarlo Basado en Cadenas de Markov”, Universidad de Barcelona, Barcelona.

- [10] DIMOV Ivan, “Monte Carlo Methods for Applied Scientists”, World Scientific Publishing, Singapore, 2008.
- [11] FAHRMEIR Ludwig y TUTZ Gerhard, “Multivariate Statistical Modelling Based on Generalized Linear Models”, Springer, New York, 1994.
- [12] GILKS Walter y RICHARDSON Sylvia y SPIEGELHALTER David, “Markov Chain Monte Carlo in Practice”, Chapman & Hall/CRC, 1996.
- [13] HAIR Joseph, ANDERSON Rolph, TATHAM Ronald, BLACK William, “Análisis Multivariante”, Prentice Hall, 2005.
- [14] KALOS Malvin y WHITLOCK Paula, “Monte Carlo Methods”, Wiley-VCH, República Federal de Alemania, 2004.
- [15] KNIGHT Keith, “Mathematical Statistics”, Chapman & Hall/CRC, Florida, 2000.
- [16] KOTLER Philip, “Dirección de mercadotecnia”, Prentice Hall, México, 2001.
- [17] LEHMANN E.L. y CASELLA George, “Theory of point estimation”, Springer, New York, 1998.
- [18] LESAGE James, “Applied Econometrics using MATLAB”, Department of Economics, University of Toledo, Toledo, 1999.
- [19] LEE Sik, “Structural Equation Modeling A Bayesian Approach”, John Wiley & Sons, Inglaterra, 2007.
- [20] LUCEÑO Alberto y GONZÁLES Francisco Javier, “Métodos Estadísticos para medir, describir y controlar la variabilidad”, Universidad de Cantabria, 2004.
- [21] LLATAS Isabel, “Guía Rápida: WinBUGS”.
- [22] MAYORGA J, “Inferencia Estadística”, Departamento de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia, 2004.

- [23] NTZOUFRAS Loannis, “Bayesian Modeling Using WinBUGS”, John Wiley & Sons, New Jersey, 2009.
- [24] PRIETO Víctor, QUINTERO Constanza, RODRÍGUEZ Ismael, “Análisis de bayes empírico mediante un ejemplo”, Revista de Colombiana de Estadística N°31 y N°32, Departamento de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia, 1995.
- [25] TU Yundong, “Markov Chain Monte Carlo: Methods and Programming in MATLAB”, Universidad de California, California, 2009.
- [26] ROBERT Christian, CASELLA George, “Monte Carlo Statistics Method”. Université Paris Dauphine, Paris, 2001.
- [27] ROSS Sheldon, “Introduction to Probability Models”, ELSEVIER Inc, San Diego, 2007.
- [28] ROSS Sheldon, “Simulation”, ELSEVIER Inc. San Diego. 2006.
- [29] ROSSI Peter y ALLENBY Greg y McCULLOCH Robert, “Bayesian Statistics and Marketing”, John Wiley & Sons, Chichester, 2005.
- [30] SHAO Jun, “Mathematical Statistics: Exercises and Solution”, Springer, New York, 2005.
- [31] SKRONDAL Anders y RABE-HESKETH Sophia, “Generalized Latent Variable Modeling”, Chapman & Hall/CRC, Florida, 2004.
- [32] SPIEGELHALTER David y THOMAS Anders y BEST Nicky y LUNN Dave, “WinBugs User’s Manual”, Version 1.4, Enero, 2003.
- [33] SÖRBOM Dag, SATORRA Albert, SARIS Willem, “The Detection and Correction of Specification Errors in Structural Equations Models”, Sociological Methodology, Vol 17, Pg 105-129, 1987.

- [34] WU Botao, “Invariant Probability Distributions”, Department of Mathematics, Faculty of Science, National University of Singapore, 2012.
- [35] XUE Le y YANG Chung, “An Exploratory Study of Customer Satisfaction Based on ACSI Model”, University of Gävle, Diciembre, 2008.
- [36] YOHAI Víctor, BOENTE Graciela, “Notas de estadística”, Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 2006.

SITIOS WEB

PROMONEGOCIOS. Satisfacción de cliente: Método de Seguimiento

Obtenido de: <http://www.promonegocios.net/mercadotecnia/satisfaccion-cliente-seguimiento.htm>

WIKIPEDIA. Distribución Beta.

Obtenido de: http://es.wikipedia.org/wiki/Distribuci%C3%B3n_beta

WIKIPEDIA. Distribución Gamma.

Obtenido de: http://es.wikipedia.org/wiki/Distribuci%C3%B3n_gamma

WIKIPEDIA. Función Beta.

Obtenido de: http://es.wikipedia.org/wiki/Funci%C3%B3n_beta

WIKIPEDIA. Función Gamma.

Obtenido de: http://es.wikipedia.org/wiki/Funci%C3%B3n_gamma

WIKIPEDIA. Método de Monte Carlo.

Obtenido de: http://es.wikipedia.org/wiki/M%C3%A9todo_de_Montecarlo

WIKIPEDIA. Thomas Bayes.

Obtenido de: http://es.wikipedia.org/wiki/Thomas_Bayes

ANEXOS

ANEXO 1

1. Función Gamma

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

(Wikipedia, 2012)

2. Función Beta

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$$

ó

$$B(x, y) = \int_0^{\infty} \frac{t^{x-1}}{(1+t)^{x+y}} dt$$

(Wikipedia, 2012)

3. Distribución Gamma

Sea $x > 0$

$$f(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{k-1}}{\Gamma(k)}$$

$$E(X) = \frac{k}{\lambda}$$

$$V(X) = \frac{k}{\lambda^2}$$

(Wikipedia, 2012)

4. Distribución Beta

Sea $0 < x < 1$

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}$$

$$E(X) = \frac{a}{a+b}$$

$$V(X) = \frac{ab}{(a+b+1)(a+b)^2}$$

(Wikipedia, 2012)

5. Sintaxis de Matlab: Gráfica Distribución Beta(0.001;0.001)

```
>>x=0:0.01:1;
>>y=betapdf(x,0.001,0.001);
>>plot(x,y); rgrid
>>xlabel('theta');
>>ylabel('f(theta)');
>>title('Distribución Beta')
```

ANEXO 2

Demostración del Teorema 4: Probabilidades Límite

Para la realización de la demostración, es necesario definir previamente el Teorema de Perron Frobenius y un Lema, la demostración puede encontrarse en el trabajo de WU (2012).

TEOREMA 9. Teorema de Perron Frobenius

Sea P una matriz $n \times n$ con todas sus entradas positivas, entonces P tiene un valor propio dominante α tal que:

- $\alpha > 0$, y el vector propio asociado tiene todas sus entradas positivas
- cualquier otro valor propio asociado k de P , satisface $|k| < \alpha$

Demostración. Sea $P = (p_{ij})$ una matriz $n \times n$ con $p_{ij} > 0$, además, si $\vec{v} > \vec{0}$, entonces $P\vec{v} > 0$. Se define $G := \{\lambda > 0 : P\vec{v} \geq \lambda\vec{v} \text{ para algún } \vec{v} \in [0, \infty[^n \text{ y } \vec{v} \neq \vec{0}\}$. Sea α el valor máximo en G , entonces $P\vec{v} \geq \alpha\vec{v}$ para algún $\vec{v} \in [0, \infty[^n$.

En verdad lo que se tiene es $P\vec{v} = \alpha\vec{v}$ para algún \vec{v} con estradas positivas, suponiendo que no es así, entonces $P\vec{v} - \alpha\vec{v} > \vec{0}$. Dado $p_{ij} > 0$, se tiene $P(P\vec{v}) - P(\alpha\vec{v}) > \vec{0}$, entonces existe $\alpha' > \alpha > 0$ tal que $P(P\vec{v}) - \alpha'(P\vec{v}) \geq 0$, como $P\vec{v} > 0$, entonces $P\vec{v} \in [0, \infty[^n$. Así $\alpha' \in G$, esto contradice que α es el máximo de G . Entonces $P\vec{v} = \alpha\vec{v}$.

Para continuar con la demostración, se debe probar que existe un único \vec{v} tal que $P\vec{v} = \alpha\vec{v}$.

Suponiendo que no es así, $P\vec{w} = \alpha\vec{w}$ con \vec{w} distinto de cualquier múltiplo de \vec{v} . Por lo probado anteriormente se sabe que $\vec{w} > \vec{0}$, entonces existe un $c \in \mathbb{R}$ tal que $\vec{w} + c\vec{v} \geq \vec{0}$, $\vec{w} + c\vec{v} \neq \vec{0}$, y $\vec{w} + c\vec{v}$ tiene entradas con valor de 0.

Como $P\vec{w} = \alpha\vec{w}$ y $cP\vec{v} = P(c\vec{v}) = c\alpha\vec{v} = \alpha(c\vec{v})$ se tiene $P\vec{w} + Pc\vec{v} = \alpha\vec{w} + \alpha(c\vec{v})$ y $P(\vec{w} + c\vec{v}) = \alpha(\vec{w} + c\vec{v})$ como $\vec{w} + c\vec{v} \geq \vec{0}$ se tiene $P(\vec{w} + c\vec{v}) > \vec{0}$, así $\alpha(\vec{w} + c\vec{v}) > \vec{0}$, en particular $\vec{w} + c\vec{v} > \vec{0}$ dado $\alpha > 0$, lo cual es una contradicción.

Se supone que $P\vec{w} = k\vec{w}$ para $k \neq \alpha$, entonces $|k||\vec{w}| = |k\vec{w}| = |P\vec{w}| \leq |P||\vec{w}| = P|\vec{w}|$ donde $|\vec{w}| = (|w_1|, \dots, |w_n|)$ por tanto $|k| \in G$, y $|k| \leq \alpha$. \square

LEMA 1. Sea P una matriz de transición de una cadena de Markov irreducible y aperiódica, entonces, existe $M > 0$, tal que, para todo $n > M$, P^n tiene todas sus entradas estrictamente positivas.

Demostración. Como la cadena de Markov es irreducible, entonces para cualquier par de estados i, j existe m_{ij} tal que $P_{ij}^{m_{ij}} > 0$, además, la cadena de Markov es aperiódica, entonces existe M_i tal que para todo $n > M_i$, $P_{ij}^n > 0$. Así, para cualquier $n \geq M_i$, $P_{ij}^{n+m_{ij}} \geq P_{ij}^n P_{ij}^{m_{ij}}$, como el espacio de estados es finito, se puede considerar el valor máximo M de $M_i + m_{ij}$ para todo i, j . Entonces para cualquier $n > M$ y para todo i, j , $P_{ij}^n > 0$. \square

Con estos antecedentes, se enuncia nuevamente el Teorema 4.

Si P es una matriz estocástica para una cadena de Markov irreducible y aperiódica entonces:

- existe una única probabilidad invariante tal que $\vec{\pi} P = \vec{\pi}$
- $\vec{\pi}$ puede ser calculada con

$$\lim_{n \rightarrow \infty} P_{ij}^n = \pi_j$$

y si \vec{v} es cualquier vector de probabilidad inicial

$$\lim_{n \rightarrow \infty} \vec{v} P_{ij}^n = \vec{\pi}$$

Demostración. De la aplicación del teorema de Perron Frobenius se sabe que se cumple con $\vec{\pi} P = \vec{\pi}$, con ello se demuestra el primer literal. Para el segundo literal, sea P una matriz con todas sus entradas positivas, si P tuviera entradas con el valor de 0, el lema anterior menciona que que P elevada a una potencia $n(P^n)$ es una matriz con todas sus entradas positivas. Considerando la entrada i, j de $P^{t+1} = P P^t$, que es $P_{ij}^{t+1} = \sum_k P_{ik} P_{kj}^t$.

Sea $m_j^{(t)} := \min_i P_{ij}$, y $M_j^{(t)} = \max_i P_{ij}^t$ tal que $0 \leq m_j^{(t)} \leq M_j^{(t)} \leq 1$. Entonces:

$$\begin{aligned}
 m_j^{(t+1)} &= \min_i \sum_k P_{ik} P_{kj}^t \\
 &\geq \min_i \sum_k P_{ik} P_{ij}^t \\
 &= \min_i P_{ij}^t \sum_k P_{ik} \\
 &= m_j^{(t)} \cdot 1 \\
 &= m_j^{(t)}
 \end{aligned}$$

Así, la sucesión $\{m_j^{(1)}, m_j^{(2)}, \dots\}$ es no decreciente, de la misma forma la sucesión $\{M_j^{(1)}, M_j^{(2)}, \dots\}$ es no creciente, además, $\{m_j^{(1)}, m_j^{(2)}, \dots\}$ está acotada por $M_j^{(1)}$, y $\{M_j^{(1)}, M_j^{(2)}, \dots\}$ está acotada por $m_j^{(1)}$.

Existe m_j y M_j tal que:

$$\lim_{t \rightarrow \infty} m_j^{(t)} = m_j$$

y

$$\lim_{t \rightarrow \infty} M_j^{(t)} = M_j$$

y

$$m_j \leq M_j$$

Ahora se probará que $m_j = M_j$. Para ello se calcula:

$$\begin{aligned}
 M_j^{(t+1)} - m_j^{(t+1)} &= \max_i \sum_k P_{ik} P_{kj}^t - \min_l \sum_k P_{lk} P_{kj}^t \\
 &= \max_{il} \sum_k P_{kj}^t (P_{ik} - P_{lk}) \\
 &= \max_{il} \left[\sum_k P_{kj}^t (P_{ik} - P_{lk})^+ + \sum_k P_{kj}^t (P_{ik} - P_{lk})^- \right] \\
 &\leq \max_{il} \left[M_j^{(t)} \sum_k (P_{ik} - P_{lk})^+ + m_j^{(t)} \sum_k (P_{ik} - P_{lk})^- \right] (*)
 \end{aligned}$$

donde:

$\sum_k (P_{ik} - P_{lk})^+$: es la suma de términos positivos de $P_{ik} - P_{lk}$

$\sum_k (P_{ik} - P_{lk})^-$: es la suma de términos negativos de $P_{ik} - P_{lk}$

Además

$$\begin{aligned}
 \sum_k (P_{ik} - P_{lk})^- &= \sum_k P_{ik}^- - \sum_k P_{lk}^- \\
 &= 1 - \sum_k P_{ik}^+ - 1 + \sum_k P_{lk}^+ \\
 &= \sum_k P_{lk}^+ - \sum_k P_{ik}^+ \\
 &= - \sum_k (P_{ik} - P_{lk})
 \end{aligned}$$

Así, (*)

$$M_j^{(t+1)} - m_j^{(t+1)} \leq (M_j^{(t)} - m_j^{(t)}) \max_{il} \sum_k (P_{ik} - P_{lk})^+$$

Si $\max_{il} \sum_k P_{ik} - P_{lk} = 0$, entonces $M_j^{(t+1)} = m_j^{(t+1)}$.

Por otro lado, sea r el número de términos en k tal que $P_{ik} - P_{lk} > 0$, y, sea s el número de términos en k tal que $P_{ik} - P_{lk} < 0$. Entonces $r \geq 1$. Sea $n := r + s \geq 1$. Se denota $\delta := \min_{ij} P_{ij}$, con esto se puede escribir:

$$\begin{aligned}
 \sum_k (P_{ik} - P_{lk})^+ &= \sum_k P_{ik}^+ - \sum_k P_{lk}^+ \\
 &= 1 - \sum_k P_{ik}^- - \sum_k P_{lk}^+ \\
 &\leq 1 - s\delta - r\delta \\
 &= 1 - n\delta
 \end{aligned}$$

Así, (*)

$$\begin{aligned}
 M_j^{(t+1)} - m_j^{(t+1)} &\leq (1 - \delta)(M_j^{(t)} - m_j^{(t)}) \\
 &\leq (1 - \delta)^t (M_j - m_j) \rightarrow 0
 \end{aligned}$$

cuando $t \rightarrow \infty$.

Denotando $\pi_j = M_j = m_j > 0$.

Dado $m_j^{(t)} \leq P_{ij}^t \leq M_j^{(t)}$, entonces

$$\begin{aligned}\lim_{t \rightarrow \infty} P_{ij}^{(t)} &= \lim_{t \rightarrow \infty} m_j^{(t)} \\ &= \lim_{t \rightarrow \infty} M_j^{(t)} \\ &= \pi_j\end{aligned}$$

de esta forma:

$$\lim_{t \rightarrow \infty} P_{ij}^{(t)} = \pi_j$$

Además, para cualquier vector de probabilidad \vec{v} con $\sum_i v_i = 1$, se tiene:

$$\begin{aligned}\lim_{t \rightarrow \infty} \vec{v} P_{ij}^{(t)} &= (v_1, \dots, v_n) \begin{pmatrix} \pi_1 & \dots & \pi_n \\ \vdots & \ddots & \vdots \\ \pi_n & \dots & \pi_n \end{pmatrix} \\ &= (\pi_1 \sum_i v_i, \dots, \pi_n \sum_i v_i) \\ &= (\pi_1, \dots, \pi_n) \\ &= \vec{\pi}\end{aligned}$$

□

■ **Sintaxis de Matlab: Muestra de la distribución beta utilizando el algoritmo de Hasting Metropolis**

```
function MH = Hastingmetropolis(N, a, b)
MH(1) = rand(1);
x = MH(1);
for i = 1 : N
y = rand(1, 1);
p = (y(a-1) * (1 - y)(b-1)) / (x(a-1) * (1 - x)(b-1));
alpha = min(1, p);
if rand(1) > alpha
y = x;
```

```

end
MH(i + 1) = y;
i = i + 1;
x = y;
end

```

- **Sintaxis de R: Muestra de la distribución normal bivariada utilizando el algoritmo de Hasting Metropolis componente a componente**

```

library(MASS);
mu = c(0, 0);
sigma = matrix(c(1, 0, 3, 0, 3, 1), 2);
T = 5000;
propsigma = 1;
thetamin = c(-3, -3);
thetamax = c(3, 3);
set.seed = (1);
theta1 = runif(1, thetamin[1], thetamax[1]);
theta2 = runif(1, thetamin[2], thetamax[2]);
state = mat.or.vec(2, T);
rho <- -sigma[1, 2]/(sqrt(sigma[1, 1]) * sqrt(sigma[2, 2]));
f <- -function(x1, x2){
term1 <- -1/(2 * pi * sqrt(sigma[1, 1] * sigma[2, 2] * (1 - rho^2)))
term2 <- - - 1/(2 * (1 - rho^2))
term3 <- -(x1 - mu[1])^2/sigma[1, 1]
term4 <- -(x2 - mu[2])^2/sigma[2, 2]
term5 <- - - 2 * rho * ((x1 - mu[1]) * (x2 - mu[2]))/(sqrt(sigma[1, 1]) *
sqrt(sigma[2, 2]))
term1 * exp(term2 * (term3 + term4 - term5))}
for(t in 1 : T){

```

```

newtheta1 = rnorm(1, theta1, propsigma);
pratio = f(newtheta1, theta2)/f(theta1, theta2);
alpha = min(c(1, pratio));
u = runif(1, 0, 1);
if(u < alpha){
theta1 = newtheta1; }
newtheta2 = rnorm(1, theta2, propsigma);
pratio = f(theta1, newtheta2)/f(theta1, theta2);
alpha = min(c(1, pratio));
u = runif(1, 0, 1);
if(u < alpha){
theta2 = newtheta2; }
state[1, t] = theta1;
state[2, t] = theta2; }
x1 <- seq(thetamin[1], thetamax[1], length = 41);
x2 <- seq(thetamin[2], thetamax[2], length = 41);
hist(state[1, ], breaks = 12, col = "greenyellow", border = "black", xlim =
c(-3, 3), xlab = ' theta1', ylab = ' Frecuencia', main = ' Histograma de theta1')
hist(state[1, ], breaks = 12, , freq = FALSE, col = "forestgreen", border =
"black", xlim = c(-3, 3), xlab = ' theta1', ylab = ' Probabilidad',
main = ' Función de probabilidad ( theta 1- > N(0, 1) )')
curve(dnorm(x, 0, 1), -3, 3, add = TRUE, col = "firebrick1", lwd = 4)
hist(state[2, ], breaks = 12, col = "greenyellow", border = "black",
xlim = c(-3, 3), xlab = ' theta2', ylab = ' Frecuencia',
main = ' Histograma de theta2')
hist(state[2, ], breaks = 12, , freq = FALSE,
col = "forestgreen", border = "black",
xlim = c(-3, 3), xlab = ' theta2', ylab = ' Probabilidad',
main = ' Función de probabilidad ( theta 2- > N(0, 1) )')

```

```

curve(dnorm(x, 0, 1), -3, 3, add = TRUE, col = "firebrick1", lwd = 4)
par(mfrow = c(1, 1), pty = "s")
jet.colors <- colorRampPalette(c("springgreen", "springgreen4"))
nbc <- 100
color <- jet.colors(nbc)
nrz <- nrow(z)
ncz <- ncol(z)
zfacet <- z[-1, -1] + z[-1, -ncz] + z[-nrz, -1] + z[-nrz, -ncz]
facetcol <- cut(zfacet, nbc)
persp(x1, x2, z, main = "Distribución Normal Bivariante",
xlab = 'theta1', ylab = 'theta2', zlab = 'f',
col = color[facetcol],
theta = 45, phi = 30,
r = 10, d = 0.1, expand = 0.7,
ticktype = "detailed", nticks = 6)

```

■ **Sintaxis de Matlab: Estimación de momentos de la función de distribución binomial utilizando el muestreador de Gibbs**

```

function [x, y, mediox, mediy] = mgibbs(k, m, a, b, n)
xx = zeros(1, k);
yy = zeros(1, k);
xx(1) = binornd(n, 0.5, 1, 1);
yy(1) = betarnd(xx(1) + a, n - xx(1) + b, 1, 1);
for i = 2 : k
xx(i) = binornd(n, yy(i - 1), 1, 1);
yy(i) = betarnd(xx(i) + a, n - xx(i) + b, 1, 1);
end
x(1, 1 : k - m) = xx(1, m + 1 : k);
y(1, 1 : k - m) = yy(1, m + 1 : k);

```

```
mediax = zeros(1, k - m - 1);  
mediax(1) = x(1, 1);  
mediay = zeros(1, k - m - 1);  
mediay(1) = y(1, 1);  
for i = 2 : k - m  
  mediax(i - 1) = mean(x(1, 1 : i));  
  mediay(i - 1) = mean(y(1, 1 : i));  
end
```

ANEXO 3

Medidas de Ajuste para Modelos de Ecuaciones Estructurales

Una vez planteado uno o varios modelos de ecuaciones estructurales surge la pregunta ¿El modelo tiene buen ajuste?, no existe una única medida de ajuste que nos pueda indicar que tan bueno o malo es un modelo, para esto, se han incorporado un conjunto de medidas para analizar su ajuste y adecuación a los datos usados, las medidas que se han considerado se toman del trabajo de Coba (2006).

1. Ajuste Global del Modelo

Estadístico Ji-cuadrado es una de las medidas comúnmente usadas. Es una prueba estadística donde se utiliza la distribución χ^2 y mide las distancias entre la matrices de covarianza observada y la estimada.

Bajo la suposición de que los datos siguen una distribución normal multivariante la prueba de hipótesis se plantea de la siguiente forma.

$$H_0 : S = \Sigma$$

$$H_1 : S \neq \Sigma$$

Si el estadístico χ^2 tiene valores altos se rechazará la hipótesis nula, y si el estadístico χ^2 se aceptará la igualdad de las matrices.

Estadístico Ji-cuadrado no centrado surge del estadístico χ^2 usando anteriormente, en este caso se le hace una corrección utilizando los grados de libertad (calculados con la ecuación 3.1).

$$NCP = \chi^2 - \text{grados de libertad}$$

Los grados de libertad son los calculados en un modelo de ecuaciones estructurales. Para aceptar la hipótesis de igualdad de la matriz de covarianza poblacional con la muestral el estadístico NCP tiene que ser cercano a cero.

Raíz Cuadrada del Error Medio Cuadrático (RMSEA) Cuando la muestra supera los 400 casos en la prueba χ^2 aumenta la probabilidad de rechazar el modelo, para evitar este tipo de problema surge esta medida, la cual, mide la cercanía de la matriz de covarianza como si fuera calculada en la población y no en la muestra.

$$RMSEA = \sqrt{\left[\frac{\chi^2 - gl}{gl(n - 1)} \right]}$$

Donde n es el tamaño de la muestra. Para valores entre $[0; 0,5]$ se puede decir que el modelo tiene buen ajuste, a medida que va creciendo el valor de $RMSEA$ el modelo se ajusta menos.

Índice de Bondad de Ajuste (GFI) este índice muestra la variabilidad del modelo sus valores están entre 0 y 1, a medida que GFI se acerca a 1 el modelo tiene mejor ajuste, caso contrario, si GFI tiende hacia el valor de 0 el modelo ofrece un ajuste pobre. Este índice tiene influencia del tamaño de la muestra. Tiene el mismo papel que el coeficiente de determinación R^2 . Para el cálculo de GFI se utiliza la siguiente ecuación.

$$GFI = 1 - \frac{tr(\Sigma^{-1}S - I)^2}{tr(\Sigma^{-1}S)^2}$$

I es la matriz identidad.

Índice de Bondad de Ajuste Relativo (RGFI) está dado por el coeficiente calculado en la parte anterior GFI estimado y el GFI esperado considerando el tamaño

de la muestra y la cantidad de indicadores del modelo. Se tiene:

$$EGFI = \frac{1}{1 + \left(\frac{2gl}{n(p+q)}\right)}$$

Y para $RGFI$ se tiene:

$$RGFI = \frac{GFI}{EGFI}$$

Si el valor de $RGFI$ tiende hacia 1 el modelo tendrá mejor ajuste.

Índice de Bondad de Ajuste Ajustado (AGFI) si el índice presenta un valor superior a 0,90 se puede considerar que tiene un buen ajuste. Se calcula de la siguiente forma:

$$AGFI = 1 - \left[\frac{(p+q)(p+q+1)}{2gl} \right] (1 - GFI)$$

Índice de Bondad de Ajuste Ajustado Relativo (RAGFI) debido a que el $AGFI$ tiene influencia debido a la muestra el cálculo de este indicador ayuda a tener una mejor perspectiva del ajuste del modelo. Para calcular $EAGFI$ se usa la siguiente ecuación:

$$EAGFI = 1 - \left[\frac{(p+q)(p+q+1)}{2gl} \right] (1 - EGFI)$$

Para luego calcular $RAGFI$ así:

$$RAGFI = \frac{AGFI}{EGFI}$$

Índice de Bondad de Ajuste Normado (NFI) , este índice compara el ajuste entre el modelo propuesto y el modelo nulo(donde no se consideran relaciones). Mide la variación proporcional del ajuste al pasar del modelo nulo hacia el modelo propuesto. Se calcula así:

$$NFI = \frac{(\chi_{M.Nulo}^2 - \chi_{M.propuesto}^2)}{\chi_{M.Nulo}^2}$$

Valor cercanos a 1 del índice *NFI* muestran buen ajuste en el modelo.

Índice de Ajuste Normado Parsimonioso (PNFI) es similar al anterior, en este caso se consideran los grados de libertad. El nivel adecuado de parsimonia sería 1 grado de libertad por coeficiente a ser estimado, entonces, lo lógico es buscar mayores valores de parsimonia en los modelos. Se utiliza la siguiente ecuación:

$$PNFI = \left(\frac{gl_{M.Propuesto}}{gl_{M.Nulo}} \right) \times NFI$$

Índice de Bondad de Ajuste Normado Parsimonioso (PGFI) se busca ajustar el *GFI* de forma similar al *AGFI* considerando los grados de libertad, es decir, la parsimonia del modelo. Se prefieren valores altos.

$$PGFI = \frac{2 gl_{M.Propuesto}}{(p + q) (p + q + 1)}$$

Ji-cuadrado Normado es el cociente entre el valor ji-cuadrado sobre los grados de libertad, esta medida permite la comparación para modelo que tienen grados de libertad mayores o iguales a 0 .

Criterio de Información de Akaike (AIC) es una de las medias utilizadas frecuentemente para la comparación de modelo que poseen diferente número de variables latentes. Tenemos:

$$AIC = \chi^2 + 2t$$

A medida que AIC se acerca a cero, indica que el modelo tiene un alta parsimonia.

N-Crítico este valor indica una cantidad de casos a ser consideradas para el modelo para lograr aceptar el ajuste del modelo, mientras más grande sea la muestra mejores serán los ajustes logrados.

$$CN = \frac{\chi_{1-\alpha}^2}{\chi_{M.Propuesto}^2} + 1$$

Esta medida debe ser considerada para tamaños de muestra mayores a 200, a partir del cual los modelos de ecuaciones estructurales aseguran convergencia.

2. **Ajuste del Modelo de Medida y Estructural** Las medidas consideradas anteriormente sirven para evaluar el ajuste del modelo en conjunto (parte estructural y parte de medida), ahora se consideran medidas que permitan concluir acerca del nivel de ajuste del modelo de medida y el modelo estructural independientemente. Este análisis debe hacerse por separado para cada una de las partes del modelo de ecuaciones estructurales.

Significación estadística esta medida tiene el mismo papel que en la regresión múltiple, en el cual se realiza la prueba de hipótesis $H_0 : \xi_i = 0$, $H_1 : \xi_i \neq 0$ para verificar que los coeficientes sean estadísticamente distintos de cero, por lo tanto, se evidencia si un indicador no aporta para predecir el comportamiento de

una variable latente .

Fiabilidad mide la proporción de la varianza que tiene en común un indicador con una variable latente, un indicador deberá tener al menos el 50 % de su varianza en común con la variable latente a la cual busca predecir. La fiabilidad de un indicador se calcula utilizando la siguiente ecuación:

$$Fiabilidad = \frac{(\sum \text{Cargas factoriales estandarizadas})^2}{(\sum \text{Cargas factoriales estandarizadas})^2 + (\sum \text{Errores de medida})}$$

Varianza Extraída es una medida que indica la cantidad de la varianza de los indicadores explicada por una variable latente. SI el valor es grande esto indica que la variable latente ha logrado explicar una gran cantidad de varianza contenida en los indicadores. Se usa la siguiente ecuación:

$$Varianza Extraida = \frac{\sum \text{Cargas Estandarizadas}^2}{\sum \text{Cargas Estandarizadas}^2 + \sum \text{Errores de medida}}$$

ANEXO 4

Cálculo del tamaño muestral

El objetivo es calcular un tamaño muestral que asegure obtener resultados representativos de un segmento del mercado de las comunicaciones, para ello se ha realizado un muestreo aleatorio irrestricto considerando un tamaño de la población de $N = 165000$, este es el tamaño del mercado al cual este medio de comunicación está dirigido, mediante registros administrativos se ha estimado la venta de $n = 121000$ unidades promedio, esta es la población penetrada¹, con estos datos se ha calculado la proporción del mercado que usa este medio de comunicación:

$$\begin{aligned}\hat{p} &= \frac{121000}{165000} \\ &= 0,7333\end{aligned}$$

Además, se ha fijado un nivel de confianza del $\gamma = 95\% = 1 - \alpha$, esto como un estándar al momento de realizar estimaciones, además se desea que la precisión sea de $B = 0,05$.

Inicialmente como no se tiene el tamaño muestral

$$\begin{aligned}B &= \sqrt{\chi_{1,\alpha}^2 \cdot \frac{(N-n)}{(N-1)} \cdot \frac{p \cdot (1-p)}{n}} \\ \frac{B^2}{\chi_{1,\alpha}^2} \cdot \frac{(N-1)}{p(1-p)} &= \frac{N-n}{n} \\ n \left(\frac{B^2}{\chi_{1,\alpha}^2} \cdot \frac{(N-1)}{p(1-p)} + 1 \right) &= N \\ n &= \frac{N}{\frac{B^2}{\chi_{1,\alpha}^2} \cdot \frac{(N-1)}{p(1-p)} + 1} \\ n &= \frac{165000}{\frac{0,05^2}{3,841} \cdot \frac{(165000-1)}{0,7333(1-0,7333)} + 1} \\ n &= 300\end{aligned}$$

Se puede notar que el tamaño muestral óptimo para que se cumplan las condiciones establecidas (nivel de confianza, precisión) es 300 casos, este número de casos debe ser reducido

¹Población penetrada: son todos los usuarios actuales del medio de comunicación

debido al alto costo de las encuestas, por lo tanto se ha decidido relajar la precisión para obtener un menor número de casos que muestrear. Se ha considerado un valor de $B = 0,06$ para la precisión con este cambio el tamaño muestral es el siguiente:

$$n = \frac{165000}{\frac{0,06^2}{3,841} \cdot \frac{(165000 - 1)}{0,7333(1 - 0,7333)} + 1}$$

$$n = 209$$

El tamaño muestral se ha reducido en 91 casos, por cuestiones prácticas se ha redondeado el tamaño muestral a 210 casos.

Cuestionario

Cuadro 5.1: Cuestionario: Satisfacción del cliente

Pregunta	Variable	Escala
Calidad esperada		
¿Cómo calificaría la credibilidad?	X1	1-7
¿Cómo calificaría la cantidad de información?	X2	1-7
¿Cómo calificaría el precio?	X3	1-7
Calidad percibida		
¿Qué tanto le gusta el diario?	X4	1-7
¿Qué tanto le incita a comprar este diario?	X5	1-3
¿Qué tan cercano diría usted que está el diario a sus gastos y necesidades?	X6	1-4
Valor percibido		
Es mi diario y nunca dejará de serlo	X7	1-7
Es un diario que siempre está cambiando	X8	1-7
Está muy caro para mi presupuesto	X9	1-7
Los diarios gratuitos ofrecen lo mismo	X10	1-7
El diario ha cambiado para mal	X11	1-7

Cuadro 5.2: Cuestionario: Satisfacción del cliente (Parte 2)

Pregunta	Variable	Escala
Satisfacción del cliente		
Diría usted que lee más días, lee igual número o lee menos que antes	X12	1-3
Caparando con un año atrás, compra más, igual o menos que antes	X13	1-3

Modificación de la estructura del cuestionario después del Análisis Factorial Exploratorio (AFE)

Cuadro 5.3: Cambio de nomenclatura

Pregunta	Antes del AFE	Después del AFE
Calidad esperada		
¿Cómo calificaría la credibilidad?	X1	X1
¿Cómo calificaría la cantidad de información?	X2	X2
¿Cómo calificaría el precio?	X3	X3
Calidad percibida		
¿Qué tanto le gusta el diario?	X4	X4
¿Qué tan cercano diría usted que está el diario a sus gastos y necesidades?	X6	X5
Valor percibido		
Es mi diario y nunca dejará de serlo	X7	X6
Es un diario que siempre está cambiando	X8	X7
Satisfacción del cliente		
Está muy caro para mi presupuesto	X9	X8
Diría usted que lee más días, lee igual número o lee menos que antes	X12	X9
Caparando con un año atrás, compra más, igual o menos que antes	X13	X10

ANEXO 5

Comparación de histogramas de las estimaciones bayesianas en el Análisis Factorial Exploratorio

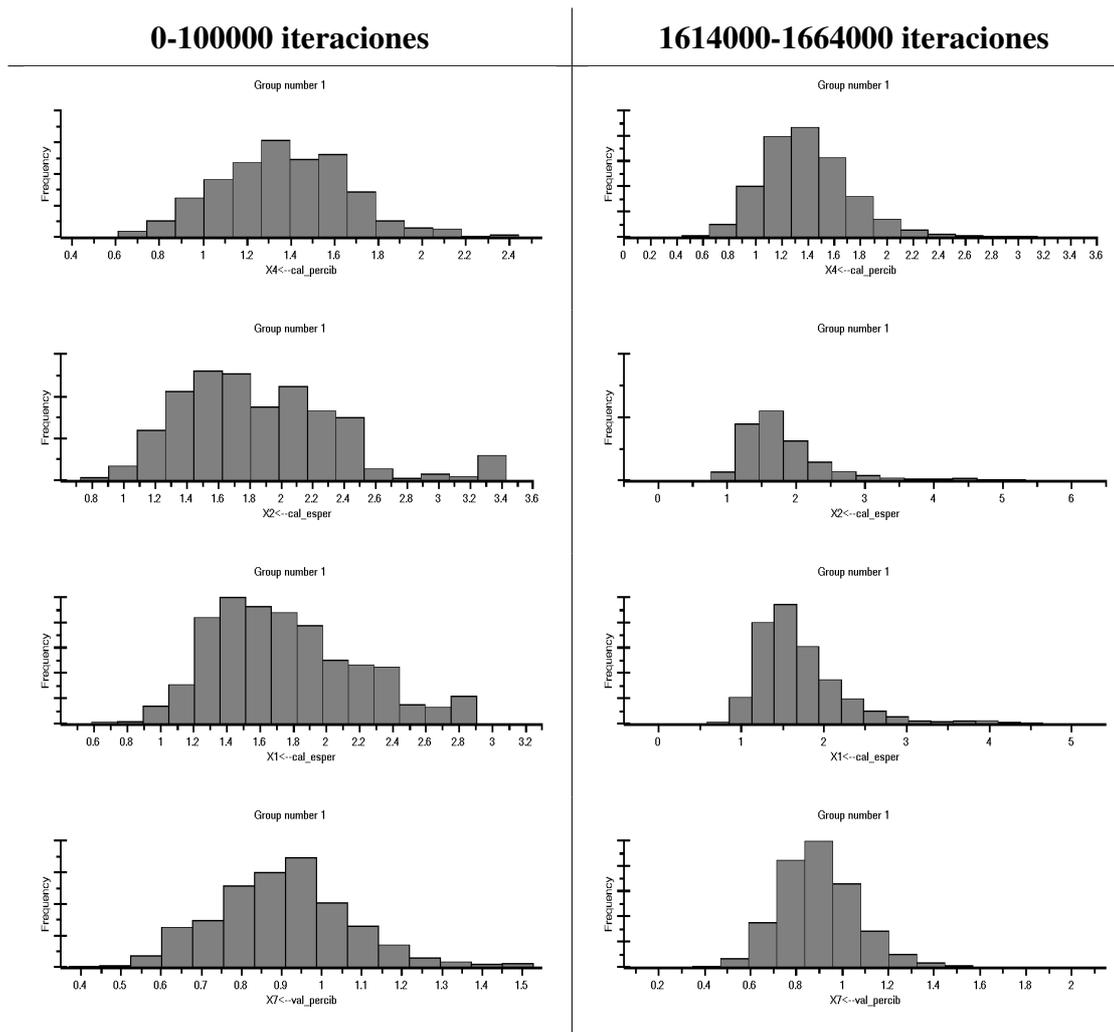


Figura 5.1: Comparación de histogramas de estimaciones bayesianas en el análisis factorial exploratorio

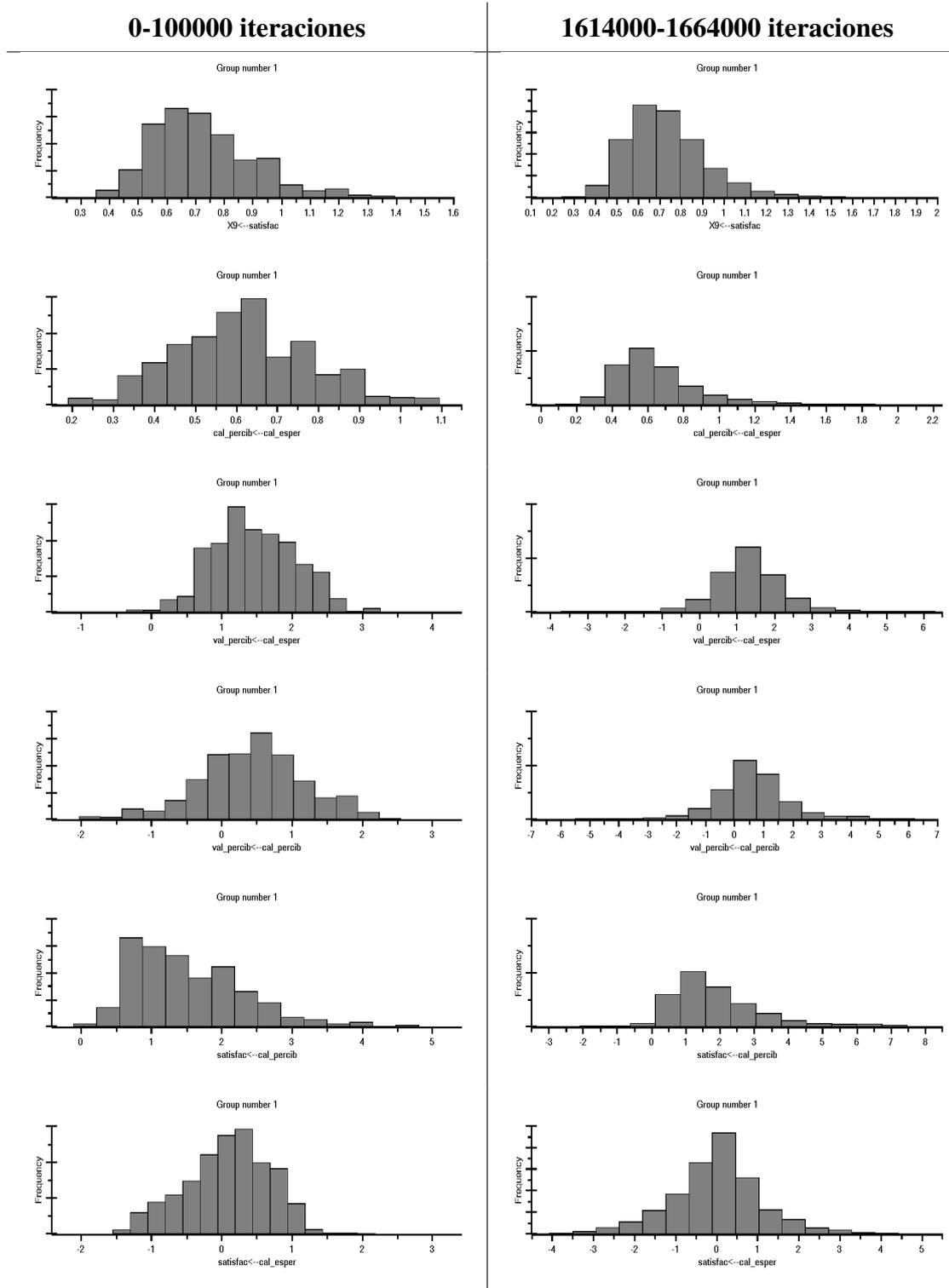


Figura 5.2: Comparación de histogramas de estimaciones bayesianas en el análisis factorial exploratorio (continuación)

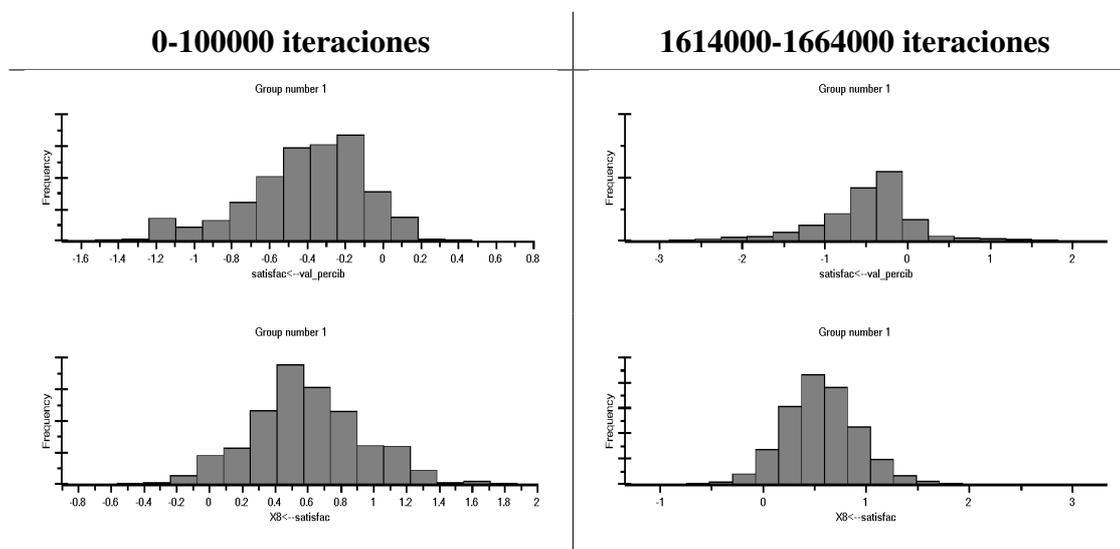


Figura 5.3: Comparación de histogramas de estimaciones bayesianas en el análisis factorial exploratorio (continuación)

La comparación de los histogramas para cada una de las estimaciones apoya la interpretación de reducción de la varianza, y por ende, muestra mayor concentración de los valores de las estimaciones bayesianas alrededor de sus medias, además se puede notar que en algunas de las estimaciones inicialmente se encontraba sesgada, en algunos histogramas se evidencia una reducción drástica del sesgo (por ejemplo $X_1 < -cal_esp$), en otras todavía se puede apreciar la presencia de sesgo (por ejemplo $X_9 < -satisfac$).

Comparación de histogramas de las estimaciones bayesianas en el Análisis Factorial Confirmatorio

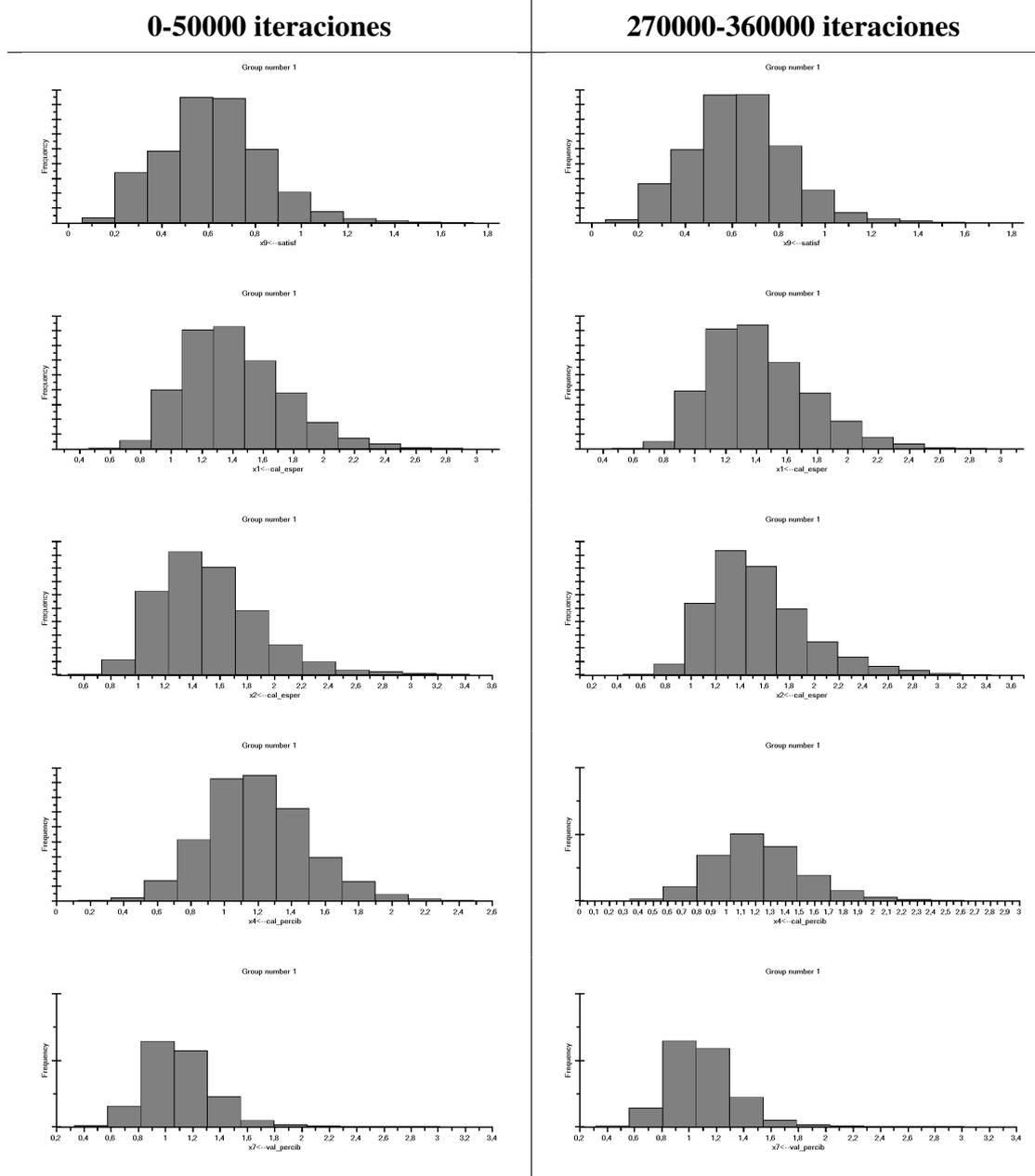


Figura 5.4: Comparación de histogramas de estimaciones bayesianas en el Análisis Factorial Confirmatorio

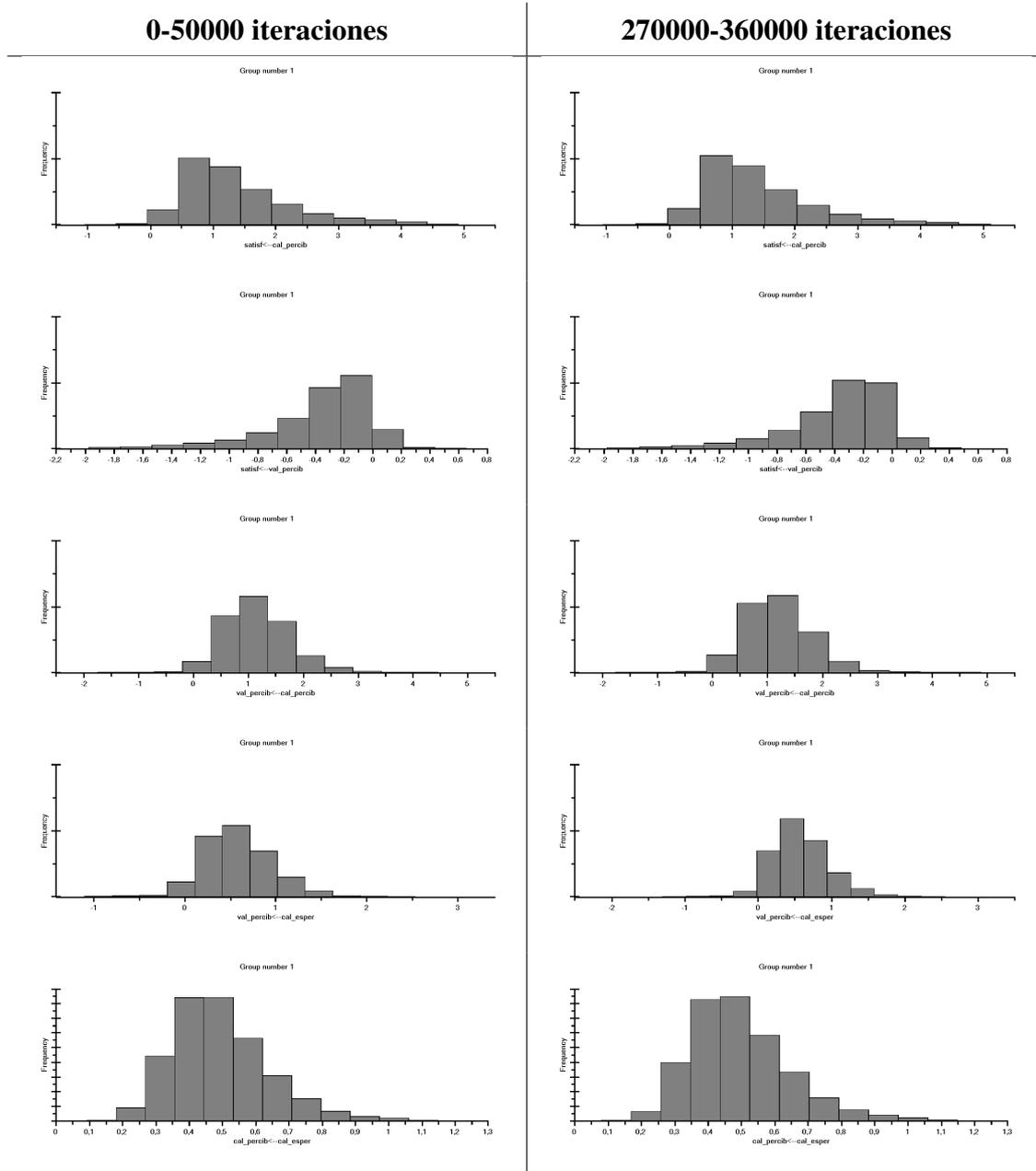


Figura 5.5: Comparación de histogramas de estimaciones bayesianas en el Análisis Factorial Confirmatorio (continuación)

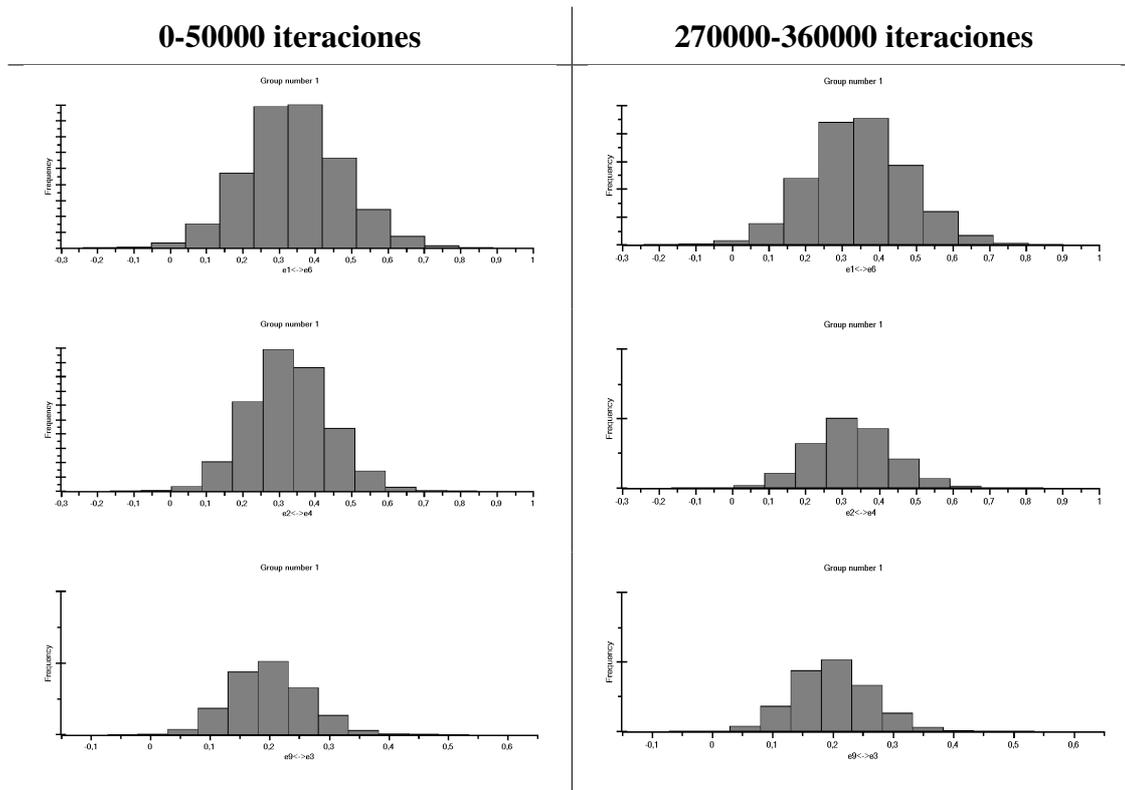


Figura 5.6: Comparación de histogramas de estimaciones bayesianas en el Análisis Factorial Confirmatorio (continuación)