

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO



“MODELOS ALTERNATIVOS DE RESPUESTA  
GRADUADA CON APLICACIONES  
EN LA CALIDAD DE SERVICIOS”

TESIS PARA OPTAR POR EL GRADO DE MAGISTER EN  
ESTADÍSTICA

AUTOR

Enver Gerald Tarazona Vargas

ASESOR

Dr. Jorge Luis Bazán Guzmán

JURADO

Dr. Luis Hilmar Valdivieso Serrano

Dr. Cristian Luis Bayes Rodríguez

LIMA-PERÚ

2013

## Dedicatoria

A los que ya partieron: a mi padre Pablo, mi abuelo Virgilio y mi tío Erick, quienes cultivaron en mí desde pequeño el cariño hacia los libros.

A mi madre Gladis y mis hermanos Henry y Elard, por su apoyo constante.

### *Primera Canción*

No he inventado ninguna melodía.  
Los que amaron dirán:  
“Conozco esta canción...  
y me había olvidado de lo hermosa que era...”  
Y habrá de parecerles  
la primera  
canción con que soñaron.

*Juan Gonzalo Rose, 1960*

### *Escrito a ciegas*

¿Quieres tú saber de mi vida?  
Yo sólo sé de mi paso,  
De mi peso,  
De mi tristeza y de mi zapato.  
¿Por qué preguntas quién soy,  
Adónde voy?... Porque sabes harto  
Lo del Poeta, el duro  
Y sensible volumen de ser mi humano,  
Que es un cuerpo y vocación,  
Sin embargo.  
Si nací, lo recuerda el Año  
Aquel de quien no me acuerdo,  
Porque vivo, porque me mato. ...

*Martín Adan, 1961*

*Tabacaria*

Não sou nada.

Nunca serei nada.

Não posso querer ser nada.

Janelas do meu quarto,

Do meu quarto de um dos milhões do mundo que ninguém sabe quem é  
(E se soubessem quem é, o que saberiam?),

Dais para o mistério de uma rua cruzada constantemente por gente,

Para uma rua inacessível a todos os pensamentos,

Real, impossivelmente real, certa, desconhecidamente certa,

Com o mistério das coisas por baixo das pedras e dos seres,

Com a morte a por umidade nas paredes e cabelos brancos nos homens,

Com o Destino a conduzir a carroça de tudo pela estrada de nada.

Estou hoje vencido, como se soubesse a verdade.

Estou hoje lúcido, como se estivesse para morrer,

E não tivesse mais irmandade com as coisas

Senão uma despedida, tornando-se esta casa e este lado da rua

A fileira de carruagens de um comboio, e uma partida apitada

De dentro da minha cabeça,

E uma sacudidela dos meus nervos e um ranger de ossos na ida.

Estou hoje perplexo, como quem pensou e achou e esqueceu.

Estou hoje dividido entre a lealdade que devo

À Tabacaria do outro lado da rua, como coisa real por fora,

E à sensação de que tudo é sonho, como coisa real por dentro.

Falhei em tudo.

Como não fiz propósito nenhum, talvez tudo fosse nada.

A aprendizagem que me deram,

Desci dela pela janela das traseiras da casa.

Fui até ao campo com grandes propósitos.

Mas lá encontrei só ervas e árvores,

E quando havia gente era igual à outra.

Saio da janela, sento-me numa cadeira. Em que hei de pensar?

...

*Álvaro de Campos (Fernando Pessoa), 1928*

## Agradecimientos

A Jorge Bazán, por su orientación, consejos y paciencia sin las cuales este trabajo no hubiera sido posible, y sobre todo por animarme a seguir el camino de la Estadística del cual siempre suelo alejarme.

A los profesores de la Maestría en Estadística de la Pontificia Universidad Católica del Perú (PUCP) por sus recomendaciones de mejora y por lo aprendido durante estos dos años.

Al Vicerrectorado de Investigación (VRI) de la Pontificia Universidad Católica del Perú (PUCP) por el financiamiento otorgado a través del premio PAIP 2012 para que el presente proyecto pueda desarrollarse.

A Telefónica Móviles S.A. por el apoyo a la presente investigación durante el Segundo Congreso Brasileiro de Teoria de Resposta ao Item (II CONBRATRI) realizado en Salvador (Bahia) el 2011 y la II Jornada Internacional de Probabilidad y Estadística (JIPE II) llevada a cabo en la PUCP el 2012.

A Jose Cruzado, Director de la Dirección de Calidad, Procesos y Lealtad, y Patricia Vega, Jefa de Planificación y Mejores Practicas en Telefónica Móviles S.A., por su apoyo en darme el tiempo necesario para culminar este trabajo.

A mis compañeros y amigos de la Dirección de Calidad, Procesos y Lealtad de Telefónica Móviles S. A. por compartir conmigo su experiencia y conocimientos sobre el estudio de la satisfacción de clientes.

A Anya Neverova por sus correcciones para la elaboración del Abstract.

## Resumen

Los modelos politómicos de la Teoría de Respuesta al Ítem (TRIP) tienen como finalidad explicar la interacción existente entre los sujetos evaluados y los atributos de un test en aquellas situaciones en las cuales los atributos que lo componen tienen varias categorías de respuesta. Dentro de los distintos tipos de modelos TRIP, el Modelo de Respuesta Graduada General (GRM) propuesto originalmente por Samejima (1969, 2010), es un conjunto de modelos diseñados para aplicarse en aquellas situaciones en las cuales las categorías de respuesta son ordinales.

En este trabajo se presenta una formulación general para los GRM, su clasificación y principales propiedades desde el punto de vista bayesiano. De manera específica, se muestra el Modelo de Respuesta Graduada Logístico de dos parámetros (2PL-GRM) como un caso particular de los GRM simétricos y el Modelo de Respuesta Graduada Logístico de Exponente Positivo (LPE-GRM) como un modelo asimétrico derivado de incorporar un parámetro de penalización que controla la curvatura de las Funciones de Respuesta a las Etapas de los Ítems (FREI). La estimación de ambos modelos fue realizada usando la inferencia bayesiana con Métodos Montecarlo vía Cadenas de Markov (MCMC) e implementada en R y WinBUGS.

Se realizó un estudio de simulación con el fin de estudiar la precisión en la recuperación de parámetros para el Modelo 2PL-GRM obteniéndose resultados apropiados para las medidas de ajuste consideradas.

Los modelos 2PL-GRM y LPE-GRM estudiados fueron aplicados al estudio de un cuestionario acerca de la satisfacción de clientes y comparados con el tradicional análisis clásico de los test. La muestra del estudio está formada por 5354 clientes de una empresa de telecomunicaciones que se comunicaron con el Call Center de atención al cliente por algún motivo (consulta, reclamo, pedido, etc.). A través del análisis de dimensionalidad de la escala se encontró que el cuestionario evalúa dos dimensiones de la satisfacción con la atención al cliente: la Accesibilidad (4 ítems) y el Desempeño del asesor (7 ítems). Los resultados indican, considerando diferentes criterios, que en ambas dimensiones el modelo LPE-GRM es mejor. Adicionalmente, ambos modelos ofrecen mejor información que el tradicional análisis clásico.

Se sugiere realizar diferentes estudios de simulación para evaluar distintas condiciones para la inferencia del modelo LPE-GRM puesto que para las mismas condiciones de estimación MCMC se observa que puede ser más demorado debido a que presenta mayor autocorrelación que el modelo 2PL-GRM.

**Palabras clave:** teoría de respuesta al ítem (TRI), modelos politómicos de la teoría de respuesta al ítem (PTRI), variables latentes, inferencia bayesiana, modelo de respuesta graduada, Métodos Montecarlo vía Cadenas de Markov (MCMC), asimetría, calidad de servicios, satisfacción de clientes.

## Abstract

Polytomous item response theory (PIRT) models explain the interaction between examinees and test questions where the test questions have various response categories. Within the different types of PIRT models, the originally proposed by Samejima (1969, 2010) General Graded Response Model (GRM) is a general framework of models to be used in the situations in which the response categories are ordinal.

This work shows a general formulation for the GRM, its classification and main properties from Bayesian perspective. Specifically, the two-parameter logistic graded response model (2PL-GRM) is presented as a particular case of symmetric GRM and a logistic positive exponent graded response model (LPE-GRM) as an asymmetric model derived from incorporation of a penalty parameter that controls the curve of the boundary response functions (CBRF). The estimation of both models was performed using the Bayesian inference with Markov chain Monte Carlo (MCMC) methods and implemented in R and WinBUGS.

A simulation study is performed to evaluate the accuracy of parameter recovery for 2PL-GRM Model, obtaining adequate results for the considered adjusting measures.

The 2PL-GRM and LPE-GRM models were applied to a customer satisfaction questionnaire study and compared with the classical test analysis. The study sampling consists of 5354 customers from a Telco who have called to the call Center for some reason (query, complaint, order, etc.). Scale dimensionality analysis reveals that the questionnaire evaluates two dimensions of customer service satisfaction: Accessibility (4 items) and Agent Performance (7 items). Considering different criteria, the results indicate that in both dimensions LPE-GRM model is better. Additionally, both models offer better information than the traditional classical analysis.

It is suggested performing various simulation studies to evaluate different conditions for LPE-GRM model's inference because for the same MCMC estimation conditions it is observed that it presents greater autocorrelation than 2PL-GRM model.

**Keywords:** item response theory (IRT), polytomous item response theory models (PIRT), latent variables, Bayesian inference, graded response model, Markov chain Monte Carlo (MCMC), skewness, service quality, customer satisfaction.

# Índice general

Lista de Abreviaturas	X
Lista de Símbolos	XI
Índice de figuras	XII
Índice de cuadros	XIV
<b>1. Introducción</b>	<b>1</b>
1.1. Consideraciones Preliminares	1
1.2. Objetivos	2
1.3. Organización del Trabajo	3
<b>2. El Modelo de Respuesta Graduada de la TRIP</b>	<b>4</b>
2.1. Formulación del Modelo de Respuesta Graduada General	4
2.1.1. Presentación y Racionalidad del Modelo	4
2.1.2. Observaciones	5
2.1.3. El Modelo de Respuesta Graduada como extensión de un modelo binario	6
2.1.4. El Modelo de Respuesta Graduada usando Aumentación de Datos	7
2.2. Clasificación de los Modelos de Respuesta Graduada	8
2.2.1. Modelos de Respuesta Graduada Simétricos	8
2.2.2. Modelos de Respuesta Graduada Asimétricos	8
2.3. Supuestos de los Modelos de Respuesta Graduada	9
2.4. Características de los Modelos de Respuesta Graduada	10
2.4.1. Función de Procesamiento (FP)	10
2.4.2. Función de Respuesta a las Etapas del Ítem (FREI)	11
2.4.3. Función de Respuesta a las Categorías del Ítem (FRCI)	11
2.4.4. Función Básica (FB)	12
2.4.5. Función de Información de las Respuestas de los Ítems (FIRI)	12
2.4.6. Condición de un Máximo Único	12
2.4.7. Sobreparametrización	13
2.4.8. Presencia de Parámetros Incidentales	13
2.4.9. Identificabilidad	13
2.5. Información de Fisher en los Modelos de Respuesta Graduada	13
2.6. Casos Particulares	16
2.6.1. Modelo de Respuesta Graduada de un Parámetro (1P-GRM)	16

2.6.2. Modelo Rating Scale de Muraki (RS-GRM) . . . . .	16
2.6.3. Modelo de Respuesta Graduada Multidimensional . . . . .	16
2.6.4. Modelo Testlet de Respuesta Graduada . . . . .	17
2.7. Criterios para la selección del Modelo de Respuesta Graduada . . . . .	17
<b>3. Modelo de Respuesta Graduada Logístico y LPE</b>	<b>18</b>
3.1. Modelo de Respuesta Graduada Logístico de dos Parámetros (2PL-GRM) . . . . .	18
3.1.1. Definición del Modelo . . . . .	18
3.1.2. Características del Modelo 2PL-GRM . . . . .	19
3.1.3. Información de Fisher . . . . .	21
3.2. Modelo de Respuesta Graduada Logístico de Exponente Positivo (LPE-GRM)	23
3.2.1. Definición del Modelo . . . . .	23
3.2.2. Características del Modelo LPE-GRM . . . . .	24
3.2.3. Información de Fisher . . . . .	27
<b>4. Estimación Bayesiana del Modelo de Respuesta Graduada</b>	<b>30</b>
4.1. Estimación Bayesiana del Modelo de Respuesta Graduada Logístico (2PL-GRM)	30
4.1.1. Función de Verosimilitud para el Modelo 2PL-GRM . . . . .	30
4.1.2. Especificación de las prioris en el modelo 2PL-GRM . . . . .	31
4.1.3. Distribución a posteriori de los parámetros del modelo 2PL-GRM . . . . .	32
4.1.4. Inferencia Bayesiana del modelo 2PL-GRM . . . . .	32
4.2. Estimación Bayesiana del Modelo de Respuesta Graduada Logístico con Ex-	
ponente Positivo (LPE-GRM) . . . . .	33
4.2.1. Función de Verosimilitud para el modelo LPE-GRM . . . . .	33
4.2.2. Especificación de las prioris en el modelo LPE-GRM . . . . .	33
4.2.3. Distribución a posteriori de los parámetros del modelo LPE-GRM . . . . .	34
4.2.4. Inferencia Bayesiana del modelo LPE-GRM . . . . .	34
4.3. Criterios de Comparación de Modelos . . . . .	34
4.4. Implementación Computacional . . . . .	35
<b>5. Estudio de Simulación</b>	<b>37</b>
5.1. Objetivos . . . . .	37
5.2. Algoritmo para simular los datos . . . . .	37
5.3. Método para Estimar los parámetros . . . . .	38
5.4. Criterios para evaluar la simulación . . . . .	38
5.5. Escenario . . . . .	38
5.6. Resultados . . . . .	39
<b>6. Aplicación</b>	<b>43</b>
6.1. Introducción . . . . .	43
6.2. Instrumento de Medición . . . . .	45
6.3. Diseño Muestral . . . . .	46
6.4. Recolección de Datos . . . . .	46
6.5. Análisis del instrumento según el enfoque de la Teoría Clásica de los Test (TCT)	46



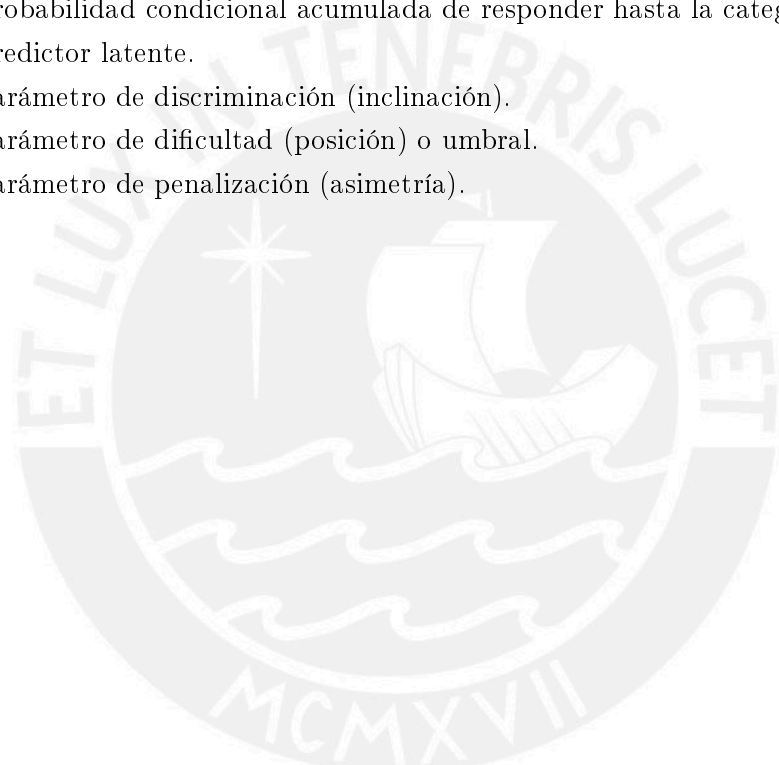
6.5.1. Validez del Instrumento . . . . .	46
6.5.2. Fiabilidad del Instrumento . . . . .	47
6.5.3. Análisis de los Ítems . . . . .	48
6.6. Estimación de los parámetros mediante los métodos desarrollados . . . . .	49
6.6.1. Modelo de Respuesta Graduada Logístico (2PL-GRM) . . . . .	49
6.6.2. Modelo de Respuesta Graduada Logístico de Exponente Positivo (LPE-GRM) . . . . .	52
6.7. Comparación de Modelos . . . . .	56
6.8. Funciones de Información . . . . .	57
6.9. Características finales de la Escala . . . . .	57
<b>7. Conclusiones</b>	<b>60</b>
7.1. Conclusiones . . . . .	60
7.2. Sugerencias para investigaciones futuras . . . . .	61
<b>A. Códigos en R y BUGS</b>	<b>62</b>
A.1. Código en R para simular datos aleatorios del GRM . . . . .	62
A.2. Función en R para realizar la estimación bayesiana usando MCMC . . . . .	65
A.3. Código en BUGS para el Modelo de Respuesta Graduada Logístico (2PL-GRM) . . . . .	67
A.4. Código en BUGS para el Modelo de Respuesta Graduada con Exponente Positivo (LPE-GRM) . . . . .	69
A.5. Código en R para el estudio de simulación . . . . .	70
A.6. Código en R para la aplicación en el análisis de satisfacción de clientes . . . . .	75
A.7. Código en R para graficar la Función de Respuesta a las Etapas de los Ítems (FREI) del GRM . . . . .	78
A.8. Código en R para graficar la Función de Respuesta a las Categorías de los Ítems (FRCI) del GRM . . . . .	79
A.9. Código en R para graficar la Función de Información de un Ítem (FII) del GRM . . . . .	81
<b>B. Diagnósticos de convergencia</b>	<b>84</b>
B.1. Diagnóstico de convergencia en la simulación . . . . .	84
B.2. Diagnóstico de convergencia en la aplicación . . . . .	85
<b>C. Cuestionario de Satisfacción: Atributos medidos</b>	<b>88</b>
<b>Bibliografía</b>	<b>89</b>

## Lista de Abreviaturas

1P-GRM	Modelo de Respuesta Graduada de 1 parámetro.
2PL-GRM	Modelo de Respuesta Graduada Logístico de 2 parámetros.
AIC	<i>Akaike Information Criterion</i>
BIC	<i>Bayesian Information Criterion</i>
CCI	Curva Característica del Ítem. Similar a FRI.
CML	Máxima Verosimilitud Condicional ( <i>Conditional Maximum Likelihood</i> )
DIC	Deviance Information Criterion
FB	Función Básica.
FIRI	Función de Información de las Respuestas del Ítem.
FII	Función de Información del Ítem.
FIT	Función de Información del Test.
FP	Función de Procesamiento.
FRCI	Función de Respuesta a las Categorías del Ítem.
FREI	Función de Respuesta a las Etapas del Ítem.
FRI	Función de Respuesta del Ítem.
fda	Función de densidad acumulada.
GRM	Modelo de Respuesta Graduada o Modelo de Respuesta Graduada General
JML	Máxima Verosimilitud Conjunta ( <i>Joint Maximum Likelihood</i> )
LPE-GRM	Modelo de Respuesta Graduada Logístico de Exponente Positivo.
RS-GRM	Modelo Rating Scale de Muraki.
MCMC	Métodos Montecarlo vía Cadenas de Markov
MML	Máxima Verosimilitud Marginal ( <i>Marginal Maximum Likelihood</i> )
TCT	Teoría Clásica de los Test.
TRI	Teoría de Respuesta al Ítem.
TRIP	Teoría de Respuesta al Ítem Politémica.

## Lista de Símbolos

$Y_{ij}$	Variable manifiesta que modela la respuesta del individuo $i$ en el ítem $j$ en un test.
$u_i$	Valor de la variable latente $U_i$ (habilidad o rasgo latente) para el sujeto $i$ .
$\xi_j$	Vector de parámetros del ítem $j$ .
$P_{ijk}$	Probabilidad condicional de responder la categoría $k$ en el ítem $j$ .
$P_{ijk}^{ac}$	Probabilidad condicional acumulada de responder hasta la categoría $k$ en el ítem $j$ .
$\eta_{ijk}$	Predictor latente.
$a_j$	Parámetro de discriminación (inclinación).
$b_{jk}$	Parámetro de dificultad (posición) o umbral.
$d_j$	Parámetro de penalización (asimetría).



## Índice de figuras

2.1. Representación gráfica tomada de Molenaar et al. (2012) que muestra las causas de asimetría en los Modelos de Respuesta Graduada: (a) Modelo Tradicional (b) $\sigma_{\varepsilon_{ij}}^2$ heterocedástico (c) $U_i$ asimétrico (d) $Z_{ij}   u_i$ asimétrico. Para mayor claridad el impacto del efecto es exagerado. . . . .	9
3.1. Representación gráfica de las FREI del Modelo 2PL-GRM para tres ítems de 5 categorías con parámetros de dificultad $b = (-1.5, -0.5, 0.5, 1.5)$ y diferentes parámetros de discriminación $a_j = 0.8, 1.6$ y $2.4$ . . . . .	20
3.2. Representación gráfica de las FRCI del Modelo 2PL-GRM para seis ítems de 5 categorías con dos parámetros de discriminación $a_j = 0.8$ y $2.4$ , y tres configuraciones de parámetros de dificultad reflejando diferentes niveles: (i) $b_{jk} = -3.0, -1.5, -0.5$ y $1.0$ (ii) $b_{jk} = -1.5, -0.5, 0.5$ y $1.5$ y (iii) $b_{jk} = -1.0, 0.5, 1.5$ y $3.0$ . . . . .	21
3.3. Representación gráfica de las FII del Modelo 2PL-GRM para seis ítems de 5 categorías con dos parámetros de discriminación $a_j = 0.8$ (ítems impares) y $2.4$ (ítems pares), y tres configuraciones de parámetros de dificultad reflejando diferentes niveles: $b_{jk} = -3.0, -1.5, -0.5$ y $1.0$ (ítems 1 y 2), $b_{jk} = -1.5, -0.5, 0.5$ y $1.5$ (ítems 3 y 4) y $b_{jk} = -1.0, 0.5, 1.5$ y $3.0$ (ítems 5 y 6) . . . . .	23
3.4. Representación gráfica de las FREI del Modelo LPE-GRM para seis ítems de 5 categorías con un parámetro de discriminación $a_j = 2.4$ , dos configuraciones de parámetros de dificultad reflejando ítems con baja y alta dificultad: $b_{jk} = -3.0, -1.5, -0.5$ y $1.0$ (ítems impares) y $b_{jk} = -1.0, 0.5, 1.5$ y $3.0$ (ítems pares), y distintos valores de asimetría $d_j = 0.3$ (ítems 1 y 2), $1.0$ (ítems 3 y 4) y $8.0$ (ítems 5 y 6) . . . . .	26
3.5. Representación gráfica de las FRCI del Modelo LPE-GRM para seis ítems de 5 categorías con un parámetro de discriminación $a_j = 2.4$ , dos configuraciones de parámetros de dificultad reflejando ítems con baja y alta dificultad: $b_{jk} = -3.0, -1.5, -0.5$ y $1.0$ (ítems impares) y $b_{jk} = -1.0, 0.5, 1.5$ y $3.0$ (ítems pares), y distintos valores de asimetría $d_j = 0.3$ (ítems 1 y 2), $1.0$ (ítems 3 y 4) y $8.0$ (ítems 5 y 6) . . . . .	27
3.6. Representación gráfica de las FII del Modelo LPE-GRM para seis ítems de 5 categorías con un parámetro de discriminación $a_j = 2.4$ , dos configuraciones de parámetros de dificultad reflejando ítems con baja y alta dificultad: $b_{jk} = -3.0, -1.5, -0.5$ y $1.0$ (ítems impares) y $b_{jk} = -1.0, 0.5, 1.5$ y $3.0$ (ítems pares), y distintos valores de asimetría $d_j = 0.3$ (ítems 1 y 2), $1.0$ (ítems 3 y 4) y $8.0$ (ítems 5 y 6) . . . . .	28

6.1. Distribución de las puntuaciones totales para cada dimensión . . . . .	48
6.2. FREI y FRCI del Modelo 2PL-GRM para los atributos que permiten diferenciar mejor entre los individuos más insatisfechos . . . . .	52
6.3. FREI y FRCI del Modelo 2PL-GRM para los atributos que permiten diferenciar mejor entre individuos más satisfechos . . . . .	53
6.4. FREI y FRCI del Modelo LPE-GRM para los atributos que permiten diferenciar mejor entre individuos más insatisfechos y satisfechos . . . . .	54
6.5. Diagramas de dispersión de las habilidades (satisfacción con la Accesibilidad y Desempeño del Asesor) estimadas con los modelos 2PL-GRM y LPE-GRM . . . . .	58
6.6. Funciones de Información de los Ítems (FII) obtenidas con los Modelos 2PL-GRM y LPE-GRM para la satisfacción con la accesibilidad y el desempeño del asesor . . . . .	59
6.7. Distribución de los puntajes de las escalas de satisfacción con la accesibilidad y el desempeño del asesor usando el Modelo LPE-GRM . . . . .	59
B.1. Histórico de dos cadena para los parámetros del Ítem 4 . . . . .	84
B.2. Gráficas de las densidades y autocorrelaciones para los parámetros del Ítem 4 . . . . .	84
B.3. Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el parámetro de discriminación del Ítem 9, tiempo de atención, estimado con el Modelo 2PL-GRM . . . . .	85
B.4. Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el segundo umbral (parámetro de dificultad) del Ítem 9, tiempo de atención, estimado con el Modelo 2PL-GRM . . . . .	85
B.5. Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el parámetro de discriminación del Ítem 9, tiempo de atención, estimado con el Modelo LPE-GRM . . . . .	86
B.6. Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el segundo umbral (parámetro de dificultad) del Ítem 9, tiempo de atención, estimado con el Modelo LPE-GRM . . . . .	86
B.7. Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el parámetro de penalización del Ítem 9, tiempo de atención, estimado con el Modelo LPE-GRM . . . . .	87

## Índice de cuadros

4.1. Especificación de prioris para los parámetros del ítem en el modelo 2PL-GRM	32
5.1. Parámetros de los Ítems usados en la simulación	39
5.2. Sesgo en la recuperación de parámetros de los ítems usando estimación bajo MCMC para el Modelo 2PL-GRM	40
5.3. Raíz del error cuadrático medio (RSME) y Media absoluta del error (MAE) de las estimaciones de los parámetros de los ítems en la simulación del Modelo de Respuesta Graduada Logístico (2PL-GRM) considerando R= 20 réplicas	41
5.4. Resultados del promedio de la Raíz del error cuadrático medio (RSME), la Media absoluta del error (MAE) y la correlación (r) de los parámetros estimados en la simulación del Modelo de Respuesta Graduada Logístico (2PL-GRM) considerando R = 20 réplicas	42
6.1. Análisis de la Dimensionalidad de la escala. Matriz de cargas factoriales rotadas con el método VARIMAX y comunalidades	47
6.2. Análisis de ítems bajo la perspectiva clásica	49
6.3. Estadísticos para las puntuaciones totales	49
6.4. Estimación de los parámetros de discriminación del Modelo 2PL-GRM	50
6.5. Estimación de los parámetros de dificultad del Modelo 2PL-GRM	51
6.6. Estimación de los parámetros de discriminación del Modelo LPE-GRM	54
6.7. Estimación de los parámetros de dificultad del Modelo LPE-GRM	55
6.8. Estimación de los parámetros de penalización del Modelo LPE-GRM	56
6.9. Medidas de comparación de los modelos para la satisfacción con la accesibilidad (Dimensión 1) y el desempeño del asesor (Dimensión 2)	57
6.10. Estadísticos para las escalas finales de satisfacción con la accesibilidad y el desempeño del asesor obtenidas con el Modelo LPE-GRM	57
C.1. Ítems considerados dentro del cuestionario de satisfacción con la atención de un Call Center	88

## Capítulo 1

# Introducción

### 1.1. Consideraciones Preliminares

La Teoría de Respuesta al Ítem (TRI), también conocida como teoría de rasgos latentes, es un enfoque psicométrico moderno que abarca un conjunto de modelos que [Pasquali \(2009\)](#) los caracteriza por adoptar dos axiomas fundamentales: En primer lugar, el desempeño o respuesta de un sujeto en una tarea (los ítems del test) son explicados por un conjunto de factores o rasgos latentes (habilidad, aptitud, etc.) a través de una relación del tipo causa y efecto, en donde el desempeño es el efecto y los rasgos latentes las causas que originan dichos efectos. La segunda característica de la TRI consiste en que existe una relación entre el desempeño de la tarea y el conjunto de rasgos latentes que puede ser descrita por una función matemática monótona decreciente, llamada Función de Respuesta a los Ítems (FRI) o Curva Característica de los Ítems (CCI). Como puntualiza [Ostini y Nering \(2006\)](#), las FRI básicamente relacionan la probabilidad de que un sujeto evaluado responda a un ítem de una manera determinada con la posición que tiene dicho sujeto en la escala de la habilidad que el ítem está midiendo. En otras palabras, la función describe en términos probabilísticos como una persona con un nivel alto de habilidad es factible que responda a diferentes categorías de respuesta de un ítem, en comparación con otra persona que tiene un nivel más bajo de la misma habilidad. De manera gráfica, las FRI suelen tener una forma similar a una ojiva.

Es posible considerar a la TRI como una extensión de la Teoría Clásica de los Test (TCT), que surge de manera formal con los trabajos iniciales de [Lawley \(1943\)](#) y [Lord y Novick \(1968\)](#), con la finalidad de superar las limitaciones prácticas que presenta la TCT en situaciones en que los test son cortos y por ignorar fuentes de error tales como la variación debido al efecto cuando las respuestas a los ítems deben ser juzgados subjetivamente. ([Bock y Moustaki, 2007](#)).

Entre los distintos tipos de modelos que abarca la TRI, [Bock y Moustaki \(2007\)](#) distinguen aquellos que son del tipo dicotómico, politómico (nominal y ordinal) y de rankings. Los modelos politómicos de la teoría de respuesta al ítem (TRIP) son modelos cuya finalidad es la de explicar la interacción que existe entre los sujetos evaluados y los atributos del test, en aquellas situaciones en las que estos atributos poseen varias categorías de respuesta ([Ostini y Nering, 2010](#)), lo cual los diferencia de los modelos tradicionales de la TRI que se encuentran enfocados en respuestas binarias (modelos dicotómicos). Dentro de la clasificación de modelos politómicos mostrada por [Bazán et al. \(2012\)](#), el Modelo de Respuesta Graduada (GRM) de [Samejima \(1969\)](#) de la TRIP está diseñado especialmente para aplicarse en instrumentos con atributos politómicos ordinales. En este modelo, las Funciones de Respuesta a las Etapas de

los Ítems (FREI, el equivalente a las FRI para el caso politómico) son usualmente consideradas simétricas, y de manera particular corresponden al Modelo de Respuesta Graduada usando enlace logístico (2PL-GRM). De manera similar a lo observado para los modelos dicotómicos por Samejima (2000), Bazán et al. (2006) y Bolfarine y Bazán (2010), en este estudio se propone usar FREI asimétricas considerando un nuevo parámetro de ítem que controle la forma de la curvatura

Por otro lado, el estudio y la medición de la calidad es uno de los elementos más importantes para poder determinar el valor de las propuestas de productos y servicios. Como menciona Allen (2004) dentro de los esfuerzos por mejorar la calidad de servicios y alinearlos a las expectativas de los clientes, surgen dos componentes muy importantes: la medición de la satisfacción y de la lealtad por parte de los clientes. Al ser tanto la satisfacción como la lealtad variables latentes, estas son medidas a través de cuestionarios conformados por atributos con respuestas múltiples, por lo general del tipo politómico (varias categorías de respuesta ordenadas o no), debido a que permiten producir mayor información sobre el nivel de actitud y percepción que tienen los clientes. A pesar de las ventajas mostradas por la TRI en psicometría para estudiar instrumentos con características similares (Lord y Novick, 1968; Pasquali, 2009), esta no ha sido aplicada regularmente en la medición de la satisfacción y lealtad de los clientes. La TRI ya es un importante área de trabajo en Estadística como se puede ver en Rao y Sinharay (2007).

Este estudio se justifica debido a que en el caso de la medición de la satisfacción y lealtad las tasas de respuesta de las categorías de los ítems no suelen tener las mismas proporciones y por tanto es necesario proponer modelos asimétricos antes que de tipo simétrico.

Consideramos que el aporte es tanto metodológico como conceptual al introducir y proponer nuevos modelos de datos para instrumentos de gran relevancia social puesto que la calidad de servicio brindada por las empresas es tanto un objetivo principal entre las organizaciones para desarrollar programas de mejora continua de sus procesos y servicios, así como una necesidad creciente entre los ciudadanos. Además, también creemos que el impacto académico es alto, ya que a nivel internacional inclusive existen pocas investigaciones aplicando modelos de la TRI en el área de la calidad de servicios. Esto es especialmente verdadero considerando que los trabajos De Battisti et al. (2005, 2010, 2012); Nicolini (2006) se basan en la aplicación del modelo Rasch de Crédito Parcial. El presente trabajo se inscribe en la línea de los trabajos recientemente presentados por Tarazona y Bazán (2011, 2012) extendiéndolo para el caso de considerar enlaces alternativos bajo inferencia bayesiana. Para este propósito la implementación computacional de los modelos a proponer será desarrollada usando programas de acceso libre como R (R Core Team, 2012) y WinBUGS (Lunn et al., 2000).

## 1.2. Objetivos

El objetivo general de este trabajo es presentar el Modelo de Respuesta Graduada para datos politómicos ordinales de la TRI propuesto por Samejima (1969) desde el punto de vista Bayesiano, así como proponer modelos alternativos basados en el uso de FREI asimétricas. Esto incluye el estudio de las propiedades del modelo, su estimación desde el punto de bayesiano, la implementación computacional del modelo así como la aplicación a un conjunto de datos reales en el ámbito de la medición de la Calidad de Servicios. De manera específica:



- Presentar una formulación general para los Modelos de Respuesta Graduada, su clasificación, características, funciones de información, casos particulares y criterios de selección del modelo.
- Presentar el Modelo de Respuesta Graduada Logístico y el Modelo de Respuesta Graduada Logístico de Exponente Positivo.
- Presentar la estimación bayesiana de los Modelos de Respuesta Graduada Logístico y Logístico de Exponente Positivo, los criterios para la comparación de modelos así como su respectiva implementación computacional usando software libre.
- Realizar estudios de simulación acerca de los modelos Respuesta Graduada considerando computación intensiva sobre diferentes escenarios y enlaces alternativos.
- Aplicar el modelo a conjunto de datos reales de calidad de servicios.

### 1.3. Organización del Trabajo

En el Capítulo 2 se presenta una formulación general para los Modelos de Respuesta Graduada (GRM), mostrando su clasificación en modelos simétricos y asimétricos, las características generales, el cálculo de las funciones de información, algunos casos particulares y los criterios necesarios para la selección del modelo. En el Capítulo 3 se define y se presentan propiedades de los Modelos de Respuesta Graduada Logístico (2PL-GRM) y Logístico de Exponente Positivo (LPE-GRM) como casos particulares de un GRM simétrico y asimétrico, respectivamente. En el Capítulo 4 se describe el proceso de estimación de los parámetros de los modelos descritos en el capítulo anterior bajo la perspectiva bayesiana usando MCMC, además de los criterios para la comparación de modelos y la implementación computacional usando R y WinBUGS. En el Capítulo 5 se presenta un estudio de simulación de recuperación de parámetros para el Modelo 2PL-GRM. En el Capítulo 6 se analizan los resultados de aplicar los modelos propuestos en el ámbito de la calidad de servicios, específicamente para el caso del análisis de la satisfacción de clientes. Finalmente se presentan conclusiones y sugerencias al presente trabajo en el Capítulo 7.

Como complemento a los capítulos anteriores se ha considerado tres anexos: En el Anexo A se incluyen los códigos de los programas de simulación y los utilizados en la aplicación sobre el estudio de satisfacción de clientes. En el Anexo B se muestran los resultados usados para el diagnóstico de la convergencia de las distribuciones a posteriori de los parámetros estimados en la simulación y la aplicación. En el Anexo Anexo C se muestran los ítems del instrumento considerados en el estudio de aplicación.

## Capítulo 2

# El Modelo de Respuesta Graduada de la TRIP

Dentro de la clasificación de modelos politómicos de la Teoría de Respuesta al Ítem (TRIP) mostrada por Bazán et al. (2012), el Modelo de Respuesta Graduada se caracteriza por estar diseñado especialmente para aplicarse en instrumentos constituidos de atributos con respuestas ordinales. Más que un modelo, el Modelo de Respuesta Graduada puede considerarse como un marco general de modelos aplicados a ítems politómicos ordinales como es descrito inicialmente en Samejima (1969) y Samejima (1972). En este capítulo se presenta una formulación general para los Modelos de Respuesta Graduada, su clasificación, características, funciones de información, casos particulares y principales criterios de selección del modelo.

### 2.1. Formulación del Modelo de Respuesta Graduada General

#### 2.1.1. Presentación y Racionalidad del Modelo

Consideremos un conjunto de  $i = 1, \dots, n$  sujetos evaluados en una prueba o cuestionario con  $j = 1, \dots, J$  ítems, donde cada uno de los ítems tiene  $m$  categorías con respuesta politómica ordinal. El Modelo de Respuesta Graduada General o simplemente Modelo de Respuesta Graduada (GRM, por las siglas en inglés de *Graded Response Model*) es definido como:

$$Y_{ij} \mid u_i, \xi_j \sim \text{Categórica}(P_{ij1}, \dots, P_{ijm}) \quad (2.1)$$

$$P_{ijk} \equiv P_{jk}(u_i) \equiv P(Y_{ij} = k \mid u_i, \xi_j) = P_{ijk}^{ac} - P_{ij(k-1)}^{ac} \quad (2.2)$$

$$P_{ijk}^{ac} \equiv P_{jk}^{ac}(u_i) \equiv P(Y_{ij} \leq k \mid u_i, \xi_j) = F(\eta_{ijk}) \quad (2.3)$$

$$\eta_{ijk} = a_j (b_{jk} - u_i) \quad (2.4)$$

$$-\infty = b_{j0} < b_{j1} < \dots < b_{j(m-1)} < b_{jm} = \infty \quad (2.5)$$

$$i = 1, \dots, n, j = 1, \dots, J, k = 1, \dots, m$$

donde:

- $Y_{ij}$  es una variable aleatoria manifiesta condicional que modela la respuesta politómica del individuo  $i$  en el ítem  $j$ , tomando distintos valores categóricos ordinales

$k \in \{1, \dots, m\}$ .

- $P_{ijk}$  son probabilidades condicionales de responder la categoría  $k$  en el ítem  $j$ .
- Considerando la distribución categórica debe cumplirse que  $\sum_{k=1}^m P_{ijk} = 1$
- Las  $P_{ijk}^{ac}$  son probabilidades condicionales acumuladas de responder en cualquier categoría inferior o igual a  $k$  en el ítem  $j$
- $u_i$  es el valor de la variable latente  $U_i$  (habilidad o rasgo latente) para el sujeto  $i$  tal que  $u_i \in \mathbb{R}$ .
- $F(\cdot)$  es una función de distribución acumulada (fda) que relaciona el predictor latente  $\eta_{ijk}$  con las probabilidades acumuladas condicionales.
- $\xi_j = (a_j, b_{j1}, \dots, b_{j(m-1)})'$  es el vector de parámetros del ítem  $j$ . El parámetro de discriminación  $a_j > 0$  está relacionado con la inclinación del ítem y los  $m-1$  parámetros de dificultad  $b_{jk}$  representan la posición de los umbrales (*thresholds*) entre dos categorías del ítem.

### 2.1.2. Observaciones

- El modelo presentado es considerado un modelo de variable latente monótono unidimensional  $(Y, U_i)$ .
- Aunque el número de categorías por cada ítem es asumido el mismo por comodidad, se puede tener un modelo más general considerando un número  $m_j$  diferente de categorías para cada ítem.
- Cuando es considerada la distribución acumulada de la distribución logística tenemos el modelo de respuesta graduada usual propuesto por Samejima (1969, 1972)
- La restricción dada en 2.5 de que  $b_{jk} > b_{j(k-1)}$  y considerando  $b_{j0} = -\infty$  y  $b_{jm} = +\infty$  es una condición importante que define el tipo de modelo de respuesta politómica. Esto es denominado por Samejima (1969) como respuesta graduada y define el nombre del modelo considerado.
- Las probabilidades de respuesta a las categorías  $P_{ijk}$  son consideradas como el **componente aleatorio observable** (Bazán et al., 2012) y como fue mostrados se obtienen substrayendo dos probabilidades acumuladas adyacentes  $P_{ijk}^{ac}$ . Además se considera de que  $P_{ij0}^{ac} = 0$  y  $P_{ijm}^{ac} = 1$ , por lo que de modo particular:

$$P_{ij1} = P_{ij1}^{ac} - P_{ij0}^{ac} = P_{ij1}^{ac}, \quad P_{ijm} = P_{ijm}^{ac} - P_{ij(m-1)}^{ac} = 1 - P_{ij(m-1)}^{ac}$$

teniendo finalmente que:

$$P_{ijk} \equiv P_{jk}(u_i) \equiv P(Y_{ij} = k \mid u_i, \xi_j) = \begin{cases} F(\eta_{ijk}) & \text{si } k = 1 \\ F(\eta_{ijk}) - F(\eta_{ij(k-1)}) & \text{si } 2 < k < m \\ 1 - F(\eta_{ij(k-1)}) & \text{si } k = m \end{cases} \quad (2.6)$$

- La habilidad o rasgo latente puede ser cualquier constructo que se supone que está detrás de las respuestas observables en los ítems. Se asume que  $U$  es continua en concordancia al supuesto de *continuum* dado por Pfanzagl et al. (1971), es decir que puede asumir cualquier valor dentro de la escala de los números reales  $(-\infty, \infty)$ . Es importante resaltar que a lo largo de este trabajo se ha preferido denotar a la habilidad como  $U$  en vez de la notación clásica psicométrica  $\theta$ , debido al estatus de variable aleatoria, en concordancia con los trabajos de Bazán et al. (2012) y Bartholomew et al. (2011).
- El parámetro de inclinación  $a_j$  se encuentra asociado a la discriminación del ítem  $j$ ; es decir, indica que tan útil es dicho ítem para diferenciar entre individuos con habilidades distintas. Este parámetro toma valores  $a_j > 0$  y para el GRM permanece constante a lo largo de todas las categorías del mismo ítem y tiene un valor proporcional a la inclinación de la FREI en los puntos  $b_{jk}$ . De acuerdo a la clasificación de Bazán et al. (2012) a los modelos con intercepto restringido al ítem se les denota como modelos  $2P(j)$ , por lo que una manera alternativa de nombrar al modelo descrito en 2.1 sería  $2P(j) - GRM$ . Según el criterio de Hafsteinsson et al. (2007) cuando  $a_j > 1$  se considera al ítem con una discriminación aceptable y cuando  $a_j > 2$  posee una discriminación bastante alta.
- Los parámetros  $b_{jk} \in \mathbb{R}$  y controlan la “posición o dificultad” de la respuesta ordinal  $k$  del ítem  $j$  al indicar el nivel de habilidad  $u_i$  en el cual la probabilidad de responder en alguna de las categorías posteriores a la  $k$  es igual a 0.5. Es por ello que es común denominar a estos parámetros como umbrales (*thresholds*) o etapas de ítem, ya que pueden ser vistos como un valor de frontera en la habilidad que debe de tener un individuo para que su selección tenga una transición de la categoría  $k$  hacia la  $k + 1$ .
- Bazán et al. (2010) recomienda usar la siguiente parametrización para el predictor latente lineal definido en 2.4 cuando se aplique la Inferencia Bayesiana:

$$\eta_{ijk} = \kappa_{jk} - a_j u_i \quad (2.7)$$

Esta forma alternativa es preferible debido a un criterio de optimización computacional ya que facilita el tiempo de convergencia. Al usar esta parametrización, la dificultad puede obtenerse haciendo  $b_{jk} = \frac{\kappa_{jk}}{a_j}$ .

### 2.1.3. El Modelo de Respuesta Graduado como extensión de un modelo binario

Baker y Kim (2004) señalan que el GRM tiene una contribución importante en la TRI en el sentido de que puede ser visto como una extensión de un modelo TRI binario y de

esta forma poder aplicarse en escalas ordinales tipo Likert. Sobre la base de lo anterior, las respuestas de los individuos presentadas en el modelo 2.1 pueden ser expresadas como:

$$W_{ijk} = \begin{cases} 1 & Y_{ij} = k \\ 0 & \text{otros casos} \end{cases} \quad (2.8)$$

De lo que se sigue que es equivalente expresar:

$$P_{ijk} = P(Y_{ij} = k | u_i, \xi_j) = P(W_{ijk} = 1 | u_i, \xi_j) \quad (2.9)$$

Por lo que  $W_{ijk} \sim \text{Bernoulli}(P_{ijk})$  con  $P_{ijk}$  definida como en 2.6

#### 2.1.4. El Modelo de Respuesta Graduada usando Aumentación de Datos

Asumamos que en el GRM existe una variable continua auxiliar  $Z_{ij}$  subyacente a la variable aleatoria de respuesta ordinal  $Y_{ij}$  de tal forma que:

$$Y_{ij} = \begin{cases} k & \tau_{j(k-1)} < Z_{ij} < \tau_{jk} \\ 0 & \text{otros casos} \end{cases} \quad (2.10)$$

donde  $T = (\tau_{j0}, \tau_{j1}, \dots, \tau_{jm})$  son puntos de corte que limitan a  $Z_{ij}$ , considerándose que  $\tau_{j0} = -\infty$  y  $\tau_{jm} = \infty$ .

Para realizar la inferencia en relación a la variable latente  $U_i$ ,  $Z_{ij}$  es expresada como una regresión en función de  $U_i$  resultando:

$$Z_{ij} = v_j + \lambda_j u_i + \varepsilon_{ij} \quad (2.11)$$

donde  $v_j$  es el intercepto,  $\lambda_j$  la ponderación de la regresión o carga factorial y  $\varepsilon_{ij}$  es el término del error, el cual es una variable aleatoria con fda similar a la utilizada en 2.3.

La distribución marginal de  $Z_{ij}$ ,  $g(\cdot)$ , está dada por:

$$g(Z_{ij}) = \int_{-\infty}^{\infty} f(Z_{ij} | u_i) h(U_i) dU_i \quad (2.12)$$

donde  $h(\cdot)$  es la función de densidad para  $U_i$  y  $Z_{ij} | u_i$  es la puntuación de  $Z_{ij}$  condicionada sobre  $u_i$  con función de densidad condicional  $f(\cdot)$  y

$$E[Z_{ij} | u_i] = v_j + \lambda_j u_i \quad (2.13)$$

y

$$\text{Var}[Z_{ij} | u_i] = \sigma_{\varepsilon_{ij}}^2 \quad (2.14)$$

La probabilidad de que un individuo con una habilidad  $u_i$  responda en la categoría  $k$  del ítem  $j$  estará dado por :

$$P(Y_{ij} = k | u_i, \xi_j) = \int_{\tau_{j(k-1)}}^{\tau_{jk}} f(Z_{ij} | u_i) dZ_{ij} | u_i \quad (2.15)$$

De la Ecuación 2.11 se obtiene:

$$\begin{aligned}
 P(Y_{ij} = k | u_i, \xi_j) &= P(\tau_{j(k-1)} \leq v_j + \lambda_j u_i + \varepsilon_{ij} \leq \tau_{jk}) \\
 &= P(\tau_{j(k-1)} - v_j - \lambda_j u_i \leq \varepsilon_{ij} \leq \tau_{jk} - v_j - \lambda_j u_i) \\
 &= F(\tau_{jk} - v_j - \lambda_j u_i) - F(\tau_{j(k-1)} - v_j - \lambda_j u_i)
 \end{aligned}$$

considerando  $v_j = 0$ ,  $a_j = \lambda_j$ ,  $b_{jk} = \frac{\tau_{jk}}{\lambda_j}$ , obtenemos:

$$P(Y_{ij} = k | u_i, \xi_j) = F(a_j(b_{jk} - u_i)) - F(a_j(b_{j(k-1)} - u_i))$$

que es similar a lo expresado en 2.6 para el GRM.

## 2.2. Clasificación de los Modelos de Respuesta Graduada

### 2.2.1. Modelos de Respuesta Graduada Simétricos

Asumiendo que existe una variable latente continua auxiliar  $Z_{ij}$  subyacente a la respuesta ordinal  $Y_{ij}$ , se dice que un GRM es simétrico cuando se cumplen las siguientes características (Samejima, 2010):

1.  $Y_{ij}$  puede ser definido como en 2.10.
2. La regresión de  $Z_{ij}$  en función de  $U_i$  definida en 2.11 es lineal.
3. La función de densidad condicional  $f(Z_{ij} | u_i)$  puede ser cualquier distribución de densidad condicional simétrica o asimétrica con tal que se mantenga para todos los valores fijos de  $U_i$ .
4. La varianza  $\sigma_{\varepsilon_{ij}}^2$  es la misma para todo  $U_i$ .

La otra forma de expresar un modelo de respuesta graduada simétrico es considerando  $F(\cdot)$  como siendo la función de distribución acumulada de una distribución simétrica como por ejemplo la Normal o Logística.

### 2.2.2. Modelos de Respuesta Graduada Asimétricos

Aceptando la relación lineal entre  $Z_{ij}$  y  $U_i$  dada en 2.11, entonces cualquier violación de la simetría en la distribución marginal de  $Z_{ij}$  es causada por las siguientes causas mutuamente no excluyentes (Molenaar et al., 2012):

1.  $\sigma_{\varepsilon_{ij}}^2$  es heterocedástica (varía a lo largo de  $U_i$ ). Ver Figura 2.1(b)
2. La distribución de  $U_i$  es asimétrica. Ver Figura 2.1(c)
3. La distribución de  $Z_{ij} | u_i$  es asimétrica 2.1(d).

La otra forma de expresar un modelo de respuesta graduada asimétrico es considerando  $F(\cdot)$  como siendo la función de distribución acumulada de una distribución asimétrica.

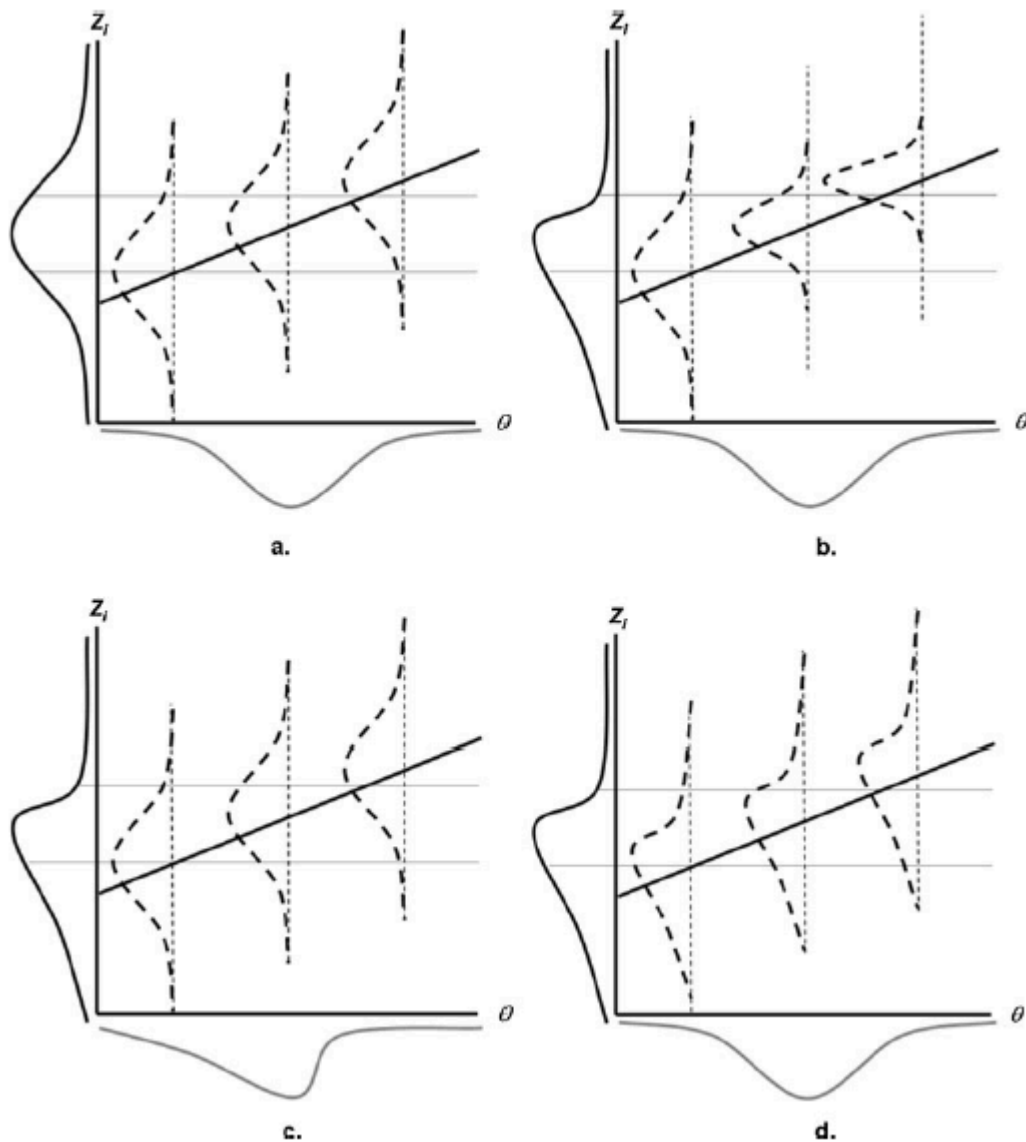


Figura 2.1: Representación gráfica tomada de Molenaar et al. (2012) que muestra las causas de asimetría en los Modelos de Respuesta Graduada: (a) Modelo Tradicional (b)  $\sigma_{\varepsilon_{ij}}^2$  heterocedástico (c)  $U_i$  asimétrico (d)  $Z_{ij} | u_i$  asimétrico. Para mayor claridad el impacto del efecto es exagerado.

### 2.3. Supuestos de los Modelos de Respuesta Graduada

Consideremos las siguientes notaciones iniciales. Para un conjunto de  $J$  ítems el patrón o conjunto integral de respuestas del individuo  $i$ , denotado como  $\mathbf{Y}_i$ , indica la secuencia de  $Y_{ij}$  para  $j = 1, \dots, J$  y su realización  $\mathbf{y}_i$  puede ser escrita como:

$$\mathbf{y}_i = (y_{i1}, \dots, y_{ij}, \dots, y_{iJ})'$$

De esta forma podemos expresar al conjunto de respuestas de  $n$  individuos como:

$$\mathbf{y} = \mathbf{y}_{..} = (\mathbf{y}_{1.}', \dots, \mathbf{y}_{i.}', \dots, \mathbf{y}_{n.}')'$$

Usando la parametrización alternativa planteada en 2.8 tenemos de manera equivalente



que la realización del conjunto integral de respuestas  $\mathbf{W}_{i..}$  estará dada por:

$$\mathbf{w}_{i..} = \left( \mathbf{w}'_{i1.}, \dots, \mathbf{w}'_{ij.}, \dots, \mathbf{w}'_{iJ.} \right)'$$

y el conjunto total de respuestas:

$$\mathbf{w} = \mathbf{w}_{...} = \left( \mathbf{w}_{1..}', \dots, \mathbf{w}_{j..}', \dots, \mathbf{w}_{n..}' \right)'$$

en donde  $\mathbf{w}_{ij.} = (w_{ij1}, \dots, w_{ijk}, \dots, w_{ijm})'$  es el vector aleatorio de respuestas del individuo  $i$  en el ítem  $j$ .

De acuerdo con [Azevedo \(2003\)](#), en el GRM deben hacerse tres suposiciones que son imprescindibles para el desarrollo de los procesos de estimación:

- (S1) Las respuestas provenientes de individuos distintos son independientes entre si. De esta forma, la probabilidad del conjunto global de respuestas estará dado por:

$$P(\mathbf{Y} = \mathbf{y} \mid u_i, \xi_j) = \prod_{i=1}^n P(\mathbf{Y}_i = \mathbf{y}_i \mid u_i, \xi_j) \quad (2.16)$$

- (S2) Dado un nivel de habilidad  $u_i$ , las respuestas  $\mathbf{Y}_i$  del individuo  $i$  en los  $J$  ítems son independientes (independencia local). En consecuencia, la probabilidad del patrón de respuesta  $\mathbf{y}_i$  puede ser escrita como:

$$P(\mathbf{Y}_i = \mathbf{y}_i \mid u_i, \xi_j) = \prod_{j=1}^J P(Y_{ij} = y_{ij} \mid u_i, \xi_j) \quad (2.17)$$

- (S3) La probabilidad de la respuesta  $y_{ij}$  puede ser representada por un **modelo multivariado de Bernoulli**

$$P(Y_{ij} = y_{ij} \mid u_i, \xi_j) = P(\mathbf{W}_{ij.} = \mathbf{w}_{ij.} \mid u_i, \xi_j) \equiv P(\mathbf{W}_{ij.} = \mathbf{w}_{ij.}) = \prod_{k=1}^m [P_{ijk}]^{w_{ijk}} \quad (2.18)$$

## 2.4. Características de los Modelos de Respuesta Graduada

Las principales características del GRM se encuentran descritas por [Samejima \(1997b, 2004, 2010\)](#), las cuales son presentadas a continuación:

### 2.4.1. Función de Procesamiento (FP)

Es posible considerar a  $Y_{ij} = k$  como la puntuación ordinal a ser asignada a las personas evaluadas que satisfactoriamente han completado la etapa  $k - 1$  pero que han fallado en completar la etapa  $k$ , así de esta manera se puede definir a  $M_{jk}(u_i)$ , la Función de Procesamiento (FP) para la respuesta  $Y_{ij} = k$ , con  $k = 1, 2, 3, \dots, m, m + 1$ , como la función de probabilidad condicionada conjuntamente a que el examinado tenga un mismo nivel de habilidad  $u_i$  y que haya superado los procesos o etapas anteriores. Sea  $(m + 1)$  la puntuación ordinal siguiente



a la última categoría  $m$ . Dado de que cualquiera puede obtener al menos una puntuación de 1, y nadie es capaz de obtener una puntuación de  $(m + 1)$ , es razonable establecer:

$$M_{jk}(u_i) = \begin{cases} 1 & \text{si } y_{ij} = 1 \\ 0 & \text{si } y_{ij} = m + 1 \end{cases}$$

para todo  $U$ . Se asume que la FP es no decreciente en  $U$ .

#### 2.4.2. Función de Respuesta a las Etapas del Ítem (FREI)

La FREI para un ítem  $j$  se encuentra definida como:

$$P_{jk}^+(u_i) \equiv P(Y_{ij} \geq k | u_i, \xi_j) = 1 - F(\eta_{ijk}) \quad (2.19)$$

con  $k = 1, 2, \dots, m$ , donde  $P_{jk}^+(u_i)$  es la probabilidad condicional de que el individuo con habilidad  $u_i$  logre superar al menos la etapa  $k$  (responder en una categoría  $k$  o superior).

La ecuación anterior puede ser también expresada en términos de la FP:

$$P_{jk}^+(u_i) \equiv P(Y_{ij} \geq k | u_i, \xi_j) = \prod_{h=1}^k M_{jh}(u_i) \quad (2.20)$$

En particular,  $P_{jk}^+(u_i) = 1$  cuando  $Y_{ij} = 1$  y  $P_{jk}^+(u_i) = 0$  cuando  $Y_{ij} = m + 1$ . Además se cumple que para cualquier valor de  $k = 1, \dots, m$ :

$$P_{jk}^+(u_i) \geq P_{j(k+1)}^+(u_i)$$

De acuerdo a [Hemker \(2001\)](#), estas funciones corresponden a las del tipo acumulativo, puesto que como se definió en 2.5, un orden estricto en las etapas del ítem es asumido. Esta restricción en el orden implica que cualquier respuesta obtenida dice algo sobre todas las etapas del ítem. Los modelos de esta naturaleza también son conocidos como modelos de probabilidades acumuladas, modelos de diferencia, modelos de Thurstone y modelos de gradación ([Bazán et al., 2012](#)).

#### 2.4.3. Función de Respuesta a las Categorías del Ítem (FRCI)

La FRCI para el ítem  $j$  fue definida en 2.2 y 2.6, pudiendo ser también expresada en términos de las FREI como:

$$P_{jk}(u_i) \equiv P(Y_{ij} = k | u_i, \xi_j) = P_{jk}^+(u_i) - P_{j(k+1)}^+(u_i) \quad (2.21)$$

con  $k = 1, 2, \dots, m$ , donde  $P_{jk}(u_i)$  modela la probabilidad de responder en una categoría específica del ítem como función del nivel de habilidad  $U_i$  y corresponden a las  $P_{ijk}$  definidas en 2.6.

Es importante notar que cuando  $m = 2$  las ecuaciones 2.19 y 2.21 para  $P_{jk}^+(u_i)$  como  $P_{jk}(u_i)$  corresponden a la Función de Respuesta del Ítem (FREI) o Curva Característica del Ítem (CCI) de un modelo TRI binario de dos parámetros. En otras palabras, un GRM definido de esta forma modela respuestas binarias como un caso especial, de esta forma el GRM como marco general también aplica para este tipo de ítems.

#### 2.4.4. Función Básica (FB)

La FB para el ítem  $j$  se encuentra definida como:

$$A_{jk}(u_i) \equiv \frac{\partial}{\partial u_i} \log P_{jk}(u_i) = [P_{jk}(u_i)]^{-1} \frac{\partial}{\partial u_i} P_{jk}(u_i) \quad (2.22)$$

con  $k = 1, 2, \dots, m$ , siendo evidente que la FB existe mientras que  $P_{jk}(u_i)$  sea positiva para todo  $U$  y sea diferenciable con respecto a  $U$ .

#### 2.4.5. Función de Información de las Respuestas de los Ítems (FIRI)

Dada la Función Básica (FB), la FIRI para el ítem  $j$  se encuentra definida como:

$$\begin{aligned} I_{jk}(u_i) &= -\frac{\partial}{\partial u_i} A_{jk}(u_i) = -\frac{\partial^2}{\partial u_i^2} \log P_{jk}(u_i) \\ &= -[P_{jk}(u_i)]^{-2} \left[ P_{jk}(u_i) \frac{\partial^2}{\partial u_i^2} P_{jk}(u_i) - \left\{ \frac{\partial}{\partial u_i} P_{jk}(u_i) \right\}^2 \right] \\ &= -[P_{jk}(u_i)]^{-1} \left[ \frac{\partial^2}{\partial u_i^2} P_{jk}(u_i) \right] + [P_{jk}(u_i)]^{-2} \left[ \frac{\partial}{\partial u_i} P_{jk}(u_i) \right]^2 \\ &= [P_{jk}(u_i)]^{-2} \left[ \frac{\partial}{\partial u_i} P_{jk}(u_i) \right]^2 - [P_{jk}(u_i)]^{-1} \left[ \frac{\partial^2}{\partial u_i^2} P_{jk}(u_i) \right] \end{aligned} \quad (2.23)$$

con  $k = 1, 2, \dots, m$ . Es posible concluir que las FIRI existen mientras que  $P_{jk}(u_i)$  sea positiva para todo el rango de  $U$  y sea doblemente diferenciable con respecto a  $U$ .

#### 2.4.6. Condición de un Máximo Único

Sea  $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{ij}, \dots, y_{iJ})'$  un patrón de respuesta específico para el individuo  $i$  en los  $J$  ítems. Debido a los supuestos (S2) de independencia local podemos escribir la verosimilitud del patrón específico de respuesta como:

$$L(\mathbf{y}_i | u_i, \xi_j) = P(\mathbf{Y}_i = \mathbf{y}_i | u_i, \xi_j) = \prod_{j=1}^J P(Y_{ij} = y_{ij} | u_i, \xi_j) \quad (2.24)$$

Usando las Ecuaciones 2.22 y 2.24 tenemos que:

$$\frac{\partial}{\partial u_i} \log L(\mathbf{y}_i | u_i, \xi_j) = \sum_{y_{ij} \in \mathbf{y}_i} \frac{\partial}{\partial u_i} \log P_{jk}(u_i) = \sum_{y_{ij} \in \mathbf{y}_i} A_{jk}(u_i) \equiv 0 \quad (2.25)$$

Es decir, existe una condición suficiente de que un GRM en específico proporciona un máximo único para cualquier función de verosimilitud (para cualquier y todos los patrones de respuesta) y para ello deben de cumplirse las dos siguientes condiciones:

1. La FB  $A_{jk}(u_i)$  para las puntuaciones graduadas  $y_{ij}$  de cada ítem  $j$  es estrictamente decreciente en  $U$ , es decir  $A_{j(k-1)}(u_i) < A_{jk}(u_i)$  para todo  $U$  y  $k = 1, \dots, m$
2. Las asíntotas superiores e inferiores de  $A_{jk}(u_i)$  son no negativas y no positivas de manera respectiva

Las anteriores condiciones son conocidas como la *condición del máximo único* Samejima

(1997b, 2004). Esta condición solo garantiza la unicidad de la solución pero no garantiza la existencia de un máximo global.

#### 2.4.7. Sobreparametrización

La sobreparametrización es una de las características más resaltantes de los modelos de la TRI (Bazán et al., 2010). El GRM tiene  $n + mj$  parámetros desconocidos. Esto hace que la inferencia en estos modelos sea intensiva computacionalmente. Es directo notar que a mayor tamaño de muestra y mayor número de ítems, el número de parámetros a estimar aumenta.

#### 2.4.8. Presencia de Parámetros Incidentales

Neyman y Scott (1948) señalan que las estimaciones de los parámetros estructurales no suelen ser consistentes cuando son estimados de manera conjunta con parámetros incidentales. En los modelos de la TRI, para un número fijo de ítems, los parámetros de los ítems son denominados estructurales mientras que los parámetros de la variable latente son considerados incidentales debido a que el número de parámetros de la variable latente se incrementa si se incrementa el número de individuos evaluados. Es importante resaltar que en la inferencia bayesiana la distinción entre parámetros incidentales y estructurales carece de sentido porque ambos son tratados como variables aleatorias y la incertidumbre sobre su valor es especificada en la distribución a priori. (Bazán et al., 2010).

#### 2.4.9. Identificabilidad

Un parámetro se dice identificado cuando distintos valores del mismo implican diferentes distribuciones de probabilidad para los datos (Bazán et al., 2010). Por ejemplo, para una habilidad  $u_i = 0$ , si en el predictor lineal del GRM dado en la Ecuación 2.4 se reemplaza  $a_j = 2$ ,  $b_{jk}=0.5$  y  $b_{j(k-1)} = 0$ , la probabilidad de responder en la categoría  $k$  estará dada por  $P_{ijk} = F(1) - F(0)$ , la cual es similar que cuando  $a_j = 0.5$ ,  $b_{jk} = 2$  y  $b_{j(k-1)} = 0$ . Si bien distintas combinaciones de valores no implican necesariamente diferentes distribuciones de probabilidad para los datos, la probabilidad de respuesta en una categoría  $k$  es la misma. Como resalta Bazán et al. (2010) los parámetros del modelo no pueden ser estimados de manera única a menos de que se impongan ciertas restricciones, lo cual vuelve complejo el proceso de inferencia de los parámetros. Desde el punto de vista clásico la solución más utilizada es enfrentar el problema dividiéndolo en dos etapas (p. ej. el método de Máxima Verosimilitud Marginal vía EM). Desde la perspectiva bayesiana, la sobreparametrización se mantiene y el tiempo computacional es generalmente mucho mayor que con la alternativa clásica. Sin embargo esta estimación conjunta resulta conveniente al reflejar mejor el problema original de la TRI que es estimar las características individuales y de los ítems de manera simultánea y completa

### 2.5. Información de Fisher en los Modelos de Respuesta Graduada

Reckase (2009) menciona que el término *información* en TRI es usado como un indicador estadístico de la calidad de la estimación de un parámetro, usualmente la habilidad que es el objetivo de la prueba. Debido a que las respuestas a los ítems proporcionan información para todos los niveles de la escala de U, la información es usualmente representada como una función del parámetro a ser estimado. La fórmula para hallar la Función de Información de

un Modelo puede ser derivado usando diferentes métodos.

Es conocido en la literatura estadística que el estimador de máxima verosimilitud de un parámetro, en este caso  $U$ , es distribuido con una media igual a  $u$  y varianza igual a:

$$\sigma^2(\hat{U} | U) = \frac{1}{E \left[ \left( \frac{\partial}{\partial u} \ln L \right)^2 |_{U=u} \right]} \quad (2.26)$$

donde  $L$  es la verosimilitud de los datos dado el parámetro, en este caso la verosimilitud de las respuestas a los ítems dado  $U$ . Lord (1980) demostró que el término en el denominador de 2.26 es igual a la Función de Información del Test (FIT) para el caso dicotómico. El recíproco de la varianza del estimador de máxima verosimilitud fue llamado como información por Fisher (1925) por lo que la formulación de la información es usualmente llamada como Información de Fisher (ver por ejemplo, Baker y Kim, 2004). Fisher (1925) considero a la varianza de esta estimación como una medida de *precisión intrínseca*. Los conceptos de información sobre  $U$  han sido generalizados para test compuestos por ítems politómicos ordinales como el GRM.

Consideremos un sujeto con habilidad  $u$ ,  $Y_j | u, \xi_j \sim \text{categórica}(P_{j1}, \dots, P_{jm})$  la respuesta graduada del sujeto en el ítem  $j = 1, \dots, m$  y  $f(y; u)$  la función de verosimilitud de la variable latente  $u$ . Partiendo de 2.26, es posible definir a la Información de Fisher como:

$$I(U) = E \left[ \left( \frac{\partial}{\partial u} \ln f(y; u) \right)^2 |_{U=u} \right] \quad (2.27)$$

o de manera equivalente

$$I(U) = E \left[ \left( \frac{\partial^2}{\partial u^2} \ln f(y; u) \right) |_{U=u} \right] \quad (2.28)$$

Debido a los supuestos (S2) y (S3) definidas en 2.3 podemos expresar a la verosimilitud como:

$$f(y; u) = \prod_{j=1}^J \prod_{k=1}^m [P_{jk}]^{w_{jk}} \quad (2.29)$$

siendo  $P_{jk}$  la probabilidad de respuesta del sujeto en la categoría  $k = 1, \dots, m$  del  $j$ -ésimo ítem y  $w_{jk}$  la parametrización alternativa para  $y_{jk}$  planteada en 2.8. El logaritmo de la verosimilitud estará expresado por:

$$\ln f(y; u) = \sum_{j=1}^J \sum_{k=1}^m w_{jk} \ln P_{jk}$$

tomando la primera derivada:

$$\begin{aligned}\frac{\partial}{\partial u} \ln f(y; u) &= \sum_{j=1}^J \sum_{k=1}^m w_{jk} \frac{\partial}{\partial u} \ln P_{jk} \\ &= \sum_{j=1}^J \sum_{k=1}^m \frac{w_{jk}}{P_{jk}} \frac{\partial}{\partial u} P_{jk}\end{aligned}$$

reemplazando  $P'_{jk}(u) = \frac{\partial}{\partial u} P_{jk}$ :

$$\frac{\partial}{\partial u} \ln f(y; u) = \sum_{j=1}^J \sum_{k=1}^m \frac{w_{jk}}{P_{jk}} P'_{jk}(u)$$

reemplazando la expresión anterior en 2.27:

$$\begin{aligned}I(U) &= E \left[ \left( \sum_{j=1}^J \sum_{k=1}^m \frac{w_{jk}}{P_{jk}} P'_{jk}(u) \right)^2 \middle| U=u \right] \\ &= E \left[ \sum_{j=1}^J \sum_{k=1}^m \frac{w_{jk}^2}{P_{jk}^2} [P'_{jk}(u)]^2 \right] \\ &= \sum_{j=1}^J \sum_{k=1}^m \frac{[P'_{jk}(u)]^2}{P_{jk}^2} E[w_{jk}^2]\end{aligned}$$

Como  $W_{jk} \sim \text{Bernoulli}(P_{jk})$  y  $\text{Var}[W_{jk}] = E[W_{jk}^2] - E[W_{jk}]^2$ :

$$\begin{aligned}I(U) &= \sum_{j=1}^J \sum_{k=1}^m \frac{[P'_{jk}(u)]^2}{P_{jk}^2} [P_{jk}(1 - P_{jk}) + P_{jk}^2] \\ &= \sum_{j=1}^J \sum_{k=1}^m \frac{[P'_{jk}(u)]^2}{P_{jk}}\end{aligned}\tag{2.30}$$

Partiendo de 2.30, definimos a la Función de Información de un Ítem (FII) como:

$$I_j(U) = \sum_{k=1}^m \frac{[P'_{jk}(u)]^2}{P_{jk}}\tag{2.31}$$

En consecuencia, la Función de Información del Test (FIT) es definida como la suma de las FII de los  $J$  ítems:

$$I(U) = \sum_{j=1}^J I_j(U)\tag{2.32}$$

Siguiendo la convención habitual utilizada en los Modelos TRI binarios, (ver por ejem-

plo Hambleton y Swaminathan, 1985), la función para el error estándar puede ser obtenida reemplazando la expresión anterior en la Ecuación 2.26:

$$SE(U) = \left[ \sum_{j=1}^J I_j(U) \right]^{-\frac{1}{2}} = \frac{1}{\sqrt{I(U)}} \quad (2.33)$$

Como comenta Reckase (2009), existe una literatura emergente sobre otras formas alternativas de funciones de información adicionales a la de Fisher descritas en este trabajo. Al respecto se recomienda revisar los trabajos de Savalei (2006) y Chang y Ying (1996) sobre el uso de la información de Kullback-Leibler como una alternativa a la de Fisher.

## 2.6. Casos Particulares

### 2.6.1. Modelo de Respuesta Graduada de un Parámetro (1P-GRM)

El Modelo de Respuesta Graduada de un Parámetro (1P-GRM) es una versión restringida del MRG bajo la asunción de que todos los ítems tienen la misma discriminación, es decir, el predictor lineal estará dado por:

$$\eta_{ijk} = a(b_{jk} - u_i), \quad i = 1, \dots, n, \quad j = 1, \dots, J, \quad k = 1, \dots, m$$

### 2.6.2. Modelo Rating Scale de Muraki (RS-GRM)

De acuerdo a Zhu y Stone (2012), otra restricción del GRM es el Modelo Rating Scale (RS-GRM) propuesto por Muraki (1990). El predictor lineal en este modelo es:

$$\eta_{ijk} = a_j [(b_j - c_k) - u_i], \quad i = 1, \dots, n, \quad j = 1, \dots, J, \quad k = 1, \dots, m$$

Como puede observarse, los parámetros de dificultad ( $b_{jk}$ ) del GRM son divididos en dos términos: un parámetro de posición ( $b_j$ ) distinto para cada ítem, y un conjunto de parámetros umbrales de categorías ( $c_k$ ) que se asumen similares para todos los ítems. Como resultado de lo anterior, el número de parámetros en el modelo RS-GRM se reduce de manera significativa en comparación con el GRM.

### 2.6.3. Modelo de Respuesta Graduada Multidimensional

El Modelo GRM y los casos particulares anteriores son apropiados para analizar respuestas de ítems que se asumen están determinadas por una sola habilidad o rasgo latente (supuesto de unidimensionalidad). Cuando un instrumento está diseñado para evaluar más de una habilidad, las respuestas de los ítems suelen mostrar una estructura multidimensional. De Ayala (1994) discute una versión multidimensional para el Modelo GRM donde el predictor lineal está expresado como:

$$\eta_{hijk} = \sum_h a_{jh} (b_{jk} - u_{ih}), \quad i = 1, \dots, n, \quad j = 1, \dots, J, \quad k = 1, \dots, m$$

donde  $u_{ih}$  es el nivel de habilidad para el individuo  $i$  en la dimensión  $h$ ,  $a_{jh}$  es el parámetro de discriminación (pendiente) del ítem  $j$  en la dimensión  $h$  y  $b_{jk}$  son los umbrales para la categoría  $k$  del ítem  $j$ .

#### 2.6.4. Modelo Testlet de Respuesta Graduada

Una modificación del modelo GRM para testlets fue propuesto por Wang et al. (2002), donde el predictor lineal está expresado por:

$$\eta_{ijk} = a_j [(b_j + \gamma_{id(j)}) - u_i], \quad i = 1, \dots, n, \quad j = 1, \dots, J, \quad k = 1, \dots, m$$

En este modelo, se introduce un efecto aleatorio testlet ( $\gamma_{id(j)}$ ) para reflejar la interacción entre el individuo  $i$  con el testlet  $d(j)$ . Se suele asumir que la habilidad  $u_i \sim N(0, 1)$  y  $\gamma_{d(j)} \sim N(0, \sigma_{d(j)}^2)$ . La varianza de  $\gamma_{d(j)}$  varía a lo largo de los testlets y su valor indica la cantidad de dependencia local en cada testlet. A medida de que  $\sigma_{d(j)}^2$  se incrementa, la cantidad de dependencia local también aumenta. Cuando  $\sigma_{d(j)}^2 = 0$  los ítems dentro de un teslet son condicionalmente independientes.

#### 2.7. Criterios para la selección del Modelo de Respuesta Graduada

Samejima (1996) propuso cinco criterios distintos para evaluar los GRM:

1. Los principios detrás del modelo y el conjunto de asunciones deben de tener concordancia con la naturaleza psicológica que subyace en los datos.
2. Aditividad 1, es decir, si las categoría de respuesta graduada existentes pudieran subdividirse (v.gr. aprobar o desaprobado son divididos en A, B, C, D y F), sus FRCI pueden ser especificadas del mismo modo.
3. Aditividad 2, es decir, después de una combinación de dos o más categorías de respuesta graduada adyacente (v. gr. A, B, C, D y F en aprobar y desaprobado), las FRCI de las nuevas categorías combinadas pueden seguir siendo especificadas de la misma forma (si la aditividad 1 y 2 se mantienen, el modelo puede ser naturalmente expandido a un modelo de respuesta continuo)
4. El modelo satisface la condición del máximo único.
5. Los puntos modales de las FRCI para las  $m$  categorías de respuesta graduada son ordenadas de acuerdo con las puntuaciones ordinales de los ítems,  $y_{ij} = 1, 2, \dots, m$ .



## Capítulo 3

# Modelo de Respuesta Graduada Logístico y LPE

En este capítulo se definen y presentan las propiedades de los Modelos de Respuesta Graduada Logístico (2PL-GRM) y Logístico de Exponente Positivo (LPE-GRM). El modelo 2PL-GRM es uno de los GRM más estudiados y en la literatura se le suele denominar del mismo modo que a la generalización de modelos de respuesta graduada, sin diferenciarlos. En este trabajo se le da una terminología distinta en concordancia a Samejima (2010) y es presentado como un caso particular de un GRM simétrico. Por otro lado, el modelo LPE-GRM es propuesto como una alternativa de GRM asimétrico (Samejima, 1969, 2010) y extensión de los estudios realizados para los casos de modelos TRI asimétricos binarios (Bolfarine y Bazán, 2010; Flores, 2012).

### 3.1. Modelo de Respuesta Graduada Logístico de dos Parámetros (2PL-GRM)

#### 3.1.1. Definición del Modelo

Dada una matriz de datos  $\mathbf{Y} = [Y_{ij}]_{n \times J}$  que representa las respuestas ordinales de los  $n$  individuos evaluados en una prueba o cuestionario de  $J$  ítems con  $Y_{ij}$  definida como en 2.1. Se puede considerar al Modelo de Respuesta Graduada Logístico de dos Parámetros (2PL-GRM) como un caso particular de un GRM simétrico en el cual  $F(\cdot)$  es la fda de la distribución logística estándar. El modelo 2PL-GRM queda definido como:

$$\begin{aligned}
 Y_{ij} \mid u_i, \xi_j &\sim \text{Categórica}(P_{ij1}, \dots, P_{ijm}) \\
 P_{ijk} &\equiv P_{jk}(u_i) \equiv P(Y_{ij} = k \mid u_i, \xi_j) = P_{ijk}^{ac} - P_{ij(k-1)}^{ac} \\
 P_{ijk}^{ac} &\equiv P_{jk}^{ac}(u_i) \equiv P(Y_{ij} \leq k \mid u_i, \xi_j) = \Psi(\eta_{ijk})
 \end{aligned} \tag{3.1}$$

$$P_{ijk} \equiv P_{jk}(u_i) \equiv P(Y_{ij} = k \mid u_i, \xi_j) = \begin{cases} \Psi(\eta_{ijk}) & \text{si } k = 1 \\ \Psi(\eta_{ijk}) - \Psi(\eta_{ij(k-1)}) & \text{si } 2 < k < m \\ 1 - \Psi(\eta_{ij(k-1)}) & \text{si } k = m \end{cases} \tag{3.2}$$

$$\begin{aligned}
 \eta_{ijk} &= a_j (b_{jk} - u_i) \\
 -\infty &= b_{j0} < b_{j1} < \dots < b_{j(m-1)} < b_{jm} = \infty
 \end{aligned}$$



$$i = 1, \dots, n, j = 1, \dots, J, k = 1, \dots, m$$

donde  $P_{ijk}$  son las probabilidades condicionales de responder la categoría  $k$  en el ítem  $j$ ,  $P_{ijk}^{ac}$  las probabilidades condicionales acumuladas de responder en cualquier categoría inferior o igual a  $k$  en el ítem  $j$ ,  $\eta_{ijk}$  el predictor latente lineal relacionado con  $u_i \in \mathbb{R}$ ,  $\xi_j = (a_j, b_{j1}, \dots, b_{j(m-1)})'$  el vector de parámetros del ítem  $j$  relacionado con la inclinación ( $a_j$ ) y posición ( $b_{jk}$ ) del ítem y  $\Psi(\cdot)$  la fda de la distribución logística estándar:

$$\Psi(\eta_{ijk}) = \frac{1}{1 + e^{-\eta_{ijk}}} \quad -\infty < \eta_{ijk} < \infty \quad (3.3)$$

### 3.1.2. Características del Modelo 2PL-GRM

#### Función de Respuesta a las Etapas del Ítem (FREI)

De acuerdo a lo discutido en 2.4.2, las FREI para el modelo 2PL-GRM estarán definidas como:

$$P_{jk}^+(u_i) = \Psi(-\eta_{ijk}) = \frac{1}{1 + e^{\eta_{ijk}}} \quad (3.4)$$

$$\eta_{ijk} = a_j (b_{jk} - u_i), \quad i = 1, \dots, n, j = 1, \dots, J, k = 1, \dots, m$$

Es directo notar que  $\Psi(-\eta_{ijk}) = 1 - \Psi(\eta_{ijk})$  al ser una distribución simétrica.

En la Figura 3.1 se muestra las FREI para tres ítems de 5 categorías graduadas con umbrales  $b = (-1.5, -0.5, 0.5, 1.5)$  para los tres casos y con tres parámetros de discriminación distintos para cada figura: 0.8, 1.6 y 2.4 (reflejando una discriminación baja, promedio y alta). Como puede observarse en las tres situaciones las FREI se encuentran ordenadas en concordancia con el umbral respectivo  $b_{j1} < b_{j2} < b_{j3} < b_{j4}$  y tienen la misma inclinación al ser el parámetro  $a_j$  similar dentro de cada gráfica. Se puede observar que a medida que los ítems tienen un mayor valor de  $a_j$  las FREI se vuelven más inclinadas (mayor pendiente).

#### Función de Respuesta a las Categorías del Ítem (FRCI)

Siguiendo las definiciones dadas en 2.4.3, las FRCI para el modelo 2PL-GRM estarán definidas como:

$$P_{jk}(u_i) = \frac{1}{1 + e^{-\eta_{ijk}}} - \frac{1}{1 + e^{-\eta_{ij(k-1)}}} \quad (3.5)$$

$$\eta_{ijk} = a_j (b_{jk} - u_i), \quad i = 1, \dots, n, j = 1, \dots, J, k = 1, \dots, m$$

La forma que toman las FRCI depende también de los parámetros de los ítems ( $a_j$  y  $b_{jk}$ ). En la Figura 3.2 se ilustra el comportamiento de la FRCI de 6 ítems con cinco categorías de respuesta graduada. De manera global, estas curvas difieren de las tradicionales FRI, ya que cada curva solo brinda información de una categoría en específico y no del resto. Una consecuencia de ello es que pierden la característica de ser exclusivamente monótonas. Para el caso del 2PL-GRM, sólo las funciones para las categorías extremas negativas y positivas (la menor y mayor) son monótonas decrecientes y crecientes respectivamente. En la Figura 3.2 también se puede observar cambios específicos en las FRCI ante variaciones de los parámetros  $a_j$  y  $b_{jk}$ . Los ítems de la izquierda (ítems 1, 3 y 5) presentan un índice de discriminación bajo

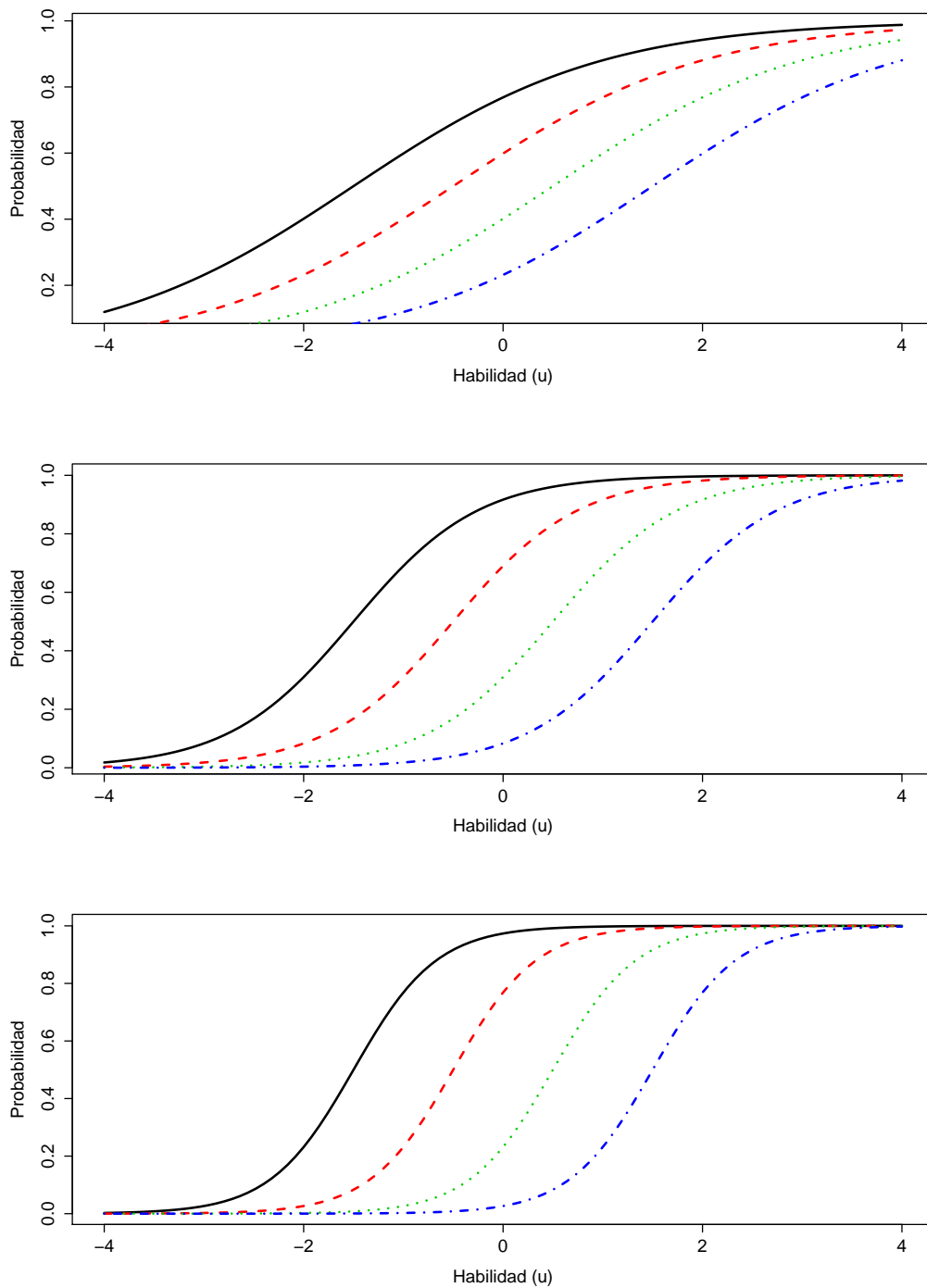


Figura 3.1: Representación gráfica de las FREI del Modelo 2PL-GRM para tres ítems de 5 categorías con parámetros de dificultad  $b = (-1.5, -0.5, 0.5, 1.5)$  y diferentes parámetros de discriminación  $a_j = 0.8, 1.6$  y  $2.4$

$a_j=0.8$  y los de la derecha un índice de discriminación alto  $a_j=2.4$  lo cual hace que las curvas se vuelvan más empinadas. Por otro lado, las dos primeras gráficas representan ítems con índices de dificultad más bajos  $b_{jk}=-3.0, -1.5, -0.5$  y  $1.0$  (ítems fáciles), los ítems 3 y 4 están más simétricamente distribuidos con parámetros de dificultad  $b_{jk}=-1.5, -0.5, 0.5$  y  $1.5$ ,

y los dos últimos ítems presentan índices de dificultad más altos  $b_{jk}=-1.0, 0.5, 1.5$  y  $3.0$  (ítems difíciles).

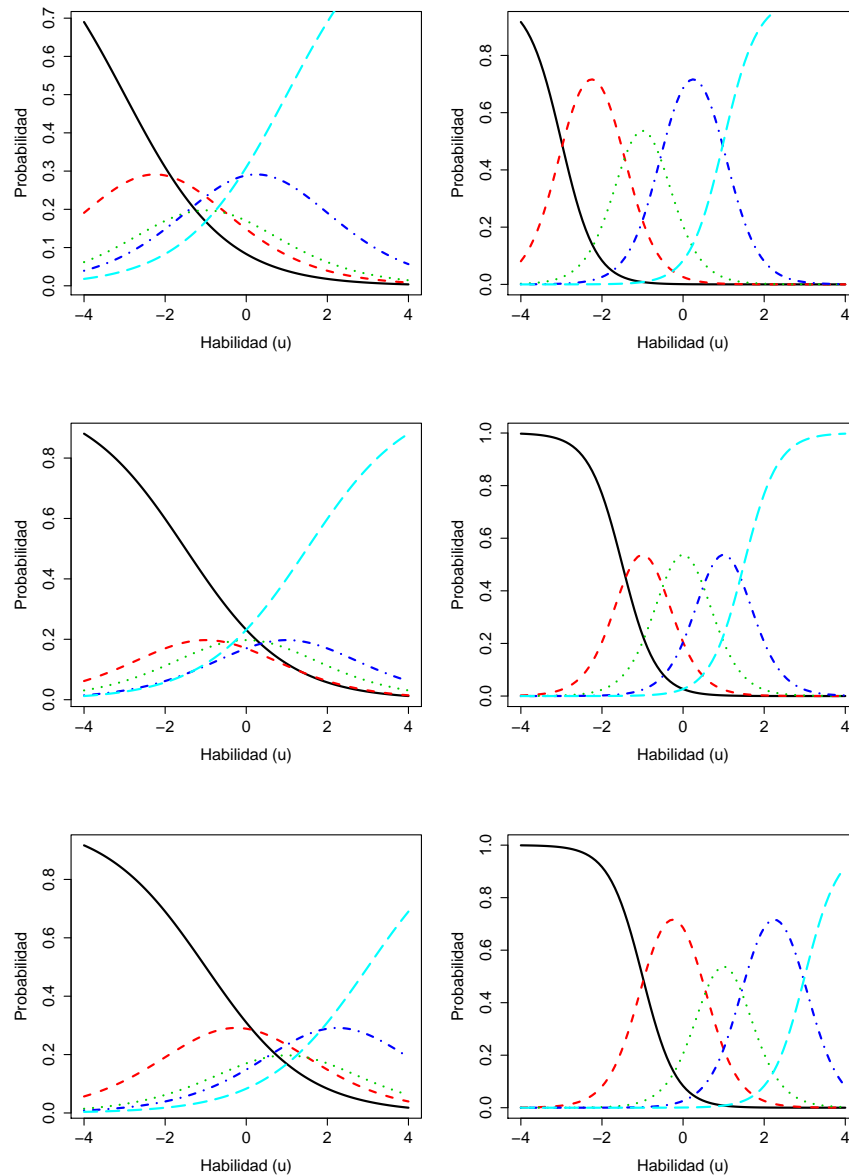


Figura 3.2: Representación gráfica de las FRCI del Modelo 2PL-GRM para seis ítems de 5 categorías con dos parámetros de discriminación  $a_j = 0.8$  y  $2.4$ , y tres configuraciones de parámetros de dificultad reflejando diferentes niveles: (i)  $b_{jk}=-3.0, -1.5, -0.5$  y  $1.0$  (ii)  $b_{jk}=-1.5, -0.5, 0.5$  y  $1.5$  y (iii)  $b_{jk}=-1.0, 0.5, 1.5$  y  $3.0$

### 3.1.3. Información de Fisher

Como se vio en la sección 2.5, la Función de Información de un Ítem (FII) para los GRM está dado por:

$$I_j(U) = \sum_{k=1}^m \frac{[P'_{jk}(u)]^2}{P_{jk}}$$

donde

$$\begin{aligned}
 P'_{jk}(u) &= \frac{\partial}{\partial u} [\Psi(\eta_{ijk}) - \Psi(\eta_{ij(k-1)})] \\
 &= \frac{\partial}{\partial u} \left[ \frac{1}{1 + e^{-\eta_{ijk}}} - \frac{1}{1 + e^{-\eta_{ij(k-1)}}} \right] \\
 &= \frac{\partial}{\partial u} \left[ \frac{1}{1 + e^{-a_j(b_{jk}-u)}} - \frac{1}{1 + e^{-a_j(b_{j(k-1)}-u)}} \right] \\
 &= -\frac{1}{(1 + e^{-a_j(b_{jk}-u)})^2} \frac{\partial}{\partial u} [e^{-a_j(b_{jk}-u)}] + \frac{1}{(1 + e^{-a_j(b_{j(k-1)}-u)})^2} \frac{\partial}{\partial u} [e^{-a_j(b_{j(k-1)}-u)}] \\
 &= -\frac{a_j e^{-a_j(b_{jk}-u)}}{(1 + e^{-a_j(b_{jk}-u)})^2} + \frac{a_j e^{-a_j(b_{j(k-1)}-u)}}{(1 + e^{-a_j(b_{j(k-1)}-u)})^2} \\
 &= -a_j \left[ \frac{e^{-a_j(b_{jk}-u)}}{1 + e^{-a_j(b_{jk}-u)}} \cdot \frac{1}{1 + e^{-a_j(b_{jk}-u)}} - \frac{e^{-a_j(b_{j(k-1)}-u)}}{1 + e^{-a_j(b_{j(k-1)}-u)}} \cdot \frac{1}{1 + e^{-a_j(b_{j(k-1)}-u)}} \right] \\
 &= -a_j \left[ \frac{1}{1 + e^{a_j(b_{jk}-u)}} \cdot \frac{1}{1 + e^{-a_j(b_{jk}-u)}} - \frac{1}{1 + e^{a_j(b_{j(k-1)}-u)}} \cdot \frac{1}{1 + e^{-a_j(b_{j(k-1)}-u)}} \right] \\
 &= -a_j \left[ \frac{1}{1 + e^{\eta_{ijk}}} \cdot \frac{1}{1 + e^{-\eta_{ijk}}} - \frac{1}{1 + e^{\eta_{ij(k-1)}}} \cdot \frac{1}{1 + e^{-\eta_{ij(k-1)}}} \right] \\
 &= -a_j \{ [\Psi(\eta_{ijk})] [1 - \Psi(\eta_{ijk})] - [\Psi(\eta_{ij(k-1)})] [1 - \Psi(\eta_{ij(k-1)})] \}
 \end{aligned}$$

Reemplazando el resultado de la expresión anterior en la Ecuación 2.31 tenemos que la FII para el Modelo 2PL-GRM está dado por:

$$I_j(U) = a_j^2 \sum_{k=1}^m \frac{\{ [\Psi(\eta_{ijk})] [1 - \Psi(\eta_{ijk})] - [\Psi(\eta_{ij(k-1)})] [1 - \Psi(\eta_{ij(k-1)})] \}^2}{\Psi(\eta_{ijk}) - \Psi(\eta_{ij(k-1)})} \quad (3.6)$$

En la Figura 3.3 se muestra la FII de los 6 ítems comentados en la Sección 3.1.2. Como se puede observar, aquellos ítems que tienen un parámetro de discriminación alto ( $a = 2.4$ ) son los que brindan mayor información.

La Función de Información del Test (FIT) estará definida por la Ecuación 2.32:

$$I(U) = \sum_{j=1}^K I_j(U)$$

La función para el error estándar estará dado por la Ecuación 2.33:

$$SE(U) = \left[ \sum_{j=1}^J I_j(U) \right]^{-\frac{1}{2}} = \frac{1}{\sqrt{I(U)}}$$

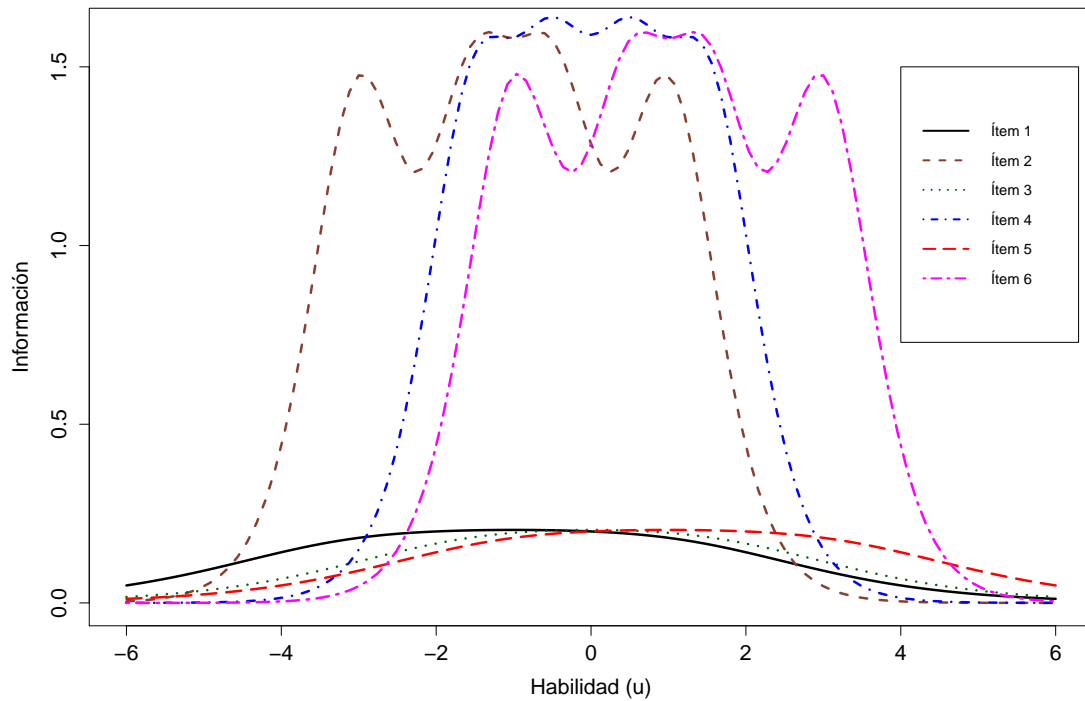


Figura 3.3: Representación gráfica de las FII del Modelo 2PL-GRM para seis ítems de 5 categorías con dos parámetros de discriminación  $a_j = 0.8$  (ítems impares) y  $2.4$  (ítems pares), y tres configuraciones de parámetros de dificultad reflejando diferentes niveles:  $b_{jk} = -3.0, -1.5, -0.5$  y  $1.0$  (ítems 1 y 2),  $b_{jk} = -1.5, -0.5, 0.5$  y  $1.5$  (ítems 3 y 4) y  $b_{jk} = -1.0, 0.5, 1.5$  y  $3.0$  (ítems 5 y 6)

### 3.2. Modelo de Respuesta Graduada Logístico de Exponente Positivo (LPE-GRM)

En el Modelo de Respuesta Graduada Logístico de Exponente Positivo fue propuesto por Samejima (2000).

#### 3.2.1. Definición del Modelo

Dado una matriz de datos  $\mathbf{Y} = [Y_{ij}]_{n \times J}$  que representa las respuestas ordinales de los  $n$  individuos evaluados en una prueba o cuestionario de  $J$  ítems con  $Y_{ij}$  definida como en 2.1. Se puede considerar al Modelo de Respuesta Graduada Logístico de Exponente Positivo (LPE-GRM) como un caso particular de un GRM asimétrico en el cual  $F(\cdot)$  es la fda de la distribución burr Tipo II estándar. El modelo LPE-GRM queda definido como:

$$Y_{ij} \mid u_i, \xi_j \sim \text{Categorica}(P_{ij1}, \dots, P_{ijm})$$

$$P_{ijk} \equiv P_{jk}(u_i) \equiv P(Y_{ij} = k \mid u_i, \xi_j) = P_{ijk}^{ac} - P_{ij(k-1)}^{ac}$$

$$P_{ijk}^{ac} \equiv P_{jk}^{ac}(u_i) \equiv P(Y_{ij} \leq k \mid u_i, \xi_j) = F_{d_j}(\eta_{ijk}) = [\Psi(\eta_{ijk})]^{d_j} \quad (3.7)$$

$$P_{ijk} \equiv P_{jk}(u_i) \equiv P(Y_{ij} = k | u_i, \xi_j) = \begin{cases} [\Psi(\eta_{ijk})]^{d_j} & \text{si } k = 1 \\ [\Psi(\eta_{ijk})]^{d_j} - [\Psi(\eta_{ij(k-1)})]^{d_j} & \text{si } 2 < k < m \\ 1 - [\Psi(\eta_{ij(k-1)})]^{d_j} & \text{si } k = m \end{cases} \quad (3.8)$$

$$\eta_{ijk} = a_j (b_{jk} - u_i)$$

$$-\infty = b_{j0} < b_{j1} < \dots < b_{j(m-1)} < b_{jm} = \infty$$

$$i = 1, \dots, n, j = 1, \dots, J, k = 1, \dots, m$$

donde  $P_{ijk}$  son las probabilidades condicionales de responder la categoría  $k$  en el ítem  $j$ ,  $P_{ijk}^{ac}$  las probabilidades condicionales acumuladas de responder en cualquier categoría inferior o igual a  $k$  en el ítem  $j$ ,  $\eta_{ijk}$  el predictor latente lineal relacionado con  $u_i \in \mathbb{R}$ ,  $\xi_j = (a_j, b_{j1}, \dots, b_{j(m-1)}, d_j)'$  es el vector de parámetros del ítem  $j$  relacionados con la inclinación ( $a_j$ ), posición ( $b_{jk}$ ) y asimetría ( $d_j$ ), y  $F_{d_j}(\cdot)$  la fda de la distribución burr Tipo II estándar:

$$F_{d_j}(\eta_{ijk}) = [\Psi(\eta_{ijk})]^{d_j} = \frac{1}{(1 + e^{-\eta_{ijk}})^{d_j}} \quad -\infty < \eta_{ijk} < \infty \quad (3.9)$$

### 3.2.2. Características del Modelo LPE-GRM

El modelo LPE-GRM puede verse como una generalización del Modelo 2PL-GR luego de introducir un parámetro de forma que controla la asimetría del ítem, además del parámetro de inclinación  $a_j$  y los parámetros de posición  $b_{j1}, \dots, b_{j(m-1)}$ .

Basándonos principalmente en [Bolfarine y Bazán \(2010\)](#) y [Bazán \(2011\)](#) en relación al Modelo LPE para ítems binarios y [Samejima \(2010\)](#) para el caso con respuestas graduadas, podemos resumir que el parámetro  $d_j$  presenta las siguientes características:

- Si  $0 < d_j < 1$ , para una categoría en particular, los individuos con niveles de habilidad muy bajo tienen mayor probabilidad de elegir las categorías superiores del ítem en relación al caso simétrico. Dada esta situación, si un test consiste de ítems con valores comunes de  $a_j$  y diferentes valores de  $b_{jk}$ , los individuos que eligen categorías más bajas de los ítems con menor dificultad son penalizados. Mientras más cercano a 0 esté  $d_j$ , las probabilidades de elegir categorías superiores son bonificadas y la habilidad es penalizada respecto al caso de ausencia de asimetría. El ítem es considerado **menos complejo**.
- Si  $1 < d_j < \infty$ , para una categoría en particular, los individuos con niveles de habilidad muy altos tienen una menor probabilidad de elegir las categorías superiores del ítem en relación al caso simétrico. Dada esta situación, si un test consiste de ítems con valores comunes de  $a_j$  y diferentes valores de  $b_{jk}$ , los individuos que seleccionan las categorías más altas de los ítems con mayor dificultad son recompensados. Cuanto más sea  $d_j > 0$ , las probabilidades de elegir categorías superiores son penalizadas y la habilidad es bonificada respecto al caso de ausencia de asimetría. El ítem es considerado

más complejo.

- Si  $d_j = 1$ , no se cumplen ninguno de los principios anteriores y las características del modelo son similares a las del caso logístico de dos parámetros.

Se entiende por complejidad de un ítem *al nivel de exigencia al realizar una tarea* (Samejima, 1997a, pag. 483) que está siendo medida por dicho ítem. Como se resalta en Bolfarine y Bazán (2010), una tarea se considera más compleja si requiere que el individuo realice o pase de manera sucesiva por una serie de subprocesos secuenciales para resolver finalmente de manera satisfactoria un problema y que eso se vea reflejado en la selección de las categorías superiores del ítem que se encarga de medir dicha tarea. Se espera que, si un proceso es más complicado y/o contiene un gran número de subprocesos secuenciales, la probabilidad condicional, dado un nivel de habilidad, para pasar de manera satisfactoria por el proceso será pequeña. Como se indica en Samejima (2000), se espera que tales diferencias en la probabilidad condicional sean más notorias para niveles de habilidad bajos. Por conveniencia, se denomina a esta característica del modelo como **complejidad del ítem**, para diferenciarla de la dificultad de un ítem, y tomar esta complejidad en cuenta.

### Función de Respuesta a las Etapas del Ítem (FREI)

De acuerdo a lo discutido en 2.4.2, las FREI para el modelo LPE-GRM estarán definidas como:

$$P_{jk}^+(u_i) = 1 - [\Psi(\eta_{ijk})]^{d_j} = 1 - \frac{1}{(1 + e^{-\eta_{ijk}})^{d_j}} \quad (3.10)$$

$$\eta_{ijk} = a_j (b_{jk} - u_i), \quad i = 1, \dots, n, \quad j = 1, \dots, J, \quad k = 1, \dots, m$$

Notar que a diferencia del Modelo 2PL-GRM,  $F_{d_j}(-\eta_{ijk}) \neq 1 - F_{d_j}(\eta_{ijk})$  debido a que no es una distribución simétrica. Para el Modelo LPE-GRM,  $d_j$  es conocido como un parámetro de “aceleración” (Samejima, 1995), en el sentido de que acelera el punto o valor de  $u$  en el cual la pendiente o poder de discriminación de la FREI se hace mayor. Es por esto que el parámetro  $d_j$  se encuentra asociado con la forma asimétrica de la FREI y presenta las siguientes características (Bolfarine y Bazán, 2010; Bazán, 2011):

- Si  $0 < d_j < 1$ , las FREI presentan asimetría negativa y siguen el principio de penalizar la respuesta de categorías inferiores de los ítems fáciles. Las curvas son trasladadas hacia la izquierda y encima respecto del caso de ausencia de asimetría.
- Si  $1 < d_j < \infty$ , las FREI presentan asimetría positiva y siguen el principio de bonificar la respuesta de las categorías superiores de los ítems difíciles. Las curvas son trasladadas a la derecha y abajo respecto del caso de ausencia de asimetría.
- Si  $d_j = 1$ , las FREI corresponden a las del Modelo 2PL-GR.

En relación al concepto de complejidad, una eventual FREI es asimétrica dependiendo de cuantas y que tan difíciles son la secuencia de subprocesos que están relacionados en la solución de un problema. Es importante notar que la palabra secuencia es usada por Samejima

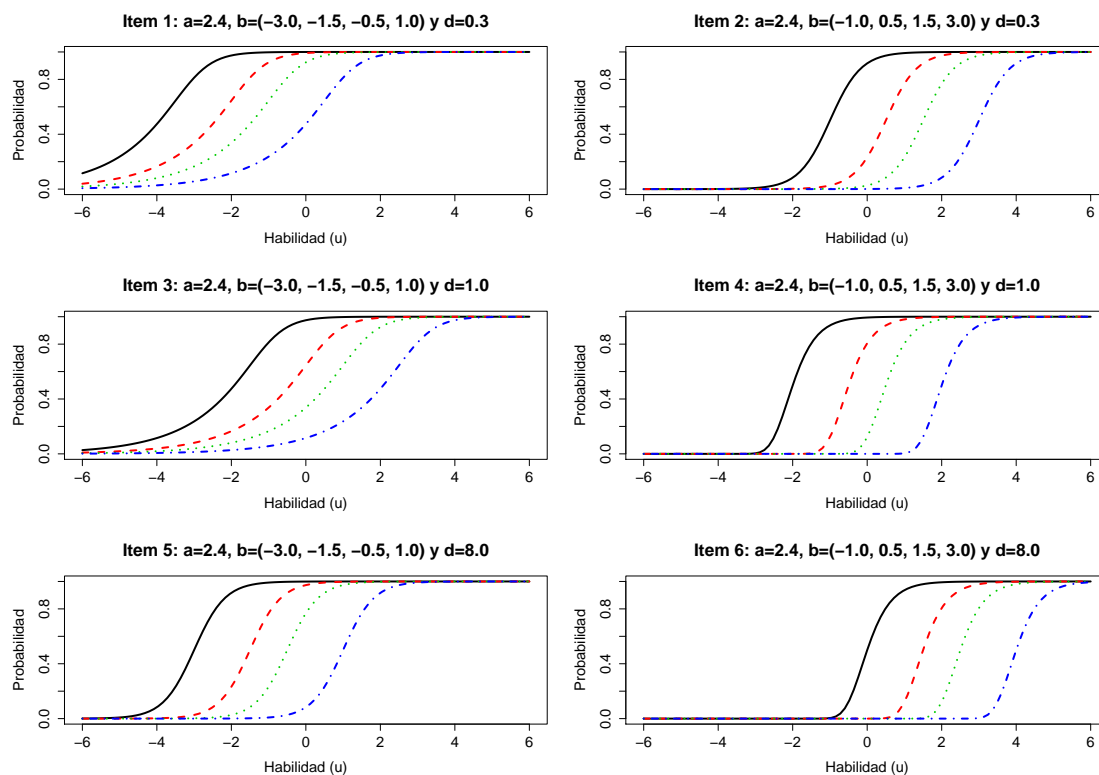


Figura 3.4: Representación gráfica de las FREI del Modelo LPE-GRM para seis ítems de 5 categorías con un parámetro de discriminación  $a_j=2.4$ , dos configuraciones de parámetros de dificultad reflejando ítems con baja y alta dificultad:  $b_{jk}=-3.0, -1.5, -0.5$  y  $1.0$  (ítems impares) y  $b_{jk}=-1.0, 0.5, 1.5$  y  $3.0$  (ítems pares), y distintos valores de asimetría  $d_j=0.3$  (ítems 1 y 2),  $1.0$  (ítems 3 y 4) y  $8.0$  (ítems 5 y 6)

en un sentido muy amplio, y los subprocesos puede ser realizado de en serie o paralelo como se detalla en Samejima (1995).

En la Figura 3.4 se muestra las FREI para seis ítems de 5 categorías graduadas con un mismo parámetro de discriminación  $a_j=2.4$ , dos configuraciones de parámetros de dificultad (i)  $b_{jk}=-3.0, -1.5, -0.5$  y  $1.0$  (ii)  $b_{jk}=-1.0, 0.5, 1.5$  y  $3.0$  reflejando ítems con baja y alta dificultad respectivamente y tres distintos valores de  $d_j=0.3, 1.0$  y  $8.0$  representando los casos asimétrico negativo, simétrico y simétrico positivo. Puede verse claramente que a medida que aumenta la complejidad del ítem las curvas se trasladan hacia la derecha y hacia abajo. En el Ítem 1 se observa como los individuos con niveles de habilidad más bajo tienen mayor probabilidad de elegir las categorías superiores del ítem en comparación al caso simétrico (Ítem 3) y con otro ítem de similar asimetría negativa pero mayor dificultad (Ítem 2). Por otro lado, el Ítem 6, muestra como los individuos con niveles de habilidad más altos tienen una menor probabilidad de elegir las categorías superiores del ítem en comparación con el caso simétrico (Ítem 4) y con otro ítem de similar asimetría positiva pero menor dificultad (Ítem 5).



### Función de Respuesta a las Categorías del Ítem (FRCI)

Siguiendo las definiciones dadas en 2.4.3, las FRCI para el Modelo LPE-GRM estarán definidas como:

$$P_{jk}(u_i) = \frac{1}{(1 + e^{-\eta_{ijk}})^{d_j}} - \frac{1}{(1 + e^{-\eta_{ij(k-1)}})^{d_j}} \quad (3.11)$$

$$\eta_{ijk} = a_j (b_{jk} - u_i), \quad i = 1, \dots, n, \quad j = 1, \dots, J, \quad k = 1, \dots, m$$

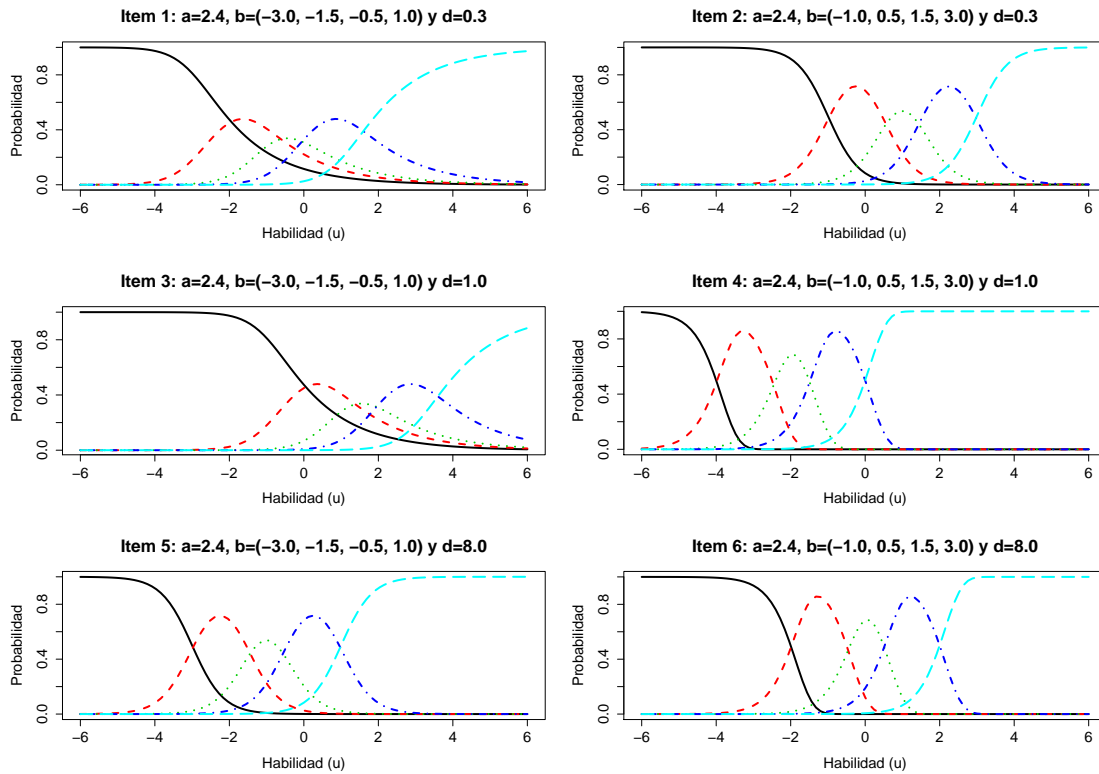


Figura 3.5: Representación gráfica de las FRCI del Modelo LPE-GRM para seis ítems de 5 categorías con un parámetro de discriminación  $a_j = 2.4$ , dos configuraciones de parámetros de dificultad reflejando ítems con baja y alta dificultad:  $b_{jk} = -3.0, -1.5, -0.5$  y  $1.0$  (ítems impares) y  $b_{jk} = -1.0, 0.5, 1.5$  y  $3.0$  (ítems pares), y distintos valores de asimetría  $d_j = 0.3$  (ítems 1 y 2),  $1.0$  (ítems 3 y 4) y  $8.0$  (ítems 5 y 6)

En la Figura 3.5 se muestra las FRCI para los seis ítems de 5 categorías graduadas del ejemplo de la sección anterior.

#### 3.2.3. Información de Fisher

Como se vio en la sección 2.5, la Función de Información de un Ítem (FII) para los GRM está dado por:

$$I_j(U) = \sum_{k=1}^m \frac{[P'_{jk}(u)]^2}{P_{jk}}$$

donde

$$\begin{aligned}
 P'_{jk}(u) &= \frac{\partial}{\partial u} \left\{ [\Psi(\eta_{ijk})]^{d_j} - [\Psi(\eta_{ij(k-1)})]^{d_j} \right\} \\
 &= d_j \left\{ [\Psi(\eta_{ijk})]^{d_j-1} \frac{\partial}{\partial u} \Psi(\eta_{ijk}) - [\Psi(\eta_{ij(k-1)})]^{d_j-1} \frac{\partial}{\partial u} \Psi(\eta_{ij(k-1)}) \right\} \\
 &= -a_j d_j \left\{ [\Psi(\eta_{ijk})]^{d_j} [1 - F(\eta_{ijk})] - [\Psi(\eta_{ij(k-1)})]^{d_j} [1 - \Psi(\eta_{ij(k-1)})] \right\}
 \end{aligned}$$

Reemplazando el resultado de la expresión anterior en la Ecuación 2.31 tenemos que la FII para el Modelo LPE-GRM está dado por:

$$I_j(U) = (a_j d_j)^2 \sum_{k=1}^m \frac{\left\{ [\Psi(\eta_{ijk})]^{d_j} [1 - \Psi(\eta_{ijk})] - [\Psi(\eta_{ij(k-1)})]^{d_j} [1 - \Psi(\eta_{ij(k-1)})] \right\}^2}{[\Psi(\eta_{ijk})]^{d_j} - [\Psi(\eta_{ij(k-1)})]^{d_j}} \quad (3.12)$$

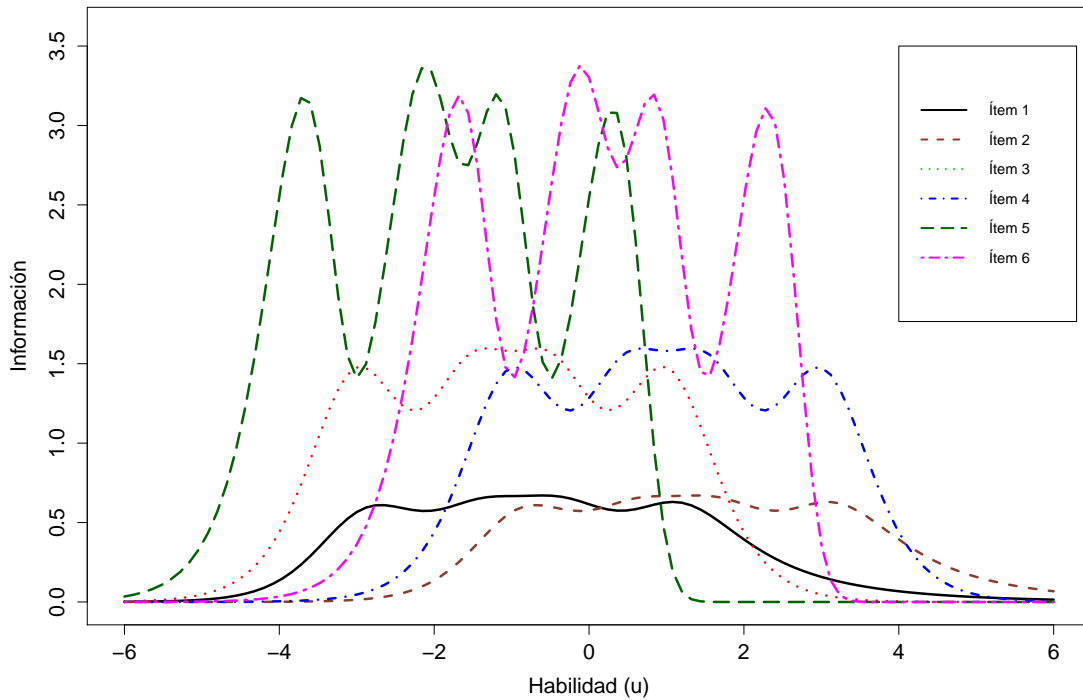


Figura 3.6: Representación gráfica de las FII del Modelo LPE-GRM para seis ítems de 5 categorías con un parámetro de discriminación  $a_j = 2.4$ , dos configuraciones de parámetros de dificultad reflejando ítems con baja y alta dificultad:  $b_{jk} = -3.0, -1.5, -0.5$  y  $1.0$  (ítems impares) y  $b_{jk} = -1.0, 0.5, 1.5$  y  $3.0$  (ítems pares), y distintos valores de asimetría  $d_j = 0.3$  (ítems 1 y 2),  $1.0$  (ítems 3 y 4) y  $8.0$  (ítems 5 y 6)

En la Figura 3.6 se muestra la FII de los 6 ítems comentados en la Sección 3.2.2.

La Función de Información del Test (FIT) estará definida por la Ecuación 2.32:

$$I(U) = \sum_{j=1}^K I_j(U)$$

La función para el error estándar estará dado por la Ecuación 2.33:

$$SE(U) = \left[ \sum_{j=1}^J I_j(U) \right]^{-\frac{1}{2}} = \frac{1}{\sqrt{I(U)}}$$



## Capítulo 4

# Estimación Bayesiana del Modelo de Respuesta Graduada

La estimación de los modelos de la TRI pueden ser agrupada en tres categorías: estimación clásica o frecuentista, estimación bayesiana y estimación con MCMC. En relación a la estimación clásica, los métodos de Máxima Verosimilitud Conjunta (JML), Máxima Verosimilitud Condicional (CML) y Máxima Verosimilitud Marginal (MML) son los métodos más usados para la estimación de los parámetros de los ítems (Johnson, 2007). Para mayor detalles sobre la estimación clásica del GRM puede consultarse a Azevedo (2003); Baker y Kim (2004). En el presente capítulo se desarrolla la metodología de inferencia Bayesiana usando métodos de MCMC. También son descritos los criterios de selección de modelos considerando el DIC (criterio de información de desvío), AIC y BIC, así como la implementación computacional necesaria para la estimación de los modelos usando R y WinBUGS.

### 4.1. Estimación Bayesiana del Modelo de Respuesta Graduada Logístico (2PL-GRM)

#### 4.1.1. Función de Verosimilitud para el Modelo 2PL-GRM

Dado una matriz de datos  $\mathbf{Y} = [Y_{ij}]_{n \times J}$  que representa las respuestas ordinales de los  $n$  individuos evaluados en una prueba o cuestionario de  $J$  ítems con  $Y_{ij}$  definida como en 2.1,  $W_{ijk}$  la parametrización alternativa dada en 2.8,  $\mathbf{U} = (U_1, \dots, U_n)$  el vector de parámetros para las personas y  $\xi = (\xi_1, \dots, \xi_J)$  el vector de parámetros de todos los ítems con  $\xi_j = (a_j, b_{j1}, \dots, b_{j(m-1)})'$ . Asumiendo que se cumple (S1), (S2) y (S3), la función de verosimilitud puede ser escrita como:

$$L(\mathbf{y} \mid \mathbf{u}, \xi) = f(\mathbf{y} \mid \mathbf{u}, \mathbf{a}, \mathbf{b}) = \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m [P_{ijk}]^{w_{ijk}} \quad (4.1)$$

$$= \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m [\Psi(\eta_{ijk}) - \Psi(\eta_{ij(k-1)})]^{w_{ijk}} \quad (4.2)$$

$$= \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m \left[ \frac{1}{1 + e^{-a_j(b_{jk} - u_i)}} - \frac{1}{1 + e^{-a_j(b_{j(k-1)} - u_i)}} \right]^{w_{ijk}} \quad (4.3)$$

$$= \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m \left[ \frac{1}{1 + e^{-(\kappa_{jk} - a_j u_i)}} - \frac{1}{1 + e^{-(\kappa_{j(k-1)} - a_j u_i)}} \right]^{w_{ijk}} \quad (4.4)$$

#### 4.1.2. Especificación de las prioris en el modelo 2PL-GRM

Para la especificación de la distribución a priori conjunta de los parámetros de las personas y de los ítems se asume que existe una estructura de prioris independientes del siguiente modo:

$$P(\mathbf{u}, \mathbf{a}, \mathbf{b} | \mathbf{y}) = P(\mathbf{u}) P(\mathbf{a}, \mathbf{b} | \mathbf{w}) = \prod_{i=1}^n P(u_i) \prod_{j=1}^J P(a_j) \prod_{k=1}^m P(\kappa_{jk}) \quad (4.5)$$

Es recomendable definir prioris informativas porque la existencia de la distribución conjunta a posteriori no esta garantizada cuando una priori impropia es usada. A continuación se detallan las prioris propuestas:

- Para el vector de parámetros de las personas (habilidades) [Bazán \(2005\)](#) considera una distribución normal e independiente:

$$u_i \sim N(0, 1) \quad (4.6)$$

Con lo que se espera obtener  $E[u_i] = 0$  y  $V[u_i] = 1$

- El parámetro de discriminación tiene una distribución a priori:

$$a_j \sim N_+(\mu_a, \sigma_a^2) \quad (4.7)$$

donde  $\mu_a$  y  $\sigma_a^2$  son constantes especificadas antes del análisis,  $N_+(\mu_a, \sigma_a^2)$  denota una distribución normal truncada positiva con parámetros de locación  $\mu_a$  y escala  $\sigma_a$ . [Sahu \(2002\)](#) y [Patz y Junker \(1999\)](#) recomiendan usar los valores de  $\mu_a = 1$  y  $\sigma_a^2 = 0.5$  para las constantes anteriores. Con estos valores se espera obtener  $E[a_j] = 1 + \sqrt{0.5} \cdot \frac{\phi(-\sqrt{2})}{1 - \Phi(-\sqrt{2})} = 1.1126$  y  $V[a_j] = 0.5 \left\{ 1 - \frac{\phi(-\sqrt{2})}{1 - \Phi(-\sqrt{2})} \left[ \frac{\phi(-\sqrt{2})}{1 - \Phi(-\sqrt{2})} + \sqrt{2} \right] \right\} = 0.3747$ .

- Las prioris para los parámetros de transición deben registrar la restricción de orden  $\kappa_{j1} < \dots < \kappa_{j,m-1}$ . Por lo anterior, estas pueden ser inducidas definiendo parámetros auxiliares sin restricciones  $\kappa_{j1}^*, \dots, \kappa_{j,m-1}^*$  tales que:

$$\kappa_{jk}^* \sim N(\mu_k, \sigma_k^2) \quad (4.8)$$

para  $k = 1, \dots, m - 1$ . Las distribuciones a priori de los parámetros de transición para el  $j$ -ésimo ítem son entonces obtenidos haciendo que  $\kappa_{jk}$  sea igual a la  $k$ -ésima estadística de orden de las variables auxiliares  $\kappa_{j1}^*, \dots, \kappa_{j,m-1}^*$  para el  $j$ -ésimo ítem, es decir:

$$\kappa_{jk} = \kappa_{j,[k]}^* \quad (4.9)$$

Se propone usar los valores de  $\mu_k = 0$  y  $\sigma_k^2 = 2$  para las constantes anteriores similar a los que [Sahu \(2002\)](#) y [Patz y Junker \(1999\)](#) usaron para los parámetros  $b_{jk}$ , con lo que se espera obtener  $E[\kappa_{jk}^*] = 0$  y  $V[\kappa_{jk}^*] = 2$

En el Cuadro 4.1 se muestra una lista de prioris para los parámetros de los ítems usadas en otros estudios relacionados con el Modelo 2PL-GRM. La notación  $I(b_{jk})$  indica que el parámetro  $b_{j(k+1)}$  fue muestreado de tal manera que sea mayor a  $b_{jk}$  como se indica en la condición 2.5 del GRM.

Autor	Prioris
Zhu y Stone (2012)	$a_j \sim \ln N(0, 1); b_{j1} \sim N(0, 4); b_{j(k+1)} \sim N(0, 4) I(b_{jk})$
Kieftenbeld y Natesan (2012)	$a_j \sim \ln N(0, 2); b_{j1} \sim U(-5, 5); b_{j(k+1)} \sim U(-5, 5) I(b_{jk})$
Curtis (2010)	$a_j \sim N^+(1, 2.5^2); \kappa_{jk}^* \sim N(0, 2.5^2), \kappa_{jk} = \kappa_{j,[k]}^*$

Cuadro 4.1: Especificación de prioris para los parámetros del ítem en el modelo 2PL-GRM

#### 4.1.3. Distribución a posteriori de los parámetros del modelo 2PL-GRM

Considerando la función de verosimilitud dada en 4.4 y las prioris consideradas en 4.7, 4.8 y 4.9, definimos la distribución a posteriori de los parámetros dado los datos como:

$$P(\mathbf{u}, \mathbf{a}, \mathbf{b} \mid \mathbf{y}) \propto P(\mathbf{u}, \mathbf{a}, \mathbf{b}) f(\mathbf{y} \mid \mathbf{u}, \mathbf{a}, \mathbf{b}) \quad (4.10)$$

$$\propto P(\mathbf{u}, \mathbf{a}, \mathbf{b}) \times \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m [\Psi(\eta_{ijk}) - \Psi(\eta_{ij(k-1)})]^{w_{ijk}} \quad (4.11)$$

$$\propto P(\mathbf{u}, \mathbf{a}, \mathbf{b}) \times \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m \left[ \frac{1}{1 + e^{-(\kappa_{jk} - a_j u_i)}} - \frac{1}{1 + e^{-(\kappa_{j(k-1)} - a_j u_i)}} \right]^{w_{ijk}} \quad (4.12)$$

#### 4.1.4. Inferencia Bayesiana del modelo 2PL-GRM

La distribución a posteriori dada en 4.12 no es identificable. Para la estimación bayesiana se consideran por tanto métodos de generación de distribuciones a partir de Cadenas de Markov de Monte Carlo (MCMC), específicamente usando el algoritmo de Metropolis-Hastings, que es implementado por defecto en el software WinBUGS, para lo cual se escribe el código del modelo, teniendo la especificación de la verosimilitud jerárquica del modelo general:

$$y \mid u_i, a_j \kappa_{jk} \sim \text{Categórica}(\Psi(\kappa_{jk} - a_j u_i) - \Psi(\kappa_{j(k-1)} - a_j u_i)) \quad (4.13)$$

$$u_i \sim N(0, 1) \quad (4.14)$$

$$a_j \sim N_+(1, 0.5) \quad (4.15)$$

$$\kappa_{jk}^* \sim N(0, 2) \quad (4.16)$$

## 4.2. Estimación Bayesiana del Modelo de Respuesta Graduada Logístico con Exponente Positivo (LPE-GRM)

### 4.2.1. Función de Verosimilitud para el modelo LPE-GRM

Dado una matriz de datos  $\mathbf{Y} = [Y_{ij}]_{n \times J}$  que representa las respuestas ordinales de los  $n$  individuos evaluados en una prueba o cuestionario de  $J$  ítems con  $Y_{ij}$  definida como en 2.1,  $W_{ijk}$  la parametrización alternativa dada en 2.8,  $\mathbf{U} = (U_1, \dots, U_n)$  el vector de parámetros para las personas y  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_J)$  el vector de parámetros de todos los ítems con  $\xi_j = (a_j, b_{j1}, \dots, b_{j(m-1)}, d_j)'$ . Asumiendo que se cumple (S1), (S2) y (S3), la función de verosimilitud puede ser escrita como:

$$L(\mathbf{y} | \mathbf{u}, \boldsymbol{\xi}) = f(\mathbf{y} | \mathbf{u}, \mathbf{a}, \mathbf{b}, \mathbf{d}) = \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m [P_{ijk}]^{w_{ijk}} \quad (4.17)$$

$$= \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m \left[ [\Psi(\eta_{ijk})]^{d_j} - [\Psi(\eta_{ij(k-1)})]^{d_j} \right]^{w_{ijk}} \quad (4.18)$$

$$= \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m \left[ \frac{1}{(1 + e^{-a_j(b_{jk} - u_i)})^{d_j}} - \frac{1}{(1 + e^{-a_j(b_{j(k-1)} - u_i)})^{d_j}} \right]^{w_{ijk}} \quad (4.19)$$

$$= \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m \left[ \frac{1}{(1 + e^{-(\kappa_{jk} - a_j u_i)})^{d_j}} - \frac{1}{(1 + e^{-(\kappa_{j(k-1)} - a_j u_i)})^{d_j}} \right]^{w_{ijk}} \quad (4.20)$$

### 4.2.2. Especificación de las prioris en el modelo LPE-GRM

Para la especificación de la distribución a priori conjunta de los parámetros de las personas y de los ítems se asume que existe una estructura de prioris independientes del siguiente modo:

$$P(\mathbf{u}, \mathbf{a}, \mathbf{b}, \mathbf{d} | \mathbf{y}) = P(\mathbf{u}) P(\mathbf{a}, \mathbf{b}, \mathbf{d} | \mathbf{w}) = \prod_{i=1}^n P(u_i) \prod_{j=1}^J P(a_j) P(d_j) \prod_{k=1}^m P(\kappa_{jk}) \quad (4.21)$$

Al igual que para el modelo 2PL-GRM, es recomendable definir prioris informativas y se usan las mismas prioris para  $u_i$ ,  $a_j$  y  $\kappa_{jk}$ . Para el parámetro de asimetría se define:

$$d_j = e^{g_j} \quad (4.22)$$

donde

$$g_j \sim N(\mu_g, \sigma_g^2) \quad (4.23)$$

Usando los valores de  $\mu_g = 0$  y  $\sigma_a^2 = 0.5$  se espera obtener  $E[d_j] = e^{2 \times 0 + \frac{0.5}{2}} = 1.2840$  y  $V[d_j] = (e^{0.5} - 1) e^{0.5} = 1.0696$

#### 4.2.3. Distribución a posteriori de los parámetros del modelo LPE-GRM

Considerando la función de verosimilitud dada en 4.20 y las prioris consideradas en 4.7, 4.8, 4.9 y 4.23, definimos la distribución a posteriori de los parámetros dado los datos como:

$$P(\mathbf{u}, \mathbf{a}, \mathbf{b}, \mathbf{d} | \mathbf{y}) \propto P(\mathbf{u}, \mathbf{a}, \mathbf{b}, \mathbf{d}) f(\mathbf{y} | \mathbf{u}, \mathbf{a}, \mathbf{b}, \mathbf{d}) \quad (4.24)$$

$$\propto P(\mathbf{u}, \mathbf{a}, \mathbf{b}, \mathbf{d}) \times \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m \left[ [\Psi(\eta_{ijk})]^{d_j} - [\Psi(\eta_{ij(k-1)})]^{d_j} \right]^{w_{ijk}} \quad (4.25)$$

$$\propto P(\mathbf{u}, \mathbf{a}, \mathbf{b}, \mathbf{d}) \times \prod_{i=1}^n \prod_{j=1}^J \prod_{k=1}^m \left[ \frac{1}{\left(1 + e^{-(\kappa_{jk} - a_j u_i)}\right)^{d_j}} - \frac{1}{\left(1 + e^{-(\kappa_{j(k-1)} - a_j u_i)}\right)^{d_j}} \right]^{w_{ijk}} \quad (4.26)$$

#### 4.2.4. Inferencia Bayesiana del modelo LPE-GRM

La distribución a posteriori dada en 4.26 no es identificable. Para la estimación bayesiana se consideran por tanto métodos de generación de distribuciones a partir de Cadenas de Markov de Monte Carlo (MCMC), específicamente Metropolis-Hastings, que es implementado por defecto en el software WinBUGS, para lo cual se escribe el código del modelo, teniendo la especificación de la verosimilitud jerárquica del modelo general:

$$y | u_i, a_j, \kappa_{jk} \sim \text{Categórica} \left( [\Psi(\kappa_{jk} - a_j u_i)]^{d_j} - [\Psi(\kappa_{j(k-1)} - a_j u_i)]^{d_j} \right) \quad (4.27)$$

$$u_i \sim N(0, 1) \quad (4.28)$$

$$a_j \sim N_+(1, 0.5) \quad (4.29)$$

$$\kappa_{jk}^* \sim N(0, 2) \quad (4.30)$$

$$d_j \sim \ln N(0, 0.5) \quad (4.31)$$

#### 4.3. Criterios de Comparación de Modelos

El principal criterio usado en este trabajo es el criterio de información de desvío (DIC, por las siglas en inglés de Deviance Information Criterion) propuesto por Spiegelhalter et al. (2002) debido a que es el criterio más usado para comparar modelos con métodos Bayesianos y es accesible a todos los investigadores al usar el WinBUGS. El DIC es un índice basado en la información similar a otros indicadores como el criterio de información bayesiana de Schwarz (BIC, Bayesian Information Criterion), y el criterio de Akaike (AIC, Akaike Information Criterion), que se caracteriza por ponderar tanto el ajuste del modelo como la complejidad en



identificar un modelo específico. El DIC es un indicador ampliamente usado en la estimación Bayesiana con métodos MCMC porque es relativamente fácil de calcular, interpretar y es satisfactorio para los modelos jerárquicos complicados. El DIC es obtenido por defecto con el WinBUGS y es aproximado a través de:

$$\widehat{DIC} = Dbar - Dhat = Dbar + 2p\widehat{D} \quad (4.32)$$

donde  $pD$  es el número efectivo de parámetros,  $Dbar$  es el esperado del desvío bayesiano a posteriori o desvío medio a posteriori (posterior mean of the deviance) y  $Dhat$  es el desvío bayesiano de las medias a posteriori o desvío bayesiano de los esperados a posteriori (deviance of posterior mean).

Los esperados del AIC y del BIC también son posibles de hallar a través de:

$$E\widehat{AIC} = Dbar + 2p$$

$$E\widehat{BIC} = Dbar + p \log(N)$$

donde  $p$  es el número de parámetros en el modelo,  $N$  es el número total de observaciones. Es importante resaltar que para el caso de los modelos estudiados,  $N = n * J$  y no  $N = n$  debido al supuesto de independencia. Estos indicadores son difíciles de estimar dado que el número actual de parámetros no puede ser claramente identificado, por lo cual en la estimación por MCMC es preferible usar el DIC. De manera similar al AIC y el BIC, valores pequeños del DIC indicarán un mejor ajuste.

#### 4.4. Implementación Computacional

Para realizar la inferencia bayesiana bajo MCMC de los modelos descritos se programaron códigos en los software WinBUGS 1.4.3 (Lunn et al., 2000) y R (R Core Team, 2012).

El código en R mostrado en el Anexo A.2 contiene a la función `grmMCMC()` que fue programada para que facilite la preparación de los datos que serán utilizados por el WinBUGS en los estudios de simulación y de aplicación que se realizarán en el presente trabajo, así como también para que permita configurar de manera sencilla los parámetros para las simulaciones y los valores de los parámetros de las distribuciones a priori consideradas en los distintos modelos. Este código llama a la función `bugs()` de la librería `R2WinBUGS` (Sturtz et al., 2005) con el fin de que envíe los datos y los parámetros necesarios para realizar el análisis en WinBUGS. Librerías adicionales requeridas por el R son automáticamente cargadas (p. ej., la librería `coda` es requerida por el `R2WinBUGS`). Una ventaja de esta librería es que muestra automáticamente dos estadísticas para cada parámetro estimado: (1) el tamaño efectivo de la muestra (tamaño de muestra ajustado por autocorrelación a lo largo de las simulaciones; y (2) el diagnóstico de convergencia de Gelman y Rubin conocido como `Rhat`.

Los códigos en BUGS para realizar la inferencia bayesiana bajo MCMC en WinBUGS de los modelos 2PL-GRM y LPE-GRM fueron adaptados de Curtis (2010) y se muestran en los Anexos A.3 y A.4. Ambos códigos pueden ser ejecutados directamente en WinBUGS. En general se requiere un número mayor de iteraciones para el modelo LPE-GRM pues los

parámetros muestran mayor autocorrelación.

Adicionalmente, en el Anexo [A.1](#) se muestra un código para simular datos provenientes del Modelo de Respuesta Graduada, mientras que en los Anexos [A.7](#), [A.8](#) y [A.9](#) se encuentran los códigos para poder graficar la Función de Respuesta a las Etapas de los Ítems (FREI), la Función de Respuesta a las Categorías de los Ítems (FRCI) y la Función de Información de los Ítems (FII) respectivamente.



## Capítulo 5

### Estudio de Simulación

#### 5.1. Objetivos

- Simular un conjunto de datos politómicos ordinales generado a partir del Modelo de respuesta Graduada Logístico (2PL-GRM).
- Evaluar la precisión del método MCMC en la recuperación de los parámetros de los ítems y las habilidades del Modelo de respuesta Graduada Logístico (2PL-GRM).

#### 5.2. Algoritmo para simular los datos

Para simular una matriz de datos politómicos ordinales seguimos el siguiente procedimiento:

1. Definir la cantidad de sujetos  $n$ , cantidad de ítems  $J$  y el número de categorías de respuesta por cada ítem  $m$ .
2. Generar los parámetros del modelo. Se simulan los parámetros para las personas tal que  $u_i \sim N(0, 1)$  y los parámetros de los ítems considerando  $a_j \sim N_+(1, 0.5)$  y  $b_{jk} \sim N(0, 2)$  o, en caso de ser posible, son fijados utilizando estimaciones obtenidas en ítems con instrumentos ya calibrados con el fin de asegurar que la simulación sea más real (Harwell et al., 1996).
3. Calcular para cada uno de los  $J$  ítems las probabilidades condicionales de responder en cualquier categoría inferior o igual a  $k$  ( $P_{ijk}^{ac}$ ), y las probabilidades condicionales de responder en la categoría  $k$  ( $P_{ijk}$ ) de acuerdo a 3.1 y 3.2.
4. Simular para cada uno de los  $n$  individuos la respuesta en cada ítem. Para ello se toma una muestra aleatoria simple de una observación para una población de  $m$  elementos con probabilidades iguales a las calculadas en el paso anterior ( $P_{ijk}$ ).

En el proceso de simulación los parámetros  $a_j$  y  $b_{jk}$  quedan fijos a lo largo de todas las réplicas de la simulación con el fin de minimizar el error experimental (Kieftenbeld y Natesan, 2012). Un nuevo conjunto de habilidades  $u_i$  son generadas para cada réplica, con las que las respuestas de los individuos son simuladas. El código para simular respuestas politómicas ordinales basadas en un Modelo 2PL-GRM, adaptado de (Rizopoulos, 2006), se muestra en el Anexo A.1.

### 5.3. Método para Estimar los parámetros

La estimación de los parámetros se realiza usando la inferencia bayesiana mediante la estimación MCMC implementada en el programa WinBUGS 1.4.3 (Lunn et al., 2000) usando el algoritmo de Metropolis-Hastings. Con este método se calcularán las medias estimadas de las distribuciones a posteriori de los parámetros a estimar. Se estiman al menos  $mJ + n$  parámetros en cada replicación de la simulación.

La implementación computacional para la inferencia bayesiana del Modelo 2PL-GRM en WinBUGS se encuentra en el Anexo A.3 y la función en R usando la librería R2WinBUGS que permite ejecutarla en el Anexo A.2.

### 5.4. Criterios para evaluar la simulación

La precisión en la recuperación de los parámetros es evaluada usando la raíz del error cuadrático medio (RMSE) y el promedio del error absoluto (MAE). El RMSE es definido como la raíz cuadrada del promedio de las diferencias entre los valores reales y los estimados a lo largo de todas las réplicas:

$$RMSE(\lambda) = \sqrt{\frac{1}{R} \sum_{s=1}^R (\hat{\lambda}_s - \lambda_s)^2} \quad (5.1)$$

donde  $\lambda$  representa tanto a los parámetros de los ítems (discriminación o dificultad) como a los parámetros para las personas (habilidad),  $\hat{\lambda}$  es el estimado del parámetro anterior y  $R$  el número de réplicas en la simulación.

Del mismo modo, el MAE es definido como :

$$MAE(\lambda) = \frac{1}{R} \sum_{s=1}^R |\hat{\lambda}_s - \lambda_s| \quad (5.2)$$

Las estimaciones de los parámetros de los ítems son calculados para cada uno de los  $J$  ítems y en el caso de los parámetros de las personas para cada uno de los  $n$  individuos.

Dentro de cada réplica también son calculadas para cada tipo de parámetro las correlaciones entre los valores estimados y los parámetros generados.

### 5.5. Escenario

Se consideran  $J = 20$  ítems con  $m = 5$  categorías de respuestas para  $n = 500$  individuos. Los parámetros de los ítems fueron fijados de acuerdo a lo recomendado por Kieftenbeld y Natesan (2012) usando los valores obtenidos de un conjunto de ítems calibrados del *Minnesota Satisfaction Questionnaire* usado por este mismo autor y por Lautenschlager et al. (2006). Los parámetros de los ítems son mostrados en el Cuadro 5.1.

La cantidad de réplicas del estudio de simulación es de  $R = 20$  considerando diferentes conjuntos de datos para cada una de ellas.

En el Anexo A.5 se muestra la implementación computacional del presente estudio de simulación así como la tablas con el resumen de los resultados obtenidos.

Ítem	b1	b2	b3	b4	a
1	-4.26	-2.90	-1.25	2.01	0.95
2	-2.45	-1.44	-0.60	1.45	1.48
3	-2.07	-1.27	0.16	2.11	1.46
4	-1.75	-0.76	0.13	2.02	1.49
5	-2.19	-1.27	-0.35	1.52	1.38
6	-2.88	-1.97	-0.51	1.87	1.35
7	-3.77	-2.23	-1.27	1.34	0.96
8	-3.24	-2.29	-0.49	1.93	1.32
9	-3.28	-2.09	0.49	3.09	1.08
10	-1.57	-0.75	-0.13	1.68	2.00
11	-1.39	0.08	1.07	2.99	1.22
12	-2.97	-1.50	-0.41	2.44	0.89
13	-2.05	-1.19	-0.15	1.87	2.05
14	-1.20	-0.24	0.61	2.48	1.59
15	-1.68	-0.95	-0.25	1.69	2.31
16	-1.90	-1.08	-0.39	1.59	2.07
17	-1.80	-0.80	0.10	1.96	1.55
18	-3.82	-2.63	-1.20	1.67	0.92
19	-1.40	-0.50	0.35	2.10	1.64
20	-1.70	-0.90	-0.06	1.81	2.35

Cuadro 5.1: Parámetros de los Ítems usados en la simulación

## 5.6. Resultados

Para obtener los resultados de las estimaciones de los parámetros en el WinBUGS para cada réplica, se realizaron estudios preliminares con el fin de determinar que tan larga debía ser la cadena para asegurar la convergencia y cuantas iteraciones eran necesarias. Finalmente se consideraron 2 cadenas con 26500 iteraciones cada una, de las cuales fueron descartadas (*burning*) las 6500 primeras y para las 20000 restantes se consideró saltos (*thin*) de 20 en cada cadena para reducir la autocorrelación entre iteraciones, quedando finalmente 1000 iteraciones efectivas por cadena con las que las distribuciones a posteriori de los parámetros fueron estimadas. Los resultados del sesgo de las medias de las estimaciones de los parámetros de los ítems recuperados se muestran en el Cuadro 5.2

Las medidas de RMSE y MAE para las estimaciones de los parámetros de cada ítem son mostrada en el Cuadro 5.3. Todos estos resultados indican que los parámetros de los ítems fueron recuperados de manera satisfactoria usando la estimación MCMC implementada en WinBUGS.

Finalmente se evaluó convergencia de la distribución a posteriori de los parámetros hacia una distribución estacionaria. La convergencia fue examinada visualmente a través de varios gráficos de diagnóstico disponibles en WinBUGS que fueron obtenidos con las librerías `mcmcplots` y `R2WinBUGS` del programa R. Una réplica distinta fue utilizada para evaluar cada criterio de convergencia y parte de los resultados son mostrados en el Anexo B.1. El primer gráfico muestra el histórico de la cadena de los resultados del muestreo (*sampling history plot*) para cada parámetro. La Figura B.1 ilustra el histórico para la pendiente del cuarto ítem y el

Ítem	b1	b2	b3	b4	a
1	-0.128	-0.049	0.009	0.119	-0.030
2	-0.052	0.016	0.056	0.087	-0.052
3	-0.042	0.049	0.053	0.091	-0.032
4	0.050	0.051	0.073	0.036	-0.008
5	0.002	0.021	0.068	0.113	-0.035
6	-0.066	0.006	0.067	0.171	-0.069
7	0.130	0.120	0.110	0.068	0.017
8	-0.060	-0.013	0.022	0.102	-0.060
9	-0.004	0.060	0.077	0.071	-0.017
10	0.022	0.057	0.085	0.108	-0.076
11	0.005	0.078	0.122	0.098	-0.033
12	-0.043	0.021	0.047	0.148	-0.018
13	0.035	0.035	0.078	0.081	-0.098
14	0.027	0.037	0.032	0.061	-0.016
15	0.036	0.071	0.090	0.102	-0.111
16	0.033	0.060	0.083	0.090	-0.094
17	0.064	0.036	0.054	0.076	-0.028
18	-0.017	0.001	0.027	0.072	0.001
19	0.014	0.082	0.078	0.091	-0.010
20	0.001	0.015	0.059	0.087	-0.157

Cuadro 5.2: Sesgo en la recuperación de parámetros de los ítems usando estimación bajo MCMC para el Modelo 2PL-GRM

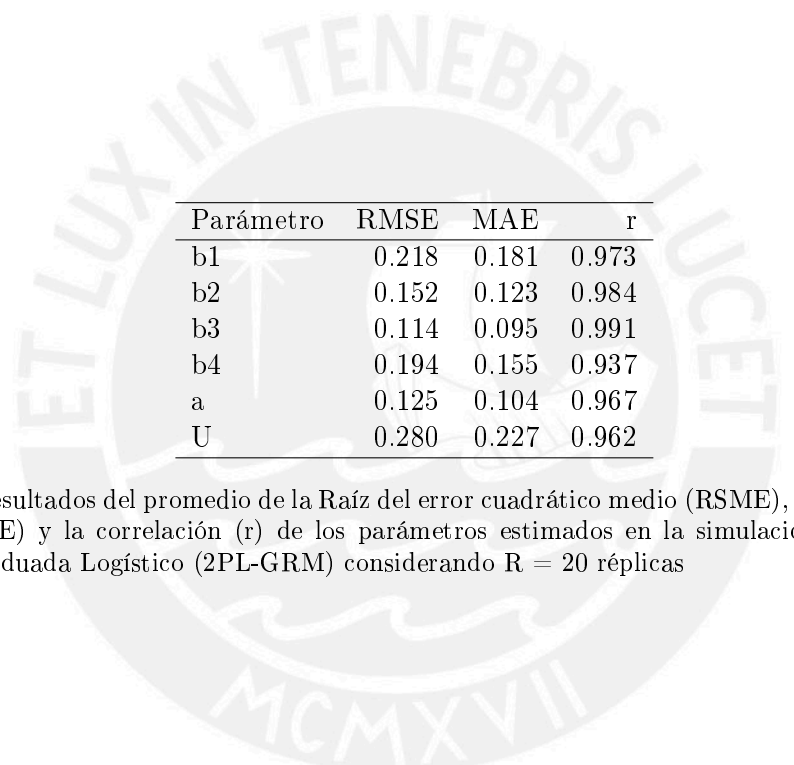
primer umbral. Este histórico muestra que cada cadena considerada converge rápidamente a una distribución estacionaria puesto que se observa que están sobrepuestas. Similares resultados fueron obtenidos para los ítems restantes. En la Figura B.2 se muestra la gráfica de las densidades y de autocorrelaciones para los parámetros del Ítem 4. Se aprecia que las distribuciones para ambas cadenas convergen y que no existe un problema de autocorrelación grave. Los ítems restantes mostraron un comportamiento similar. De forma complementaria al diagnóstico visual se obtuvieron los valores de Rhat para cada uno de los parámetros, resultando todos ellos muy cercanos a 1, lo cual es un indicio de buena precisión en los resultados.

Las conclusiones del estudio de simulación son:

1. Los parámetros estimados para los ítems y las personas del Modelo 2PL-GRM para una muestra y test de tamaño mediano ( $n=500$  y  $J=20$ ) con ítems de 5 categorías fueron recuperados de manera adecuada usando inferencia bayesiana bajo MCMC en el programa WinBUGS .
2. Para lograr una convergencia de las estimaciones fue necesario realizar un gran número de iteraciones (2 cadenas con 26500 iteraciones cada una, desechando las 6500 primeras y muestreando cada 20 iteraciones).
3. El proceso de estimación de los parámetros del modelo para una muestra y test de tamaño mediano de 5 categorías por ítem suele ser demorado (aproximadamente 1 hora con 21 minutos)

Medida de Precisión	Ítem	b1	b2	b3	b4	a
RSME	1	0.501	0.297	0.119	0.281	0.103
	2	0.228	0.131	0.108	0.139	0.128
	3	0.173	0.111	0.096	0.154	0.078
	4	0.155	0.114	0.100	0.098	0.092
	5	0.206	0.139	0.101	0.186	0.118
	6	0.300	0.223	0.133	0.246	0.134
	7	0.381	0.294	0.183	0.187	0.106
	8	0.319	0.182	0.094	0.206	0.136
	9	0.266	0.179	0.123	0.280	0.104
	10	0.090	0.107	0.098	0.162	0.163
	11	0.173	0.101	0.151	0.296	0.124
	12	0.297	0.185	0.138	0.341	0.096
	13	0.110	0.085	0.098	0.135	0.154
	14	0.123	0.079	0.076	0.186	0.091
	15	0.124	0.100	0.108	0.140	0.191
	16	0.157	0.100	0.102	0.159	0.182
	17	0.127	0.097	0.071	0.151	0.105
	18	0.454	0.339	0.186	0.238	0.098
	19	0.105	0.107	0.102	0.157	0.092
	20	0.077	0.075	0.088	0.144	0.199
MAE	1	0.420	0.228	0.090	0.209	0.087
	2	0.194	0.104	0.085	0.112	0.105
	3	0.147	0.095	0.084	0.129	0.059
	4	0.130	0.096	0.084	0.079	0.084
	5	0.165	0.112	0.084	0.155	0.103
	6	0.225	0.144	0.106	0.174	0.107
	7	0.335	0.250	0.160	0.137	0.086
	8	0.248	0.150	0.073	0.166	0.102
	9	0.218	0.144	0.088	0.238	0.091
	10	0.073	0.092	0.085	0.120	0.144
	11	0.145	0.085	0.134	0.214	0.100
	12	0.250	0.150	0.116	0.261	0.078
	13	0.086	0.073	0.090	0.115	0.126
	14	0.100	0.063	0.066	0.157	0.071
	15	0.104	0.082	0.090	0.120	0.164
	16	0.130	0.086	0.092	0.132	0.148
	17	0.104	0.080	0.060	0.129	0.090
	18	0.391	0.286	0.167	0.200	0.077
	19	0.088	0.086	0.081	0.131	0.079
	20	0.062	0.064	0.068	0.119	0.180

Cuadro 5.3: Raíz del error cuadrático medio (RSME) y Media absoluta del error (MAE) de las estimaciones de los parámetros de los ítems en la simulación del Modelo de Respuesta Graduada Logístico (2PL-GRM) considerando R= 20 réplicas



Parámetro	RMSE	MAE	r
b1	0.218	0.181	0.973
b2	0.152	0.123	0.984
b3	0.114	0.095	0.991
b4	0.194	0.155	0.937
a	0.125	0.104	0.967
U	0.280	0.227	0.962

Cuadro 5.4: Resultados del promedio de la Raíz del error cuadrático medio (RSME), la Media absoluta del error (MAE) y la correlación (r) de los parámetros estimados en la simulación del Modelo de Respuesta Graduada Logístico (2PL-GRM) considerando  $R = 20$  réplicas



## Capítulo 6

# Aplicación

La aplicación presentada en este trabajo se encuentra relacionada al ámbito de la calidad de servicios, y de manera más específica, a la medición de la satisfacción de clientes. El objetivo de la aplicación es el de estimar la habilidad latente y los parámetros de los ítems con los Modelos 2PL-GRM y LPE-GRM, en una muestra de 5354 clientes de una empresa de telecomunicaciones que se comunicaron con el Call Center de atención al cliente por algún motivo (consulta, reclamo, pedido, etc.) usando la inferencia bayesiana basada en los métodos de MCMC.

### 6.1. Introducción

El estudio y la medición de la calidad son algunos de los elementos más importantes para poder determinar el valor de las propuestas de productos y servicios. De este modo se hace necesario no solo el hecho de tener en cuenta las expectativas de los clientes sobre la calidad, sino también desarrollar una forma eficiente y precisa de poder cuantificarla. [Lewis y Booms \(1983\)](#) definen a la calidad de servicios como "la medida de que tan bien el nivel del servicio prestado coincide con las expectativas del cliente".

Como menciona [Allen \(2004\)](#) dentro de los esfuerzos por mejorar la calidad de servicios y alinearlos a las expectativas de los clientes, surgen dos componentes muy importantes: la medición de la satisfacción y de la lealtad por parte de los clientes. La satisfacción de un cliente, un término comúnmente usado en Marketing, puede ser definida como una medida de "la respuesta a la expectativa del cliente o el juicio acerca de que un rasgo del producto o servicio, o de que un producto o servicio en sí mismo, proporciona un nivel placentero de recompensa que se relaciona con el consumo" ([Oliver, 1980](#)). Por otro lado, la lealtad es un estado de actitud dirigida hacia una organización y los servicios o productos que esta ofrece, y que se traduce en comportamientos deseables, como la recompra o la permanencia del cliente ([Allen, 2004](#)). En el ámbito local, la medición de la satisfacción de los clientes es algo relativamente nuevo para muchas empresas que principalmente se orientan en objetivos financieros, mientras que aquellas que si le dan importancia suelen aplicar modelos estadísticos que no son los más apropiados para el tipo de escala de medición considerada.

Al ser la satisfacción una variable que no puede observarse ni medirse directamente (variable latente), se hace muy común la aplicación de cuestionarios muy similares a los tests utilizados en otras aplicaciones psicométricas más tradicionales presentadas en el campo de la educación y psicología. En este caso, los instrumentos utilizados suelen estar conformados por ítems con respuestas múltiples que por lo general son del tipo politómico ordinal debido

a que permiten producir mayor información sobre el nivel de actitud y percepción que tienen las personas evaluadas, en comparación al uso de escalas dicotómicas (Singh et al., 1990). A pesar de las ventajas mostradas por la TRI en psicometría para estudiar instrumentos con características similares (Lord y Novick, 1968; Pasquali, 2009) y ser una importante área de trabajo en Estadística como se puede ver en Rao y Sinharay (2007), esta no ha sido aplicada regularmente en la medición de la calidad de servicios, en especial aplicando modelos que consideren la naturaleza politómica y ordinal de las respuestas evaluadas, como sí lo hacen los modelos propuestos en esta investigación. Al contrario, el enfoque tradicional en los estudios de satisfacción de clientes es dar más importancia al análisis de los resultados obtenidos con una pregunta directa sobre la satisfacción del producto o servicio evaluado, la cual se asume como representativa y que no requiere ser estimada de los atributos como propone la TRI. Los ítems suelen ser utilizados para encontrar *drivers* o características más importantes del proceso que deben ser priorizadas y accionadas para la mejora de la calidad. Al respecto puede revisarse Busacca y Padula (2005), Mikulić y Prebežac (2008) y Dagostim y Tontini (2008).

Es de este modo que los Modelos de Respuesta Graduada (GRM) descritos a lo largo de este trabajo se muestran apropiados para este tipo de estudios, al poder establecer un modelo estadístico que describa en términos probabilísticos la respuesta que dan los sujetos evaluados y el nivel de la variable latente (habilidad) que se desea medir, que como se muestra en el presente capítulo, está relacionada a la satisfacción que presenta un cliente. Otra ventaja de la aplicación de los Modelos GRM en la medición de la satisfacción de clientes es que no sólo permite estimarla de un modo más coherente, sino que también es posible evaluar las características de los ítems que componen el cuestionario. Al respecto es importante resaltar la adaptación en la interpretación de los parámetros estimados en el ámbito de aplicación del estudio. En el ámbito de la calidad de servicios la habilidad se encuentra enfocada en constructos como en este caso lo es la satisfacción de los clientes sobre cierto producto o servicio que se le brinda, así como también lo puede ser la lealtad hacia una marca, la imagen de una empresa, etc. En relación a los parámetros de locación  $b$ , estos suelen ser asociados a la dificultad de un ítem. En otras palabras, para una categoría o etapa  $k$ , valores pequeños de  $b$  están asociados a ítems con poca dificultad o fáciles (los individuos tienen una alta probabilidad de responder en una categoría mayor a la  $k$ ) mientras que los mayores valores de  $b$  están asociados a ítems más difíciles (los individuos tienen menos probabilidad de exceder a la categoría  $k$ ). Como menciona De Battisti et al. (2008), en el ámbito de la calidad de servicios, la escala debe verse de una manera opuesta: los ítems con valores más pequeños de  $b$  identifican ítems de mayor calidad (porque la probabilidad de que el sujeto, por ejemplo, esté más satisfecho o sea leal a una marca es alta); mientras que los ítems con valores altos de este parámetro corresponden a ítems con calidad muy pobre (menor probabilidad de que, por ejemplo, un sujeto esté satisfecho o sea leal a una marca). Visto de otra forma, valores más bajos de  $b$  permitirán discriminar mejor entre individuos con menor satisfacción o lealtad, mientras que ítems con valores de  $b$  más altos permitirán discriminar mejor entre individuos con mayor satisfacción o lealtad.

## 6.2. Instrumento de Medición

Se utilizó un cuestionario orientado a evaluar la satisfacción que tienen los clientes de una empresa de telecomunicaciones con la atención recibida al llamar a un Call Center (proceso de atención telefónica).

La decisión de la elección de los atributos a incluir en este instrumento fueron derivadas de las conclusiones obtenidas de estudios cualitativos (focus groups y encuestas a profundidad) hechos con clientes que han pasado por la experiencia del proceso de atención telefónica. Para la elección de los clientes que participarían en esta fase cualitativa, se consideraron diferentes criterios como la edad, sexo, lugar de residencia, estrato socio-económico, tipo de producto adquirido, etc. Estas conclusiones ayudaron a definir los principales atributos críticos hacia la calidad (CTQ, por las siglas en inglés, *critical to quality*) del proceso en estudio. Los CTQs identificados fueron los siguientes:

- **Facilidad de comunicarse:** Evalúa la factibilidad de que el acceso de la comunicación con la central telefónica sea inmediata, entendiéndose como acceso tanto al hecho de que la llamada ingrese en un primer intento como al logro de comunicarse con el servicio de atención.
- **Utilidad de la selección del idioma:** Evalúa la posibilidad y variedad de opciones que tiene el cliente para determinar el idioma en que puede ser atendido.
- **Sencillez del menú:** Evalúa la utilidad y facilidad del uso del menú de opciones. Las opciones incluidas no sólo deben de ser las suficientes para abarcar los posibles motivos de llamada sino que deben ser fácilmente ubicables.
- **Tiempo en cola (espera):** Evalúa experiencia con el tiempo que transcurre desde que un cliente marca e ingresa la llamada hasta que un asesor contesta dicha llamada.
- **Amabilidad y cordialidad:** Evalúa que el asesor sea tolerante, aún con los clientes impacientes o molestos, y que no muestre exaltación.
- **Interés y compromiso:** Evalúa la empatía que muestra el asesor al momento de recibir la consulta o gestión del cliente, mostrando un interés en que se puedan solucionar las dificultades del cliente.
- **Capacidad y conocimiento del asesor:** Evalúa la identificación de un problema y las alternativas de solución que brinda un asesor ante las necesidades de un cliente.
- **Información adecuada que no genere confusiones:** Evalúa que la comunicación sea directa, clara y personalizada.
- **Tiempo empleado de atención:** Evalúa el tiempo que el asesor empleó para brindar la información y resolver la gestión del caso.
- **Tiempo de resolutivez:** Evalúa el tiempo global desde que el cliente llama hasta que se resolvió la gestión, este tiempo incluye el tiempo de contacto con el asesor hasta el término de la atención.

- **Solución brindada:** Evalúa la eficiencia del asesor para brindar una solución al cliente en un primer contacto y en el menor tiempo posible.

Para poder medir estos CTQs se elaboró un cuestionario compuesto por 11 ítems relacionados a estos aspectos de la satisfacción con el proceso de atención telefónica que debían ser evaluados según la opinión de los clientes. La redacción de los ítems empezaba con la frase “Dígame que tan satisfecho se encuentra con ...”. Entre los atributos considerados se encontraban, por ejemplo, “la facilidad de comunicarse con el Call Center” o “la solución obtenida al finalizar la gestión”. Las respuestas fueron obtenidas usando una escala ordinal de 10 puntos que iba desde completamente insatisfecho (1) hasta completamente satisfecho (10). Los ítems considerados en el estudio se encuentra en el Anexo C.

Adicionalmente a estos ítems, en el instrumento se midieron otras variables como la satisfacción total con el proceso (en la misma escala de los atributos y ubicada al final) y el FCR (First Contact Resolution) o solución en un primer contacto.

### 6.3. Diseño Muestral

Con el fin de comparar los estimadores que se obtengan con los modelos planteados en el presente trabajo, se analizaron las respuestas obtenidas por 5354 clientes que se comunicaron con el Call Center de atención al cliente por algún motivo (consulta, reclamo, pedido, etc.) durante el mes anterior al estudio. Se utilizó un muestreo estratificado por tipo de producto y segmento del cliente.

### 6.4. Recolección de Datos

La recolección de datos se realizó durante tres semanas por medio de entrevistas telefónicas efectuadas por teleoperadoras previamente capacitadas.

### 6.5. Análisis del instrumento según el enfoque de la Teoría Clásica de los Test (TCT)

En esta aplicación se está considerando el enfoque de la teoría de los scores mentales, donde se ve a la satisfacción como una variable latente, que al no poder ser observada no podrá ser medida directamente a través de una sola respuesta (p. ej. la pregunta global de satisfacción), por lo que es estimada a partir de las respuestas de los clientes hacia un conjunto de ítems relacionados con la experiencia en el proceso. Para poder cumplir con esta finalidad, según el enfoque de la Teoría Clásica de los Test (TCT), dicho instrumento debe de tener ciertas características básicas como el de ser válido y fiable.

#### 6.5.1. Validez del Instrumento

El instrumento es válido cuando la medida estimada se acerca a la verdadera medida que se pretende estudiar. Una forma de evaluar la validez del instrumento es mediante el análisis factorial, el cual permite comprobar el supuesto de unidimensionalidad de la escala de habilidad que se pretende evaluar, en este caso, de la satisfacción con el proceso de atención telefónica. La dimensionalidad de la escala fue evaluada usando el Análisis Factorial Exploratorio (AFE) considerando Ejes Principales, mínimos Cuadrados y mínimos cuadrados ponderados con rotación Varimax. La librería *Psych* del programa R fue utilizada para tal

fin. Los resultados del análisis factorial indicaron que la escala tiene dos dimensiones. El primer factor se encuentra compuesto por los ítems 1, 2, 3 y 4, el cual puede ser llamado como **Accesibilidad**, en el cual se obtuvo una suma de cuadrados para las cargas factoriales (SSL) igual a 3.88 con 35 % de varianza explicada. El segundo factor está compuesto por los ítems 5, 6, 7, 8, 9, 10 y 11, el cual puede ser llamado como **Desempeño del Asesor** con SSL=2.26 y 21 % de varianza explicada. En cuanto a lo índices de ajuste se obtuvo un Residuo Cuadrático Medio (RMSR) de 0.02, el índice de Tucker Lewis igual a 0.935 y un Error de Aproximación Cuadrático Medio (RMSEA) de 0.085. Los resultados obtenidos con este análisis se muestran en el Cuadro 6.1

Ítems	Factores rotados		Comunalidad
	1	2	
1	0.32	0.63	0.50
2	0.20	0.55	0.34
3	0.28	0.62	0.46
4	0.39	0.51	0.41
5	0.60	0.36	0.49
6	0.74	0.39	0.71
7	0.79	0.33	0.73
8	0.77	0.32	0.70
9	0.62	0.43	0.56
10	0.66	0.39	0.59
11	0.74	0.29	0.64
Porcentaje de varianza explicada	0.35	0.21	
Varianza acumulada	0.35	0.56	

Cuadro 6.1: Análisis de la Dimensionalidad de la escala. Matriz de cargas factoriales rotadas con el método VARIMAX y comunalidades

De acuerdo con los resultados obtenidos, se puede calcular las puntuaciones totales o “scores” para cada dimensión, de las cuales se muestran las gráficas de distribución en la Figura 6.1.

### 6.5.2. Fiabilidad del Instrumento

La fiabilidad del instrumento se refiere a la estabilidad de los resultados; es decir, que tan bien representan los puntajes observados al puntaje verdadero. La fiabilidad puede ser medida mediante la evaluación de la consistencia interna del instrumento usando el coeficiente Alfa de Cronbach (Cronbach, 1951) que es definido como:

$$\alpha = \frac{J}{J-1} \left( 1 - \frac{\sum_{j=1}^J S_{y.j}^2}{S_{y..}^2} \right)$$

donde:

- J : Número de ítems
- $S_{y.j}^2$  es la varianza del ítem j
- $S_{y..}^2$  es la varianza de la puntuación total

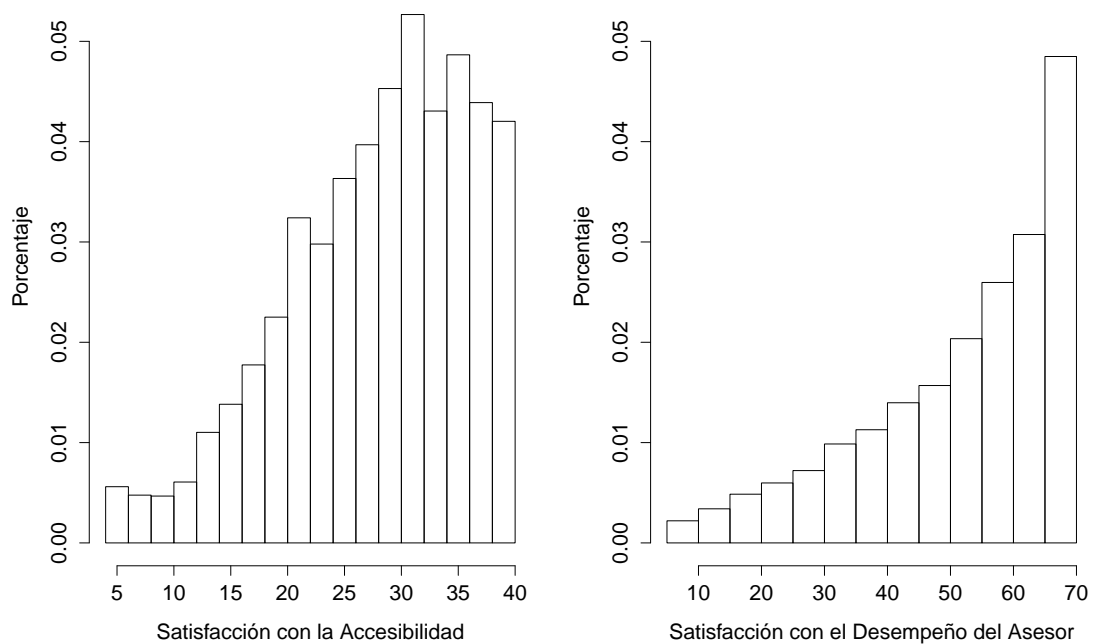


Figura 6.1: Distribución de las puntuaciones totales para cada dimensión

Se encontró un valor para el coeficiente Alfa de Cronbach igual a 0.7413 en la primera dimensión y de 0.9186 en la segunda, siendo ambos valores adecuados para garantizar la fiabilidad de la escala.

### 6.5.3. Análisis de los Ítems

Con el fin de determinar los ítems a incluir en cada una de las dos dimensiones se analizó el poder discriminativo (relevancia). Se dice que un ítem tiene poder discriminativo si distingue entre aquellos individuos que puntúan con alto nivel en la prueba y los que puntúan bajo. La discriminación fue medida con la correlación entre las puntuaciones de los individuos en el ítem y sus puntuaciones totales obtenidas con el instrumento, eliminando dicho ítem (CITC). Desde una perspectiva de la teoría clásica de los test (TCT), si el CITC es menor a 0.25 entonces el ítem no discrimina. Como se observa en el Cuadro 6.2 todos los ítems mostraron un CITC de al menos 0.47. Además del criterio anterior se evaluó el que no existiera redundancia entre los ítems, considerando el coeficiente Alfa de Cronbach excluyendo dicho ítem. En todos los casos se obtuvo un menor valor del coeficiente al ser excluido el ítem. En el mismo Cuadro 6.2 se muestran otros resultados descriptivos para los ítems (media y desviación estándar). De los resultados anteriores se puede concluir que no existe evidencia inicial para descartar algún ítem en las dimensiones halladas. Las características finales de la escala son mostradas en el Cuadro 6.3.

Dimensión	Ítem	Media	S	CITC	Alfa
Accesibilidad (Alfa = 0.7413)	1	6.628	2.904	0.602	0.641
	2	8.146	2.473	0.473	0.715
	3	7.430	2.539	0.559	0.670
	4	6.188	2.884	0.511	0.697
Desempeño del Asesor (Alfa = 0.9186)	5	8.582	2.149	0.667	0.915
	6	7.605	2.654	0.798	0.901
	7	7.439	2.745	0.810	0.900
	8	7.713	2.698	0.793	0.902
	9	7.280	2.634	0.709	0.910
	10	6.558	3.013	0.729	0.909
	11	7.081	3.194	0.763	0.906

Cuadro 6.2: Análisis de ítems bajo la perspectiva clásica

Medida	Dimensión	
	Accesibilidad	Desempeño del Asesor
Media	28.29	52.26
Mediana	30.00	56.00
Desviación Estándar	8.13	15.74
Asimetría	-0.65	-0.90
Curtosis	-0.12	-0.04

Cuadro 6.3: Estadísticos para las puntuaciones totales

### 6.6. Estimación de los parámetros mediante los métodos desarrollados

Para la implementación de los modelos se utilizó la librería R2WinBUGS del software R y el programa WinBUGS. Los códigos desarrollados en WinBUGS fueron basados en la propuesta de Curtis (2010), pero adaptados para las necesidades del estudio, en especial para el caso de los modelos con enlace asimétrico. Para comparar la estimación realizada por los diferentes modelos, se ejecutó la aplicación para 1500 iteraciones efectivas con tres cadenas (9 000 iteraciones por cadena, con un *burning* de las 4 000 primeras y un *thin* de 10) para cada modelo planteado y se utilizó la medida de ajuste DIC para la comparación de los modelos estimados (ver Sección 6.7).

#### 6.6.1. Modelo de Respuesta Graduada Logístico (2PL-GRM)

En el Modelo de Respuesta Graduada Logístico (2PL-GRM), la probabilidad de responder a una categoría  $k$  en el ítem  $j$  depende del nivel de satisfacción con el proceso de atención telefónica (habilidad  $u_i$ ) y de los parámetros de inclinación o discriminación  $a_j$  y posición o dificultad  $b_{jk}$ :

$$P_{ijk} \equiv P_{jk}(u_i) \equiv P(Y_{ij} = k | u_i, \xi_j) = \begin{cases} \Psi(\eta_{ijk}) & \text{si } k = 1 \\ \Psi(\eta_{ijk}) - \Psi(\eta_{ij(k-1)}) & \text{si } 2 < k < m \\ 1 - \Psi(\eta_{ij(k-1)}) & \text{si } k = m \end{cases}$$



$$-\infty = b_{j0} < b_{j1} < \dots < b_{j(m-1)} < b_{jm} = \infty$$

$$i = 1, \dots, n, j = 1, \dots, J, k = 1, \dots, m$$

donde  $\eta_{ijk} = a_j (b_{jk} - u_i)$  y  $\Psi(\cdot)$  la fda de la distribución logística estándar. El código para estimar el Modelo 2PL-GRM se encuentra en el Anexo A.3

Dimensión	Ítem	$a$	$ee_a$	2.50 %	Mediana	97.50 %
Accesibilidad	1	2.12	0.07	1.99	2.12	2.26
	2	1.42	0.05	1.32	1.42	1.52
	3	1.77	0.06	1.65	1.77	1.89
	4	1.59	0.05	1.50	1.59	1.70
Desempeño del Asesor	5	2.41	0.06	2.29	2.41	2.52
	6	3.27	0.08	3.10	3.27	3.42
	7	3.34	0.07	3.21	3.34	3.48
	8	3.31	0.08	3.16	3.31	3.46
	9	2.24	0.05	2.15	2.25	2.34
	10	2.30	0.05	2.20	2.30	2.39
	11	2.86	0.06	2.73	2.85	2.98

Cuadro 6.4: Estimación de los parámetros de discriminación del Modelo 2PL-GRM

Los valores estimados de los parámetros de discriminación ( $a$ ), el error estándar ( $ee_a$ ) y los principales estadísticos de orden se presentan en el Cuadro 6.4. Se observa que los ítems presentan una discriminación entre aceptable ( $a > 1$ ) y alta ( $a > 2$ ) según el criterio de [Hafsteinsson et al. \(2007\)](#), al reportar valores de  $a$  entre 1.42 a 3.34, siendo solamente tres de ellos menores a 2 (Ítems 2, 3 y 4). Para la dimensión relacionada con la Accesibilidad, la discriminación varía entre 1.42 a 2.12, siendo el Ítem 1, “Facilidad de comunicarse con la central”, el que tiene mayor valor y el Ítem 2, “Utilidad de poder seleccionar el idioma”, el de menor valor. En relación a la Gestión del Asesor, la discriminación varía entre 2.24 hasta 3.34, siendo el Ítem 7, “Capacidad o nivel de conocimiento del personal para dar una solución”, el de mayor valor y el Ítem 5, “Amabilidad y cordialidad con el asesor”, el de menor valor. Para la dimensión relacionada con la Gestión del Asesor, la discriminación varía entre 1.42 a 2.12, siendo el Ítem 1, “Facilidad de comunicarse con la central”, el que tiene mayor valor y el Ítem 2, “Utilidad de poder seleccionar el idioma”, el de menor valor. En relación al Desempeño del Asesor, la discriminación varía entre 2.24 hasta 3.34, siendo el Ítem 7, “Capacidad o nivel de conocimiento del personal para dar una solución”, el de mayor valor y el Ítem 5, “Amabilidad y cordialidad con el asesor”, el de menor valor.

Los valores estimados de los parámetros de dificultad de los ítems (umbrales) se muestran en el Cuadro 6.5. En todos los ítems la mayoría de los umbrales son negativos lo cual significa que están asociados a rasgos de buena calidad de servicio. Este hecho puede causar problemas en la estimación de las habilidades para aquellos individuos que se encuentran muy satisfechos. En la dimensión asociada a la Accesibilidad, el Ítem 2, “Utilidad de poder seleccionar el idioma”, es el que permite diferenciar mejor entre individuos insatisfechos (presenta mayor calidad para los individuos insatisfechos) mientras que el Ítem 4, “Tiempo de espera hasta que se contestó la llamada”, es el que permite diferenciar mejor entre individuos satisfechos



Dimensión	Ítem	Estadístico	b1	b2	b3	b4	b5	b6	b7	b8	b9	
Accesibilidad	1	Media	-1.77	-1.46	-1.19	-0.96	-0.45	-0.23	0.06	0.62	0.98	
		eeb	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.03
		2.5 %	-1.86	-1.53	-1.26	-1.01	-0.50	-0.27	0.02	0.57	0.93	
		Mediana	-1.77	-1.46	-1.19	-0.96	-0.45	-0.23	0.06	0.62	0.98	
		97.5 %	-1.69	-1.39	-1.13	-0.90	-0.41	-0.18	0.11	0.66	1.04	
	2	Media	-2.87	-2.48	-2.25	-2.06	-1.51	-1.26	-0.92	-0.25	0.18	
		eeb	0.09	0.07	0.07	0.06	0.05	0.04	0.03	0.03	0.03	
		2.5 %	-3.05	-2.64	-2.39	-2.19	-1.60	-1.34	-0.99	-0.30	0.13	
		Mediana	-2.87	-2.48	-2.25	-2.06	-1.50	-1.25	-0.92	-0.25	0.18	
		97.5 %	-2.71	-2.35	-2.13	-1.95	-1.42	-1.18	-0.85	-0.20	0.22	
	3	Media	-2.48	-2.15	-1.85	-1.60	-0.95	-0.68	-0.31	0.37	0.75	
		eeb	0.07	0.06	0.05	0.04	0.03	0.03	0.02	0.02	0.03	
		2.5 %	-2.61	-2.26	-1.94	-1.69	-1.01	-0.73	-0.36	0.33	0.70	
		Mediana	-2.47	-2.15	-1.85	-1.60	-0.95	-0.68	-0.31	0.37	0.75	
		97.5 %	-2.36	-2.05	-1.76	-1.52	-0.89	-0.63	-0.26	0.42	0.81	
	4	Media	-1.92	-1.51	-1.16	-0.89	-0.30	-0.04	0.33	0.98	1.47	
		eeb	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.03	0.04	
		2.5 %	-2.03	-1.59	-1.23	-0.95	-0.35	-0.08	0.29	0.92	1.39	
		Mediana	-1.92	-1.51	-1.16	-0.89	-0.30	-0.04	0.33	0.98	1.47	
		97.5 %	-1.83	-1.43	-1.10	-0.84	-0.25	0.00	0.38	1.05	1.55	
Desempeño del Agente	5	Media	-2.41	-2.21	-2.01	-1.86	-1.47	-1.25	-0.96	-0.46	-0.04	
		eeb	0.06	0.05	0.04	0.04	0.03	0.03	0.03	0.02	0.02	
		2.5 %	-2.52	-2.32	-2.10	-1.94	-1.53	-1.31	-1.02	-0.51	-0.08	
		Mediana	-2.41	-2.21	-2.01	-1.86	-1.47	-1.25	-0.96	-0.46	-0.04	
		97.5 %	-2.30	-2.11	-1.92	-1.78	-1.41	-1.20	-0.91	-0.42	0.00	
	6	Media	-1.85	-1.61	-1.39	-1.22	-0.85	-0.64	-0.37	0.10	0.48	
		eeb	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02	
		2.5 %	-1.92	-1.68	-1.45	-1.27	-0.90	-0.69	-0.40	0.07	0.44	
		Mediana	-1.85	-1.61	-1.39	-1.22	-0.85	-0.64	-0.37	0.10	0.48	
		97.5 %	-1.78	-1.55	-1.34	-1.17	-0.80	-0.60	-0.33	0.14	0.52	
	7	Media	-1.73	-1.50	-1.30	-1.13	-0.78	-0.59	-0.31	0.14	0.56	
		eeb	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.02	
		2.5 %	-1.80	-1.56	-1.36	-1.17	-0.82	-0.63	-0.35	0.10	0.52	
		Mediana	-1.73	-1.50	-1.30	-1.12	-0.78	-0.59	-0.31	0.14	0.56	
		97.5 %	-1.66	-1.45	-1.25	-1.08	-0.74	-0.55	-0.27	0.18	0.60	
	8	Media	-1.80	-1.57	-1.37	-1.19	-0.87	-0.69	-0.44	0.01	0.38	
		eeb	0.03	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02	
		2.5 %	-1.87	-1.63	-1.43	-1.25	-0.91	-0.74	-0.48	-0.02	0.34	
		Mediana	-1.80	-1.56	-1.37	-1.19	-0.87	-0.69	-0.44	0.01	0.38	
		97.5 %	-1.73	-1.51	-1.32	-1.14	-0.83	-0.65	-0.41	0.05	0.42	
	9	Media	-2.05	-1.76	-1.49	-1.27	-0.81	-0.58	-0.22	0.37	0.82	
		eeb	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.03	
		2.5 %	-2.14	-1.84	-1.55	-1.32	-0.86	-0.62	-0.26	0.33	0.77	
		Mediana	-2.05	-1.76	-1.49	-1.27	-0.81	-0.58	-0.22	0.37	0.82	
		97.5 %	-1.97	-1.69	-1.42	-1.21	-0.76	-0.53	-0.18	0.42	0.87	
	10	Media	-1.51	-1.26	-1.05	-0.87	-0.48	-0.26	0.04	0.59	1.01	
		eeb	0.03	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.03	
		2.5 %	-1.58	-1.32	-1.11	-0.92	-0.52	-0.31	-0.01	0.54	0.96	
		Mediana	-1.51	-1.26	-1.05	-0.87	-0.47	-0.26	0.04	0.59	1.01	
		97.5 %	-1.44	-1.20	-1.00	-0.82	-0.43	-0.22	0.08	0.63	1.06	
	11	Media	-1.31	-1.17	-1.03	-0.91	-0.63	-0.48	-0.26	0.13	0.48	
		eeb	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.02	
		2.5 %	-1.37	-1.22	-1.08	-0.95	-0.67	-0.52	-0.30	0.09	0.44	
		Mediana	-1.31	-1.17	-1.03	-0.91	-0.63	-0.48	-0.26	0.13	0.48	
		97.5 %	-1.26	-1.11	-0.98	-0.86	-0.59	-0.44	-0.22	0.17	0.52	

Cuadro 6.5: Estimación de los parámetros de dificultad del Modelo 2PL-GRM

(presenta menor calidad para los individuos satisfechos). En cuanto a la dimensión asociada al Desempeño del Asesor, el Ítem 5, “Amabilidad y cordialidad del asesor”, es el que presenta mayor calidad mientras que el Ítem 10, “Tiempo hasta que terminó la gestión”, es el que presenta menor calidad. En las Gráficas 6.2 y 6.3 se muestran las FREI y FRCI para estos ítems. El código en R para graficar dichas funciones se encuentran en los Anexos A.7 y A.8 de manera respectiva.

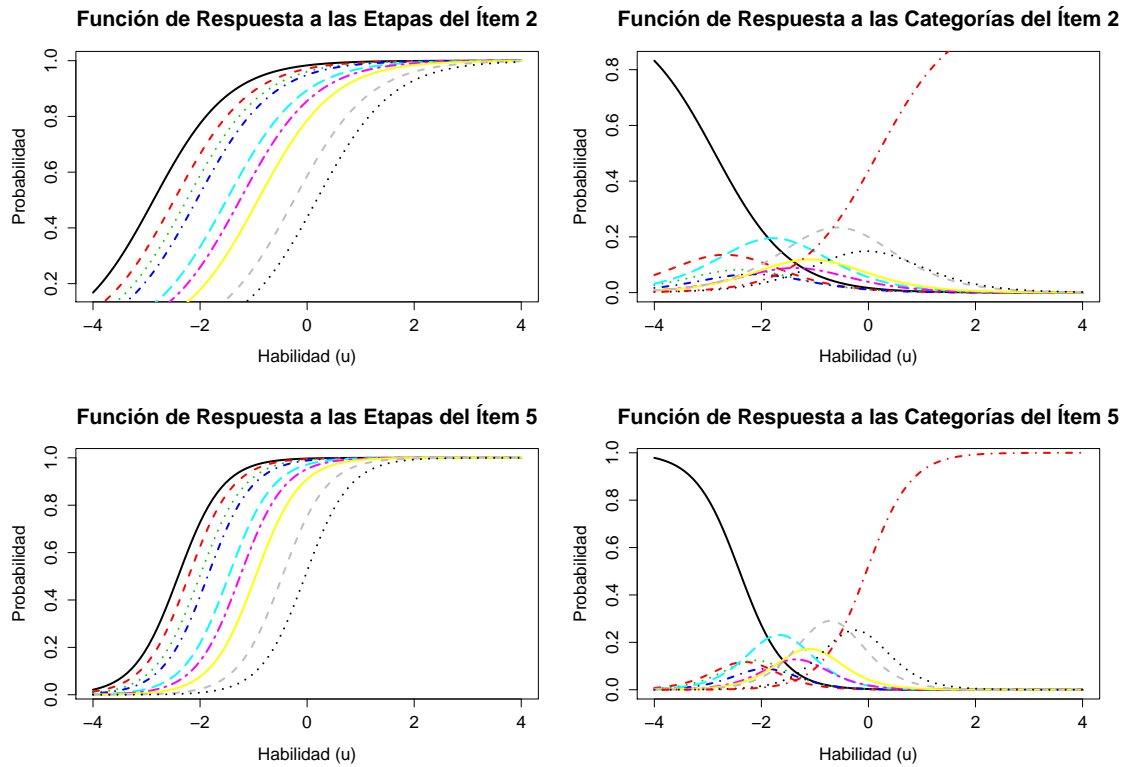


Figura 6.2: FREI y FRCI del Modelo 2PL-GRM para los atributos que permiten diferenciar mejor entre los individuos más insatisfechos

La convergencia fue examinada visualmente usando las librerías `mcmcplots` y `R2WinBUGS` del programa R como se muestra en el Anexo B.2. La Figura B.3 muestra el histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el parámetro de discriminación del Ítem 9, tiempo de atención. La Figura B.4 muestra un diagnóstico similar para el segundo umbral (parámetro de dificultad) del mismo ítem. De manera general se aprecia que las distribuciones para las cadenas convergen y que no existe un problema de autocorrelación grave. Los ítems restantes mostraron un comportamiento similar. De forma complementaria al diagnóstico visual se obtuvieron los valores de  $R_{hat}$  para cada uno de los parámetros, resultando todos ellos muy cercanos a 1, lo cual es un indicio de buena precisión en los resultados.

### 6.6.2. Modelo de Respuesta Graduada Logístico de Exponente Positivo (LPE-GRM)

En el Modelo de Respuesta Graduada Logístico (LPE-GRM), la probabilidad de responder a una categoría  $k$  en el ítem  $j$  depende del nivel de satisfacción con el proceso de atención

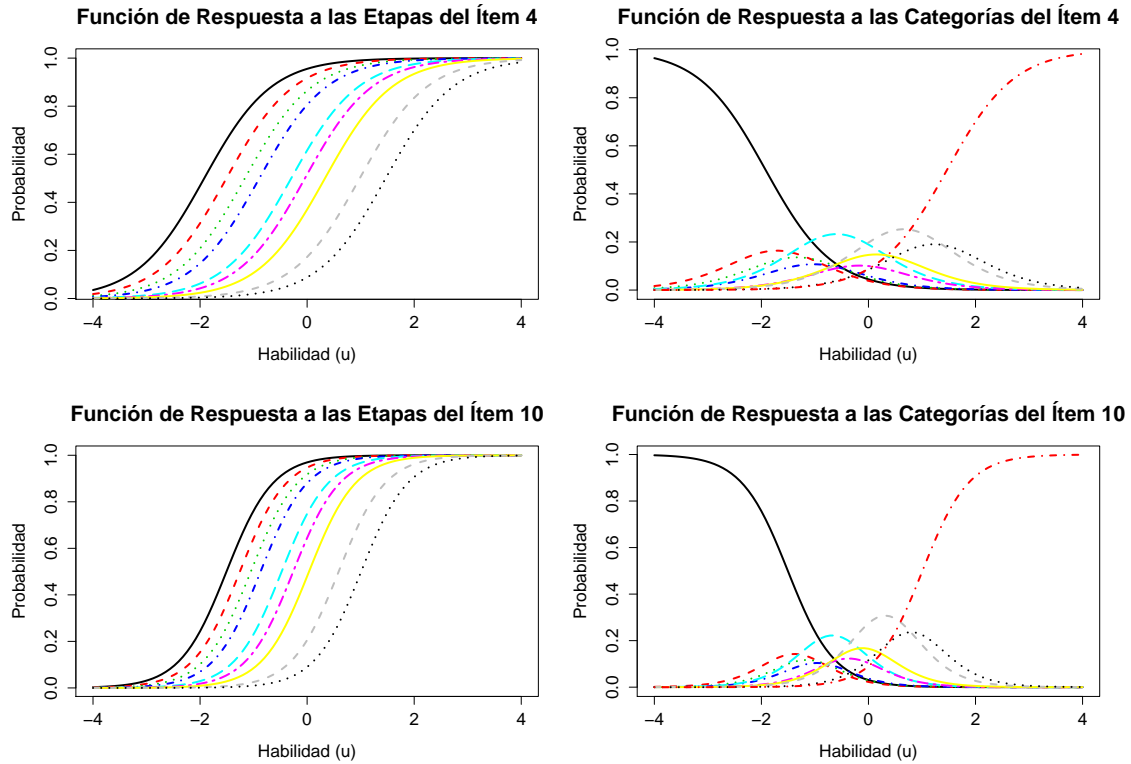


Figura 6.3: FREI y FRCI del Modelo 2PL-GRM para los atributos que permiten diferenciar mejor entre individuos más satisfechos

telefónica (habilidad  $u_i$ ) y de los parámetros de inclinación o discriminación  $a_j$ , posición o dificultad  $b_{jk}$  y el parámetro de penalidad  $d_j$  que controla la asimetría de las FREI.:

$$P_{ijk} \equiv P_{jk}(u_i) \equiv P(Y_{ij} = k | u_i, \xi_j) = \begin{cases} [\Psi(\eta_{ijk})]^{d_j} & \text{si } k = 1 \\ [\Psi(\eta_{ijk})]^{d_j} - [\Psi(\eta_{ij(k-1)})]^{d_j} & \text{si } 2 < k < m \\ 1 - [\Psi(\eta_{ij(k-1)})]^{d_j} & \text{si } k = m \end{cases}$$

$$-\infty = b_{j0} < b_{j1} < \dots < b_{j(m-1)} < b_{jm} = \infty$$

$$i = 1, \dots, n, j = 1, \dots, J, k = 1, \dots, m$$

donde  $\eta_{ijk} = a_j(b_{jk} - u_i)$  y  $\Psi(\cdot)$  la fda de la distribución logística estándar. El código para estimar el Modelo LPE-GRM se encuentra en el Anexo A.4

Los valores estimados de los parámetros de discriminación ( $a$ ), el error estándar ( $ee_a$ ) y los principales estadísticos de orden se presentan en el Cuadro 6.6. Se observa que los ítems presentan una discriminación entre 2.06 a 4.22, con mayores errores estándar que los obtenidos con las estimaciones del Modelo 2PL-GRM. Para la dimensión relacionada con la Accesibilidad, la discriminación varía entre 2.06 a 3.12, siendo el Ítem 1, “Facilidad de comunicarse con la central”, el que tiene mayor valor y el Ítem 4, “Tiempo de espera hasta que se contestó la llamada”, el de menor valor. En relación al Desempeño del Asesor, la

Dimensión	Ítem	$a$	$ee_a$	2.50 %	Mediana	97.50 %
Accesibilidad	1	3.12	0.17	2.81	3.11	3.51
	2	2.56	0.16	2.29	2.54	2.92
	3	2.63	0.16	2.33	2.63	2.92
	4	2.06	0.09	1.89	2.06	2.24
Desempeño del Asesor	5	3.73	0.16	3.40	3.73	4.03
	6	3.85	0.12	3.61	3.85	4.10
	7	3.54	0.12	3.32	3.54	3.79
	8	4.22	0.15	3.95	4.21	4.53
	9	2.57	0.08	2.41	2.57	2.73
	10	2.28	0.07	2.14	2.28	2.42
	11	2.91	0.09	2.73	2.90	3.08

Cuadro 6.6: Estimación de los parámetros de discriminación del Modelo LPE-GRM

discriminación varía entre 2.28 hasta 4.22, siendo el Ítem 8, “Claridad de la información brindada”, el de mayor valor y el Ítem 10, “Tiempo hasta que terminó la gestión”, el de menor valor.

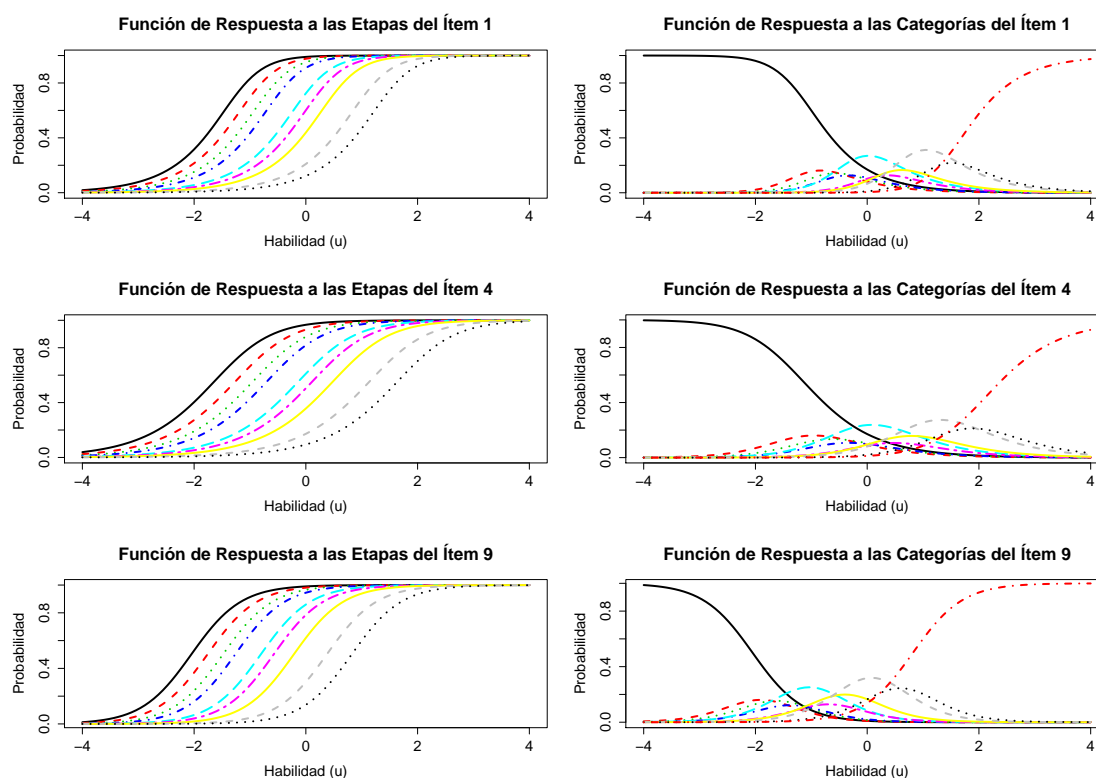


Figura 6.4: FREI y FRCI del Modelo LPE-GRM para los atributos que permiten diferenciar mejor entre individuos más insatisfechos y satisfechos

Los valores estimados de los parámetros de dificultad de los ítems (umbrales) se muestran en el Cuadro 6.7, observándose también que la mayor parte de los ítems tiene una mayor cantidad de umbrales negativos. En la dimensión asociada a la Accesibilidad, el Ítem 1, “Facilidad de comunicarse con la central”, es el que permite diferenciar mejor entre individuos

Dimensión	Ítem	Estadístico	b1	b2	b3	b4	b5	b6	b7	b8	b9
Accesibilidad	1	Media	-1.23	-0.94	-0.70	-0.48	0.00	0.22	0.51	1.08	1.46
		eeb	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.05	0.06
		2.5 %	-1.33	-1.04	-0.79	-0.57	-0.09	0.14	0.42	0.98	1.35
		Mediana	-1.23	-0.94	-0.70	-0.48	0.00	0.22	0.51	1.08	1.46
	97.5 %	-1.12	-0.84	-0.61	-0.38	0.09	0.31	0.60	1.18	1.58	
	2	Media	-1.42	-1.13	-0.94	-0.79	-0.31	-0.08	0.25	0.95	1.45
		eeb	0.08	0.07	0.07	0.06	0.06	0.05	0.05	0.06	0.08
		2.5 %	-1.57	-1.26	-1.07	-0.91	-0.41	-0.18	0.15	0.83	1.30
		Mediana	-1.42	-1.12	-0.94	-0.78	-0.30	-0.08	0.25	0.95	1.44
	97.5 %	-1.26	-0.99	-0.82	-0.67	-0.20	0.02	0.34	1.07	1.60	
	3	Media	-1.61	-1.34	-1.07	-0.85	-0.25	0.01	0.38	1.10	1.53
		eeb	0.08	0.07	0.06	0.06	0.05	0.05	0.05	0.06	0.07
		2.5 %	-1.78	-1.50	-1.21	-0.99	-0.37	-0.10	0.26	0.96	1.38
		Mediana	-1.61	-1.33	-1.07	-0.85	-0.25	0.02	0.39	1.10	1.54
	97.5 %	-1.48	-1.22	-0.97	-0.76	-0.17	0.09	0.47	1.20	1.65	
	4	Media	-1.39	-1.02	-0.70	-0.45	0.11	0.36	0.73	1.39	1.89
		eeb	0.07	0.07	0.06	0.06	0.06	0.06	0.06	0.07	0.08
		2.5 %	-1.55	-1.16	-0.83	-0.58	-0.01	0.23	0.60	1.25	1.72
		Mediana	-1.39	-1.01	-0.70	-0.45	0.12	0.37	0.74	1.39	1.89
	97.5 %	-1.27	-0.90	-0.59	-0.35	0.22	0.47	0.84	1.52	2.04	
Desempeño del Agente	5	Media	-1.67	-1.51	-1.34	-1.21	-0.85	-0.66	-0.38	0.13	0.59
		eeb	0.05	0.05	0.04	0.04	0.04	0.04	0.03	0.03	0.04
		2.5 %	-1.79	-1.61	-1.42	-1.29	-0.93	-0.73	-0.45	0.06	0.51
		Mediana	-1.67	-1.51	-1.33	-1.21	-0.85	-0.65	-0.37	0.13	0.59
	97.5 %	-1.58	-1.42	-1.25	-1.13	-0.79	-0.59	-0.32	0.19	0.66	
	6	Media	-1.61	-1.39	-1.17	-1.00	-0.64	-0.43	-0.15	0.33	0.72
		eeb	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
		2.5 %	-1.70	-1.46	-1.24	-1.06	-0.69	-0.48	-0.20	0.27	0.66
		Mediana	-1.61	-1.39	-1.17	-1.00	-0.64	-0.43	-0.15	0.33	0.72
	97.5 %	-1.53	-1.31	-1.10	-0.93	-0.58	-0.37	-0.09	0.39	0.79	
	7	Media	-1.65	-1.43	-1.23	-1.05	-0.70	-0.51	-0.23	0.22	0.63
		eeb	0.04	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.04
		2.5 %	-1.74	-1.51	-1.31	-1.13	-0.78	-0.58	-0.29	0.15	0.57
		Mediana	-1.65	-1.42	-1.23	-1.05	-0.70	-0.51	-0.23	0.22	0.64
	97.5 %	-1.57	-1.35	-1.15	-0.98	-0.64	-0.45	-0.16	0.29	0.71	
	8	Media	-1.49	-1.27	-1.08	-0.91	-0.59	-0.41	-0.16	0.31	0.70
		eeb	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
		2.5 %	-1.57	-1.34	-1.15	-0.97	-0.64	-0.47	-0.21	0.26	0.64
		Mediana	-1.49	-1.27	-1.08	-0.91	-0.59	-0.41	-0.16	0.31	0.69
	97.5 %	-1.41	-1.20	-1.02	-0.85	-0.53	-0.35	-0.11	0.37	0.76	
	9	Media	-1.77	-1.49	-1.22	-1.01	-0.56	-0.33	0.03	0.63	1.09
		eeb	0.06	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.05
		2.5 %	-1.88	-1.59	-1.32	-1.10	-0.64	-0.41	-0.05	0.54	1.00
		Mediana	-1.76	-1.49	-1.22	-1.01	-0.55	-0.33	0.03	0.63	1.09
	97.5 %	-1.66	-1.40	-1.13	-0.93	-0.49	-0.26	0.10	0.70	1.18	
	10	Media	-1.54	-1.29	-1.08	-0.90	-0.50	-0.29	0.01	0.55	0.96
		eeb	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.06
		2.5 %	-1.68	-1.42	-1.20	-1.01	-0.61	-0.40	-0.10	0.44	0.85
		Mediana	-1.54	-1.28	-1.08	-0.89	-0.50	-0.28	0.01	0.56	0.97
	97.5 %	-1.44	-1.19	-0.99	-0.81	-0.42	-0.20	0.09	0.65	1.07	
	11	Media	-1.29	-1.14	-1.00	-0.88	-0.60	-0.45	-0.23	0.17	0.51
		eeb	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.04
		2.5 %	-1.38	-1.23	-1.09	-0.96	-0.68	-0.53	-0.31	0.09	0.43
		Mediana	-1.29	-1.14	-1.00	-0.88	-0.60	-0.45	-0.23	0.16	0.51
	97.5 %	-1.20	-1.05	-0.92	-0.80	-0.52	-0.37	-0.15	0.24	0.59	

Cuadro 6.7: Estimación de los parámetros de dificultad del Modelo LPE-GRM

insatisfechos (presenta mayor calidad para los individuos insatisfechos) mientras que el Ítem 4, “Tiempo de espera hasta que se contestó la llamada”, es el que permite diferenciar mejor entre individuos satisfechos (presenta menor calidad para los individuos satisfechos). En cuanto a la dimensión asociada al Desempeño del Asesor, el Ítem 9, “Tiempo de atención”, es el que permite diferenciar mejor tanto a los individuos más satisfechos como a los más insatisfechos. En la Gráfica 6.4 se muestran las FREI y FRCI para estos ítems.

Dimensión	Ítem	$d$	$ee_d$	2.50 %	Mediana	97.50 %
Accesibilidad	1	0.46	0.04	0.39	0.46	0.54
	2	0.23	0.02	0.19	0.23	0.27
	3	0.37	0.03	0.32	0.37	0.45
	4	0.60	0.05	0.53	0.60	0.71
Desempeño del Asesor	5	0.32	0.02	0.28	0.32	0.36
	6	0.62	0.04	0.54	0.62	0.69
	7	0.88	0.07	0.75	0.88	1.01
	8	0.51	0.03	0.45	0.51	0.57
	9	0.68	0.04	0.61	0.68	0.77
	10	1.09	0.08	0.95	1.08	1.28
	11	0.98	0.08	0.85	0.97	1.14

Cuadro 6.8: Estimación de los parámetros de penalización del Modelo LPE-GRM

Los valores estimados de los parámetros de penalización ( $d$ ), el error estándar ( $ee_d$ ) y los principales estadísticos de orden se presentan en el Cuadro 6.8. La mayoría de los ítems presentan una penalidad  $d < 1$  con excepción del Ítem 10. Esto implica que individuos con niveles de satisfacción muy bajos tienen una mayor probabilidad de estar satisfecho con ese atributo, en comparación con el caso simétrico. Es importante resaltar que el ítem de mayor complejidad, el Ítem 10, mide el tiempo hasta que se terminó la gestión el cual suele estar asociado a una etapa complicada del proceso de atención al cliente porque varía mucho de acuerdo al motivo de la llamada.

En relación a la convergencia, la Figura B.5 muestra el histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el parámetro de discriminación del Ítem 9, tiempo de atención y las Figuras B.6 y B.7 muestran un diagnóstico similar para el segundo umbral (parámetro de dificultad) y el parámetro de penalización del mismo ítem. En algunos parámetros se observaron autocorrelaciones altas lo cual no implica que existan problemas de convergencia ya que los valores de Rhat resultaron en su mayoría cercanos a 1.

## 6.7. Comparación de Modelos

En el Cuadro 6.9 se muestra el resumen con los resultados de la comparación del ajuste de ambos modelos. Basados en el DIC, el modelo que estima mejor los parámetros para la dimensión asociada a la satisfacción con la accesibilidad es el Modelo 2PL-GRM y para el caso de la dimensión asociada con la satisfacción con el desempeño del asesor se obtuvo un mejor ajustes con el Modelo LPE-GRM. Considerando el resto de medidas (Dbar, EAIC y EBIC), el Modelo LPE-GRM presenta mejor ajuste en todos los casos. En la Figura 6.5 se muestra los diagramas de dispersión para las medias de las habilidades estimadas con el Modelo 2PL-GRM y el Modelo LPE-GRM. Se observa que en ambas dimensiones el primer

modelo sobrestima la satisfacción de los individuos que tienen niveles de satisfacción bajos.

Dimensión	Modelo	DIC	pD	Dbar	EAIC	EBIC	minutos
(1)	2PL	82125.30	8070.87	74054.43	84842.43	127842.83	180.04
	LPE	83040.43	9794.63	73245.80	84041.80	127074.08	330.53
(2)	2PL	114067.31	6771.78	107295.53	118143.53	164418.44	305.42
	LPE	113815.77	7053.37	106762.40	117624.40	163959.03	573.00

Cuadro 6.9: Medidas de comparación de los modelos para la satisfacción con la accesibilidad (Dimensión 1) y el desempeño del asesor (Dimensión 2)

### 6.8. Funciones de Información

En la Gráfica 6.6 se aprecian las Funciones de Información de los Ítems (FII) que fueron calculadas con la estimación promedio de los parámetros de los ítems considerando los modelos desarrollados. El código en R para realizar dichas gráficas se encuentra en el anexo A.9. Los resultados muestran que con el Modelo LPE-GRM la información para las habilidades es mayor en comparación al caso simétrico, por lo que los errores estándar son menores.

### 6.9. Características finales de la Escala

En la Figura 6.7 se presenta la distribución de las escalas de satisfacción estimadas con el Modelo LPE-GRM para la accesibilidad y el desempeño del asesor. En el Cuadro 6.10 se resumen los principales estadísticos descriptivos para las escalas estudiadas. Nótese que en ambos casos se observa una ligera asimetría negativa.

Medida	Dimensión	
	Accesibilidad	Desempeño del Asesor
Media	0.003	0.043
Mediana	0.005	0.049
Desviación Estándar	0.889	15.74
Asimetría	-0.017	-0.135
Curtosis	-0.297	-0.476

Cuadro 6.10: Estadísticos para las escalas finales de satisfacción con la accesibilidad y el desempeño del asesor obtenidas con el Modelo LPE-GRM

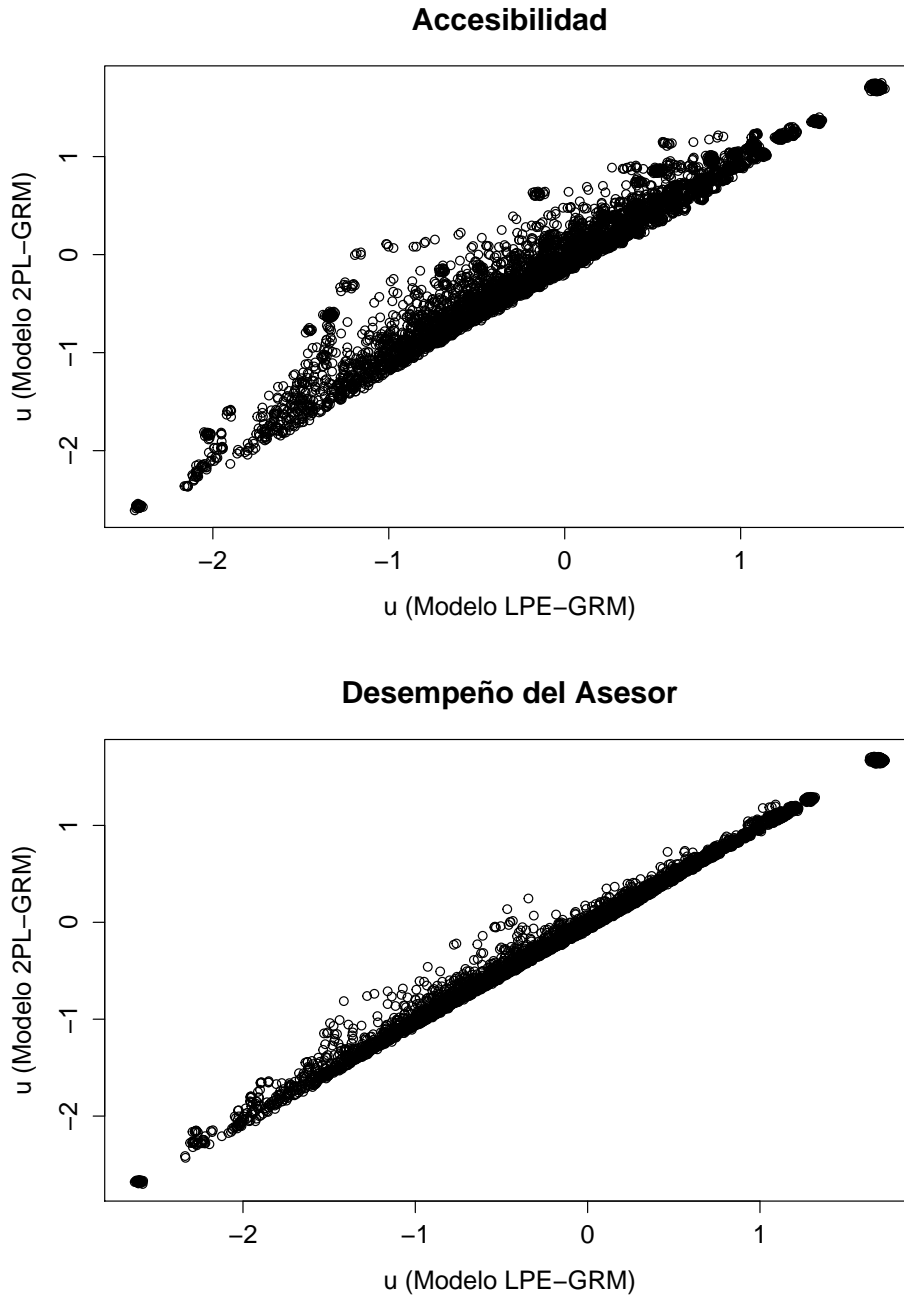


Figura 6.5: Diagramas de dispersión de las habilidades (satisfacción con la Accesibilidad y Desempeño del Asesor) estimadas con los modelos 2PL-GRM y LPE-GRM



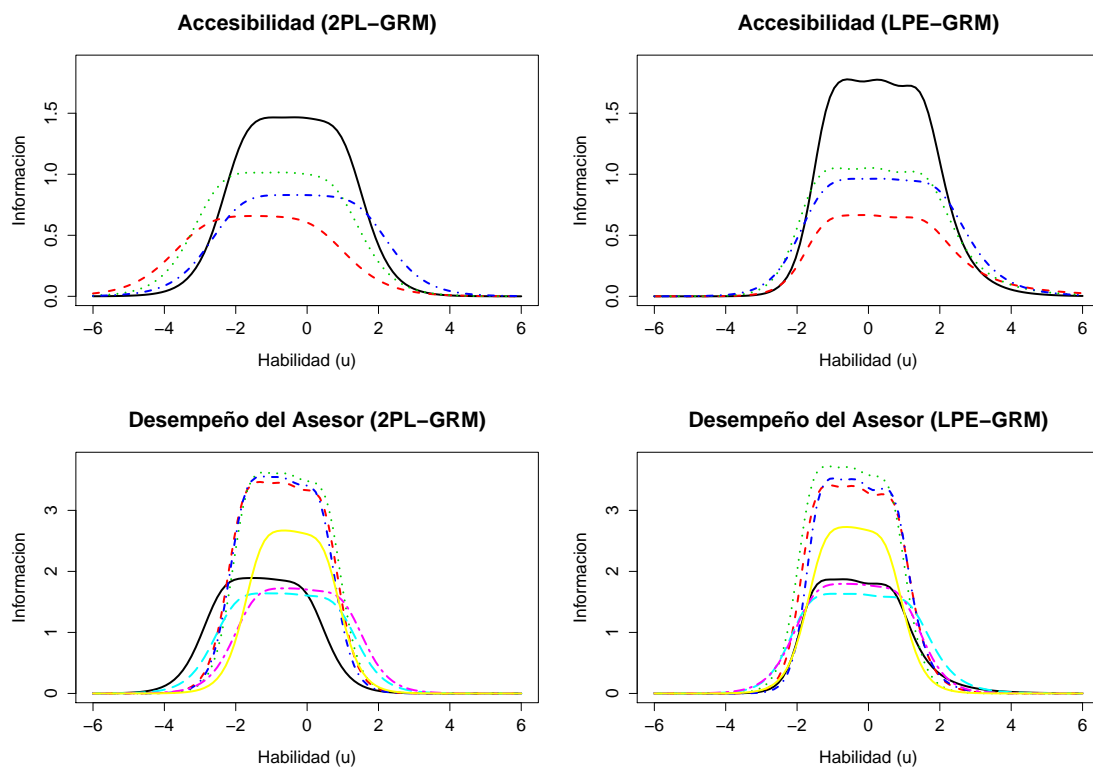


Figura 6.6: Funciones de Información de los Ítems (FII) obtenidas con los Modelos 2PL-GRM y LPE-GRM para la satisfacción con la accesibilidad y el desempeño del asesor

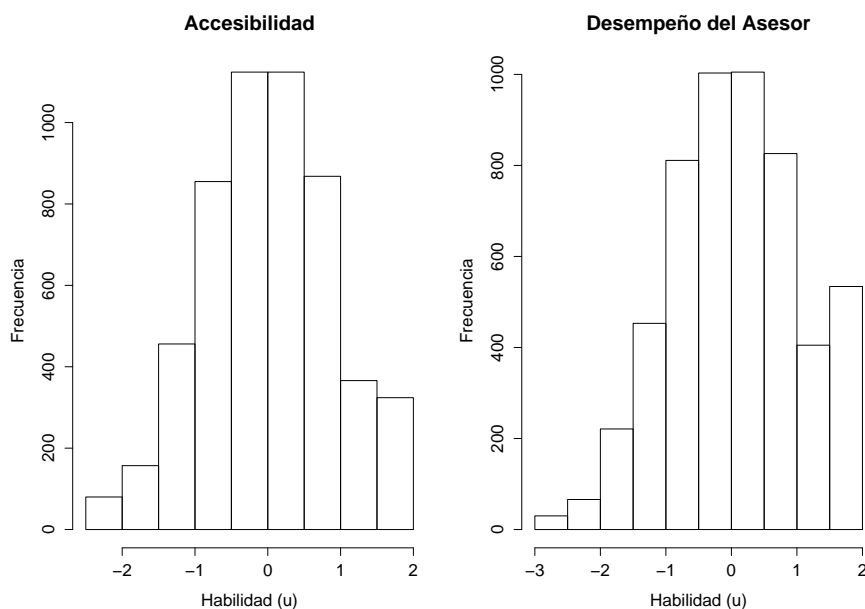


Figura 6.7: Distribución de los puntajes de las escalas de satisfacción con la accesibilidad y el desempeño del asesor usando el Modelo LPE-GRM

## Capítulo 7

# Conclusiones

### 7.1. Conclusiones

- Debido a que no existe una presentación adecuada del Modelo de Respuesta Graduada, en el Capítulo 1 fue presentado de manera general dicho modelo así como sus propiedades y características más importantes.
- Se describe el Modelo de Respuesta Graduada Logístico (2PL-GRM) que considera un enlace simétrico en la Función de Respuesta a las Etapas de los Ítems (FREI) y se presenta la estimación del modelo bajo la perspectiva bayesiana.
- Se propone, describe e implementa el Modelo de Respuesta Graduada de Exponente Positivo (LPE-GRM) como una extensión del Modelo 2PL-GRM considerando un enlace asimétrico.
- Se implementa la estimación de los parámetros de los modelos a través de la inferencia bayesiana con MCMC usando el software WinBUGS debido a la facilidad de este programa para implementar nuevos modelos, a pesar de que la estimación suele tomar mucho tiempo por la gran cantidad de parámetros que presenta el modelo.
- Se realizó un estudio de simulación con la finalidad de evaluar la precisión del método MCMC en la recuperación de los parámetros de los ítems y las habilidades del Modelo de respuesta Graduada Logístico (2PL-GRM) considerando una muestra y test de tamaño mediano con ítems de 5 categorías de respuesta, pudiéndose concluir que los parámetros fueron recuperados de manera adecuada al presentar medidas de precisión aceptables.
- La aplicación fue realizada con datos de un estudio para evaluar la satisfacción que tienen los clientes de una empresa de telecomunicaciones con la atención recibida al llamar a un Call Center (proceso de atención telefónica). De los resultados obtenidos se concluye que:
  - Los Modelos 2PL-GRM y LPE-GRM fueron aplicados de manera satisfactoria en el conjunto de datos del estudio.
  - A través del análisis de dimensionalidad de la escala se encontró que el cuestionario evalúa dos dimensiones de la satisfacción con la atención al cliente: la accesibilidad (4 ítems) y el desempeño del asesor (7 ítems)

- Ambos modelos estudiados permitieron un análisis a profundidad tanto de las propiedades de los ítems como de las escalas de habilidades (satisfacción con la accesibilidad y el desempeño del Asesor) que se pretendían estimar.
- Se obtuvo un mejor ajuste y mayor información en el cuestionario usando el Modelo LPE-GRM.
- La mayoría de los ítems mostró una asimetría  $d < 1$ , a excepción del Ítem 10, “tiempo hasta que se terminó la gestión”, el cual suele estar asociado a una etapa complicada del proceso de atención al cliente.

## 7.2. Sugerencias para investigaciones futuras

- Por motivos de comparación, en la aplicación se usó el mismo tamaño de cadena para el modelo 2PL-GRM y LPE-GRM por lo que se observa que este último presenta autocorrelaciones altas en el parámetro  $d_j$ . Por tanto se sugiere evaluar en estudios futuros el impacto de estas correlaciones considerando un número mayor de cadenas a ser simuladas.
- Realizar un estudio de simulación de recuperación de parámetros para el Modelo LPE-GRM.
- Realizar estudios de simulación para evaluar el efecto del tamaño de muestra, número de ítems y el número de categorías de respuesta en la estimación de los parámetros de los modelos estudiados.
- Implementar la estimación de los modelos estudiados usando el esquema de datos aumentados.
- Considerar otros tipos de enlace asimétrico como el de exponente positivo recíproco (RLPE) propuesto por [Bazán et al. \(2010\)](#).
- Extender el estudio del Modelo de Respuesta Graduada para el caso asimétrico considerando situaciones de multidimensionalidad .
- Optimizar el tiempo de estimación de los modelos considerando métodos de computación paralela incluidos en el software R con las librerías `bugsparell` y `dclone`.

## Apéndice A

### Códigos en R y BUGS

#### A.1. Código en R para simular datos aleatorios del GRM

```
#####
# Autor : Enver Gerald Tarazona Vargas #
# # #
# Función rgrm.pirt() #
# # #
# Fecha de inicio : 1-11-2012 #
# # #
# Función : Simular respuestas aleatorias basadas en el Modelo de Respuesta #
# Graduada Logístico (2PL-GRM) de Samejima(1969) y el Modelo de #
# Respuesta Graduada de Exponente Positivo (LPE-GRM). #
# # #
# Asesor : Jorge Luis Bazán #
# # #
# Versión : 1.1 #
# # #
# Última Modificación : 14-02-20 #
#####

rgrm.pirt<- function (n,
                    betas,
                    distr = c("normal", "logistic", "log-normal","uniform"),
                    u.vals = NULL,
                    model = c("2PL", "LPE")
                    )
{

# Argumentos:
# n = escalar indicando el número de patrones de respuesta a simular
# betas = lista de elementos con vectores numéricos, donde
```

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

# las primeras columnas corresponden a los umbrales y la última
# al parámetro de discriminación
# distr = distribución del vector de habilidades
# u.vals = vector de habilidades de las n personas
# model = modelo sobre el cual generar las respuestas aleatorias

#Simulación de las habilidades para las n personas en caso no sean
#ingresadas al llamar la función
distr <- match.arg(distr)
u<-if(is.null(u.vals) || length(u.vals) != n){
  switch(distr, normal = rnorm(n),
         logistic = sqrt(3)/pi*rlogis(n),
         'log-normal' = (rlnorm(n) - exp(0.5))/sqrt(exp(2) - exp(1)),
         uniform = runif(n, -3.5, 3.5)/sqrt(7^2/12)
        )
}
else {
  u.vals
}

# Número de ítems
p<-length(betas)

# Número de categorías por ítem
ncatg <- switch(model,
                '2PL' = sapply(betas, length),
                'LPE' = sapply(betas, length)-1
               )

# Cálculo de las probabilidades acumuladas
gammas<-pgrm.pirt(betas, u, model=model)

#Función que calcula las probabilidades individuales para cada categoría
#del ítem
prs <- lapply(gammas, function (x) {
  nc <- ncol(x)
  # cbind(x[, 1], x[, 2:nc]-x[, 1:(nc-1)])
  cbind(x[, 1], x[, 2:nc]-x[, 1:(nc-1)])
})
)

```

```

#Matriz que almacena la respuesta de los individuos
out <- matrix(0, n, p)
for (j in 1:p) {
  for (i in 1:n) {
    #selecciona al azar un elemento de entre la cantidad de categorias
    #del item, considerando las probabilidades individuales
    out[i, j] <- sample(ncatg[j], 1, prob = prs[[j]][i, ])
  }
}
out
}

#####
# Autor : Enver Gerald Tarazona Vargas #
# # #
# Función pgrm.pirt() #
# # #
# Fecha de inicio : 13-02-2013 #
# # #
# Función : Calcular para cada categoría de un ítem las probabilidades #
# acumuladas del Modelo de Respuesta Graduada. #
# # #
# Asesor : Jorge Luis Bazán #
# # #
# Versión : 1.0 #
# # #
# Última Modificación : 14-02-2013 #
#####

pgrm.pirt<- function(betas, u, model=c("2PL", "LPE")){
  # Función que devuelve, para cada ítem, una matriz con las
  # probabilidades acumuladas del Modelo de Respuesta Graduada

  n<-length(u)
  if(model=="2PL"){
    ap <- lapply(betas, function (x) {
      # número de parámetros de discriminación y dificultad
      nx <- length(x)
      #matriz con las probabilidades acumuladas
      cbind(plogis(x[nx] * (matrix(x[-nx], n, nx -1, TRUE) - u)),1)
    })
  }
}

```

```

    )
  }

  if(model=="LPE"){
    ap <- lapply(betas, function (x) {
      # número de parámetros de discriminación y dificultad
      nx <- length(x)-1
      #matriz con las probabilidades acumuladas
      cbind(plogis(x[nx] * (matrix(x[c(-nx,-(nx+1))], n, nx -1, TRUE)
        - u))^x[nx+1],1)
    })
  }

  return(ap)
}

```

## A.2. Función en R para realizar la estimación bayesiana usando MCMC

```

#####
# Autor : Enver Gerald Tarazona Vargas                                     #
#                                                                                   #
# Función grmMCMC()                                                         #
#                                                                                   #
# Fecha de inicio : 1-11-2012                                               #
#                                                                                   #
# Función : Función en R para realizar la estimación bayesiana usando MCMC #
#           en WinBUGS.                                                       #
#                                                                                   #
# Asesor : Jorge Luis Bazán                                                 #
#                                                                                   #
# Versión : 1.1                                                               #
#                                                                                   #
# Última Modificación : 14-02-2013                                           #
#####

grmMCMC <- function (data,K = NULL, model=c("2PL","LPE"),
                    program = c("WinBUGS","JAGS","OpenBUGS"),
                    bugs.dir="c:/Program Files/WinBUGS14/"
)

```

```

{

library(R2WinBUGS)

#Transformación de la data en un objeto matriz
Y<- as.matrix(data)
#Número de individuos
n <- nrow(Y)
#Número de ítems
p <- ncol(Y)
#Número de categorías por ítem
K <- if (is.null(K) || length(K) != p) {
  apply(Y, 2, max)
}
else {
  K
}
model <- match.arg(model)

#Definición de Prioris

#Discriminación
m.alpha <- 1.0
#s.alpha <- sqrt(0.5)
pr.alpha<-2

#Dificultad

m.kappa <- 0.0
#s.kappa <- sqrt(2)
pr.kappa <- 0.5

if (model == "LPE") {
  #Penalización
  m.g <- 0.0
  #s.g <- sqrt(0.5)
  pr.g <- 2
  data<-list("Y", "n", "p", "K",
            "m.alpha", "pr.alpha",
            "m.kappa", "pr.kappa", "m.g", "pr.g")
  #Monitor en BUGS

```



## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

monitor <- c("alpha", "theta", "kappa", "kappa.star", "d", "beta")
#Modelo usado en BUGS
bugs.file <- file.path(getwd(), "bugs/grm_lpe.bug")
}
else {

data<- list("Y", "n", "p", "K",
            "m.alpha", "pr.alpha",
            "m.kappa", "pr.kappa")
#Monitor en BUGS
monitor <- c("alpha", "theta", "kappa", "kappa.star", "beta")
#Modelo usado en BUGS
bugs.file <- file.path(getwd(), "bugs/grm_l.bug")
}

## Inferencia en WinBUGS

n.burn <- 40000
n.thin <- 20
n.sim <- 2000*n.thin + n.burn
n.chains <- 2

grm.out <- bugs(data=data,
               inits=NULL,
               parameters.to.save=monitor,
               model.file=bugs.file,
               bugs.directory=bugs.dir,
               n.iter=n.sim,
               n.thin=n.thin,
               n.burnin=n.burn,
               n.chains=n.chains)

}

```

### A.3. Código en BUGS para el Modelo de Respuesta Graduada Logístico (2PL-GRM)

```

#####
# Autor : Enver Gerald Tarazona Vargas                                     #
#                                                                                   #
# Programa grm_l.bug                                                       #
#####

```

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

# #
# Fecha de inicio : 1-11-2012 #
# #
# Función : Función en BUGS para realizar la inferencia del Modelo 2PL-GRM #
# #
# Asesor : Jorge Luis Bazán #
# #
# Versión : 1.0 #
# #
# Última Modificación : 28-11-2012 #
#####

model{
  for (i in 1:n){
    for (j in 1:p){
      Y[i, j] ~ dcat(prob[i, j, 1:K[j]])
    }
    theta[i] ~ dnorm(0.0, 1.0)

    ## Probabilidades acumuladas para las categorías de los ítems
    for (j in 1:p){
      for (k in 1:(K[j]-1)){
        logit(P[i, j, k]) <- kappa[j, k] - alpha[j]*theta[i]
      }
      P[i, j, K[j]] <- 1.0
    }

    ## Probabilidades para las categorías de los ítems
    for (j in 1:p){
      prob[i, j, 1] <- P[i, j, 1]
      for (k in 2:K[j]){
        prob[i, j, k] <- P[i, j, k] - P[i, j, k-1]
      }
    }
  }

  ## Prioris para los parámetros de los ítems
  for (j in 1:p){
    alpha[j] ~ dnorm(m.alpha, pr.alpha) T(0, )
  }
  pr.alpha <- pow(s.alpha, -2)
}

```

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

## Prioris para los umbrales
for (j in 1:p){
  for (k in 1:(K[j]-1)){
    kappa.star[j, k] ~ dnorm(m.kappa, pr.kappa)
  }
  kappa[j, 1:(K[j]-1)] <- sort(kappa.star[j, 1:(K[j]-1)])
}
pr.kappa <- pow(s.kappa, -2)
}

```

## A.4. Código en BUGS para el Modelo de Respuesta Graduada con Exponente Positivo (LPE-GRM)

```

#####
# Autor : Enver Gerald Tarazona Vargas #
# # #
# Programa grm_lpe.bug #
# # #
# Fecha de inicio : 1-11-2012 #
# # #
# Función : Función en BUGS para realizar la inferencia del Modelo LPE-GRM #
# # #
# Asesor : Jorge Luis Bazán #
# # #
# Versión : 1.0 #
# # #
# Última Modificación : 28-11-2012 #
#####

model{
  for (i in 1:n){
    for (j in 1:p){
      Y[i, j] ~ dcat(prob[i, j, 1:K[j]])
    }
    theta[i] ~ dnorm(0.0, 1.0)

    ## Probabilidades acumuladas para las categorías de los ítems
    for (j in 1:p){
      for (k in 1:(K[j]-1)){
        nu[i, j, k] <- kappa[j, k] - alpha[j]*theta[i]
        #logit(P[i, j, k]) <- kappa[j, k] - alpha[j]*theta[i]
        P[i, j, k] <- 1 - pow(1 + exp(nu[i, j, k]), -d[j])
      }
    }
  }
}

```

```

    P[i, j, K[j]] <- 1.0
  }

  ## Probabilidades para las categorías de los ítems
  for (j in 1:p){
    prob[i, j, 1] <- P[i, j, 1]
    for (k in 2:K[j]){
      prob[i, j, k] <- P[i, j, k] - P[i, j, k-1]
    }
  }
}

## Prioris para los parámetros de los ítems
for (j in 1:p){
  alpha[j] ~ dnorm(m.alpha, pr.alpha) I(0, )
g[j]~dnorm(m.g,pr.g)
d[j] <- exp(g[j])
}

## Prioris para los umbrales
for (j in 1:p){
  for (k in 1:(K[j]-1)){
    kappa.star[j, k] ~ dnorm(m.kappa, pr.kappa)
    kappa[j, k] <- ranked(kappa.star[j, 1:(K[j]-1)], k)
    beta[j,k]<-kappa[j,k]/alpha[j]
  }
}
}

```

#### A.5. Código en R para el estudio de simulación

```

#####
# Autor : Enver Gerald Tarazona Vargas #
# # #
# Programa simulacion_1.R #
# # #
# Fecha de inicio : 01-01-2013 #
# # #
# Función : Programa en R para realizar el estudio de simulación #
# # #
# Asesor : Jorge Luis Bazán #

```

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

# #
# Versión : 1.1 #
# #
# Última Modificación : 30-01-2013 #
#####

library(mcmcplots)
library(xtable)

#Carga los scripts de funciones auxiliares
source("grm_mcmc.R") #Inferencia Bayesiana del Modelo GRM usando MCMC
source("grm_sim.R") #Simulación de n respuestas ordinales para el Modelo GRM

#Directorio de WinBUGS
#winbugs.dir<-"D:/Archivos de programa/WinBUGS14/"
winbugs.dir<-"C:/Program Files (x86)/Winbugs14/"

#Parámetros de los ítems
item1 <-c(-4.26, -2.90, -1.25, 2.01, 0.95)
item2 <-c(-2.45, -1.44, -0.60, 1.45, 1.48)
item3 <-c(-2.07, -1.27, 0.16, 2.11, 1.46)
item4 <-c(-1.75, -0.76, 0.13, 2.02, 1.49)
item5 <-c(-2.19, -1.27, -0.35, 1.52, 1.38)
item6 <-c(-2.88, -1.97, -0.51, 1.87, 1.35)
item7 <-c(-3.77, -2.23, -1.27, 1.34, 0.96)
item8 <-c(-3.24, -2.29, -0.49, 1.93, 1.32)
item9 <-c(-3.28, -2.09, 0.49, 3.09, 1.08)
item10 <-c(-1.57, -0.75, -0.13, 1.68, 2.00)
item11 <-c(-1.39, 0.08, 1.07, 2.99, 1.22)
item12 <-c(-2.97, -1.50, -0.41, 2.44, 0.89)
item13 <-c(-2.05, -1.19, -0.15, 1.87, 2.05)
item14 <-c(-1.20, -0.24, 0.61, 2.48, 1.59)
item15 <-c(-1.68, -0.95, -0.25, 1.69, 2.31)
item16 <-c(-1.90, -1.08, -0.39, 1.59, 2.07)
item17 <-c(-1.80, -0.80, 0.10, 1.96, 1.55)
item18 <-c(-3.82, -2.63, -1.20, 1.67, 0.92)
item19 <-c(-1.40, -0.50, 0.35, 2.10, 1.64)
item20 <-c(-1.70, -0.90, -0.06, 1.81, 2.35)

items.par <- list(item1,item2,item3,item4,item5,item6,item7,item8, item9, item10,
                 item11, item12, item13, item14, item15,item16, item17,item18,
                 item19,item20)

```

```
#Número de categorías por ítem
K<-sapply(items.par, length)

#Número de individuos
n<-500

#Número de replicas
R<- 20

#Número de ítems
J<- length(items.par)

#Lista que almacena los resultados de las simulaciones
results.2plgrm<-list()

#Simulación
set.seed(666)

#Simulación de las habilidades asumiendo distribución normal estandar
u = rnorm(n)
for(s in 1:R){

  #Simulación de las respuestas de n individuos usando el Modelo 2PL-GRM
  sim<-rgrm.pirt(n=n,betas=items.par,u.vals=u)

  #Estimación de los parámetros usando MCMC
  Y<- as.matrix(sim)
  grm.bugs.time<-system.time(grm.bugs <- grmMCMC(data=Y,bugs.dir=winbugs.dir))

  #Almacenamiento de los resultados de la simulación
  results.2plgrm[[s]]<-list(model=grm.bugs, time=grm.bugs.time)
}

#Evaluación de la precisión de la simulación

#Matriz con los parámetros de los ítems
items.par<-do.call(rbind, items.par)

#Número de parámetros de los ítems
m<-dim(items.par)[2]
```

```

#Crea una matriz con los valores ajustados de los parámetros de los ítems
items.fit <- lapply(results.2plgrm,
  function (x) {
    a.fit<-x$model$mean$alpha
    b.fit<-x$model$mean$beta
    cbind(b.fit,a.fit)
  }
)
items.fit.means<-Reduce("+",items.fit)/R
items.fit.means

#Cálculo del RMSE

items.fit.rmse<-lapply(items.fit,
  function (x){
    (x-items.par)^2
  }
)
items.fit.rmse<-sqrt(Reduce("+",items.fit.rmse)/R)
items.fit.rmse

u.fit.rmse<-lapply(results.2plgrm,
  function (x){
    (x$model$mean$theta-u)^2
  }
)
u.fit.rmse<-sqrt(Reduce("+",u.fit.rmse)/R)
u.fit.rmse

#Cálculo del MAE
items.fit.mae<-lapply(items.fit,
  function (x){
    abs(x-items.par)
  }
)
items.fit.mae<-Reduce("+",items.fit.mae)/R
items.fit.mae

u.fit.mae<-lapply(results.2plgrm,
  function (x){
    abs(x$model$mean$theta-u)
  }
)

```

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

    }
  )
  u.fit.mae<-Reduce("+",u.fit.mae)/R
  u.fit.mae

#Cálculo de las Correlaciones

items.fit.cor<-matrix(0,R,m)
u.fit.cor<-matrix(0,R)

for(s in 1:R){
  for(k in 1:m){
    items.fit.cor[s,k]<-cor(items.fit[[s]][,k],items.par[,k])
  }
  u.fit.cor[s]<-cor(results.2plgrm[[s]]$model$mean$theta,u)
}

#Visualizar Resultados

#Tabla de resumen de las medidas de evaluación de la precisión de la simulación

items.fit.resume<-data.frame(
  c("b1","b2","b3","b4","a","U"),
  rbind(
    cbind(
      colMeans(items.fit.rmse),
      colMeans(items.fit.mae)
    ),
    c(mean(u.fit.rmse),mean(u.fit.mae))
  ),
  c(colMeans(items.fit.cor),mean(u.fit.cor))
)
dimnames(items.fit.resume)[[2]]<-c("Parámetro","RMSE","MAE", "r")
items.fit.resume<-xtable(items.fit.resume,
caption=c("Resultados del promedio de la Raiz del error cuadrático medio
(RSME), la Media absoluta del error (MAE) y la correlación (r) de
los parámetros estimados en la simulación del Modelo de Respuesta
Graduada Logístico (2PL-GRM) considerando R= 20 réplicas"),
digits=3,
label="sim1-fitresume")

```



## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

print(items.fit.resume, include.rownames=F)

#Tabla con RMSE, MAE y correlación promedio
#de las estimaciones de los parámetros de los ítems

items.fit.table<-data.frame(grp=rep(c("RSME","MAE"),each=J),
                           item=rep(1:J,2),
                           rbind(items.fit.rmse,items.fit.mae))
dimnames(items.fit.table)[[2]]<-c("Medida de Precisión","Ítem",
  "b1","b2","b3","b4","a")
items.fit.table<-latex.table.by(items.fit.table,
  caption=c("Raiz del error cuadratico medio (RSME) y Media absoluta
  del error (MAE)de las estimaciones de los parámetros de los
  ítems en la simulación delModelo de Respuesta Graduada Logístico
  (2PL-GRM) considerando R= 20 réplicas"),
  digits=3,
  label="sim1-fitititems")
print(items.fit.table, include.rownames=F,sanitize.text.function = force)

```

## A.6. Código en R para la aplicación en el análisis de satisfacción de clientes

```

#####
# Autor : Enver Gerald Tarazona Vargas #
# # #
# Programa aplicacion1.R #
# # #
# Fecha de inicio : 01-12-2012 #
# # #
# Función : Programa en R para realizar la aplicación de los Modelos #
# 2PL-GRM y LPE-GRM sobre un conjunto de datos de satisfaccion #
# de clientes. #
# # #
# Asesor : Jorge Luis Bazán #
# # #
# Versión : 1.1 #
# # #
# Última Modificación : 21-02-2013 #
#####

library(ltm)
library(Hmisc)
library(psych)

```

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

library(R2WinBUGS)
library(mcmcplots)

source("grm_mcmc.R")
source("grm_func.R")

#Directorio de WinBUGS
winbugs.dir<-"C:/Program Files (x86)/winbugs14/"

#Lectura de datos
baseatc<-read.table("data/Base Call Center.dat",sep="\t",header=TRUE)
baseatc<-as.data.frame(baseatc)

#Eliminación de datos perdidos
atc<-na.omit(atc)
summary(atc)

#####
# Factor Analysis      #
#####
atc.cor<-cor(atc[,c(-12,-13)], method="pearson", use = "complete.obs")
atc.cor
n<-5354
k<-2

#Principal Axis
pa <- fa(atc.cor,k,fm="pa",rotate="varimax",n.obs=n)
pa

#unweighted least squares is minres
uls <- fa(atc.cor,k,rotate="varimax",n.obs=n)
uls

#weighted least squares is minres
wls <- fa(atc.cor,k,fm="wls",rotate="varimax",n.obs=n)
wls

#####
# Análisis Clásico de los Ítems #
#####

#Dimension 1#

```

```

atc.d1<-atc[,c(1:4)]
descript(atc.d1)
describe(atc.d1)
alpha(atc.d1)

#Dimension 2#
atc.d2<-atc[,c(5:11)]
descript(atc.d2)
describe(atc.d2)
alpha(atc.d2)

#####
# Estimación del Modelo 2PL-GRM #
#####

#####
#Dimension 1#
#####

#Usar si no se hizo el análisis descriptivo
#atc.d1<-atc[,c(1:4)]

atc.d1.2l.time<-system.time(atc.d1.2l
  <- grmMCMC(data=atc.d1,bugs.dir=winbugs.dir))

#Diagnóstico
plot(atc.d1.2l)
mcmcplot(atc.d1.2l, random=20)

#####
#Dimension 2#
#####

#Usar si no se hizo el análisis descriptivo
#atc.d2<-atc[,c(5:11)]

atc.d2.2l.time<-system.time(atc.d2.2l
  <- grmMCMC(data=atc.d2,bugs.dir=winbugs.dir))

#Diagnóstico
plot(atc.d2.2l)
mcmcplot(atc.d2.2l, random=20)

```

```
#####
# Estimación del Modelo 2PL-GRM #
#####

#####
#Dimension 1#
#####

#Usar si no se hizo el análisis descriptivo
#atc.d1<-atc[,c(1:4)]

atc.d1.2lpe.time<-system.time(atc.d1.2lpe
  <- grmMCMC(data=atc.d1,bugs.dir=winbugs.dir,model="LPE"))

#Diagnóstico
plot(atc.d1.2lpe)
mcmcplot(atc.d1.2lpe)

#####
#Dimension 2#
#####

#Usar si no se hizo el análisis descriptivo
#atc.d2<-atc[,c(5:11)]

atc.d2.2lpe.time<-system.time(atc.d2.2lpe
  <- grmMCMC(data=atc.d2,bugs.dir=winbugs.dir,model="LPE"))

#Diagnóstico
plot(atc.d2.2lpe)
mcmcplot(atc.d2.2lpe, random=20)
```

#### A.7. Código en R para graficar la Función de Respuesta a las Etapas de los Ítems (FREI) del GRM

```
#####
# Autor : Enver Gerald Tarazona Vargas #
# #
# Función frei.grm.pirt() #
# #
```

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

# Fecha de inicio : 13-02-2013 #
# #
# Función : Gráficar la Función de Respuesta a las Etapas de los Ítems #
# (FREI) para los Modelos de Respuesta Graduada (GRM) #
# #
# Asesor : Jorge Luis Bazán #
# #
# Versión : 1.0 #
# #
# Última Modificación : 14-02-2013 #
#####

frei.grm.pirt<- function(item,
                        rangex=c(-4,4),
                        model=c("2PL", "LPE"),
                        main=""
                        )
{
  if(model=="2PL"){
    m<-length(item)
    for(k in 1:(m-1)){
      Pac<-function(x) 1-plogis(item[m] * (item[k]- x))
      plot(Pac,add=k>1,xlim=rangex,col=k,xlab="Habilidad (u)",
           ylab="Probabilidad",lty=k, lwd=2, main=main)
    }
  }

  if(model=="LPE"){
    m<-length(item)-1
    for(k in 1:(m-1)){
      Pac<-function(x) (plogis(-item[m] * (item[k]- x)))^item[m+1]
      plot(Pac,add=k>1,ylim=c(0,1),xlim=rangex,col=k,xlab="Habilidad (u)",
           ylab="Probabilidad",lty=k, lwd=2,main=main)
    }
  }
}
}

```

#### A.8. Código en R para graficar la Función de Respuesta a las Categorías de los Ítems (FRCI) del GRM

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```
#####
# Autor : Enver Gerald Tarazona Vargas #
# # #
# Función frci.grm.pirt() #
# # #
# Fecha de inicio : 13-02-2013 #
# # #
# Función : Gráficar la Función de Respuesta a las Categorías de los Ítems #
# (FRCI) para los Modelos de Respuesta Graduada (GRM) #
# # #
# Asesor : Jorge Luis Bazán #
# # #
# Versión : 1.0 #
# # #
# Última Modificación : 14-02-2013 #
#####

frci.grm.pirt<-function(item,
                        rangex=c(-4,4),
                        model=c("2PL", "LPE"),
                        main=""
                        )
{

if(model=="2PL"){
  m<-length(item)
  for(k in 1:m){
    P<-function(x) {
      if(k==1){
        plogis(item[m] * (item[k]- x))
      }
      else{
        if(k< m){
          plogis(item[m] * (item[k]- x))-plogis(item[m] * (item[k-1]- x))
        }
        else{
          1-plogis(item[m] * (item[k-1]- x))
        }
      }
    }
  }
  plot(P,add=k>1,xlim=rangex,col=k,xlab="Habilidad (u)",
```

```

        ylab="Probabilidad",lty=k, lwd=2,main=main)
    }

}

if(model=="LPE"){
  m<-length(item)-1
  for(k in 1:m){
    Pac<-function(x) {
      if(k==1){
        (plogis(item[m] * (item[k]- x)))^item[m+1]
      }
      else{
        if(k < m){
          (plogis(item[m] * (item[k]- x)))^item[m+1]-
            (plogis(item[m] * (item[k-1]- x)))^item[m+1]
        }
        else{
          1-(plogis(item[m] * (item[k-1]- x)))^item[m+1]
        }
      }
    }
    plot(Pac,add=k>1,xlim=rangex,col=k,xlab="Habilidad (u)",
         ylab="Probabilidad",lty=k, lwd=2,main=main)
  }
}
}

```

### A.9. Código en R para graficar la Función de Información de un Ítem (FII) del GRM

```

#####
# Autor : Enver Gerald Tarazona Vargas #
# # #
# Función fiii.grm.pirt() #
# # #
# Fecha de inicio : 13-02-2013 #
# # #
# Función : Graficar la Función de Información de los Ítems (FII) #
# para los Modelos de Respuesta Graduada (GRM) #
# # #

```

## APÉNDICE A. CÓDIGOS EN R Y BUGS

```

# Asesor : Jorge Luis Bazán                                     #
#                                                                 #
# Versión : 1.0                                                #
#                                                                 #
# Última Modificación : 14-02-2013                             #
#####

fii.grm.pirt<-function(item,
                        rangex=c(-4,4),
                        ylim=c(0,3.5),
                        col=1,lty=1,add=T,
                        model=c("2PL", "LPE"),
                        main=""
)
{

if(model=="2PL"){
  m<-length(item)
  fii<-function(x) {
    firi<-0
    for(k in 1:m){
      if(k==1){
        Pac<-plogis(item[m] * (item[k]- x))
        Pac0<-0
        p<-Pac-Pac0
      }
      else{
        if(k< m){
          Pac<-plogis(item[m] * (item[k]- x))
          Pac0<-plogis(item[m] * (item[k-1]- x))
          p<-Pac-Pac0
        }
        else{
          Pac<-1
          Pac0<-plogis(item[m] * (item[k-1]- x))
          p<-Pac-Pac0
        }
      }
    }
    firi<-firi+((1-Pac)*Pac-(1-Pac0)*Pac0)^2/p
  }
}
}

```



```

      item[m]^2*firi
    }

  }

if(model=="LPE"){
  m<-length(item)-1
  fii<-function(x) {
    firi<-0
    for(k in 1:m){
      if(k==1){
        Pac<-plogis(item[m] * (item[k]- x))
        Pac0<-0
      }
      else{
        if(k< m){
          Pac<-plogis(item[m] * (item[k]- x))
          Pac0<-plogis(item[m] * (item[k-1]- x))
        }
        else{
          Pac<-1
          Pac0<-plogis(item[m] * (item[k-1]- x))
        }
      }
    }
    p<-Pac^item[m+1]-Pac0^item[m+1]
    firi<-firi+((1-Pac)*Pac^item[m+1]-(1-Pac0)*Pac0^item[m+1])^2/p
  }
  (item[m]*item[m+1])^2*firi
}

}

plot(fii,add=add,xlim=rangex,col=col,xlab="Habilidad (u)",
      ylab="Informacion",lty=lty, lwd=2,ylim=ylim,main=main)
}

```

## Apéndice B

### Diagnósticos de convergencia

#### B.1. Diagnóstico de convergencia en la simulación

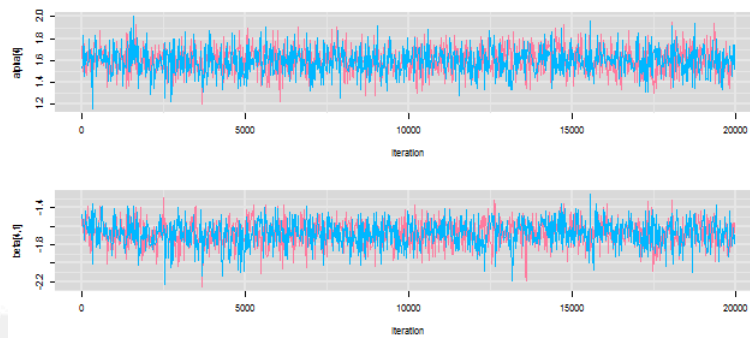


Figura B.1: Histórico de dos cadena para los parámetros del Ítem 4

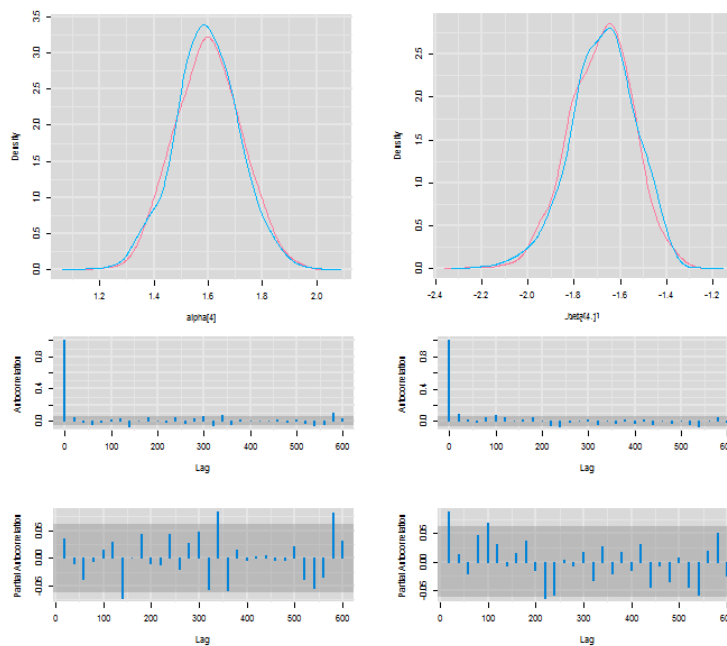


Figura B.2: Gráficas de las densidades y autocorrelaciones para los parámetros del Ítem 4

B.2. Diagnóstico de convergencia en la aplicación

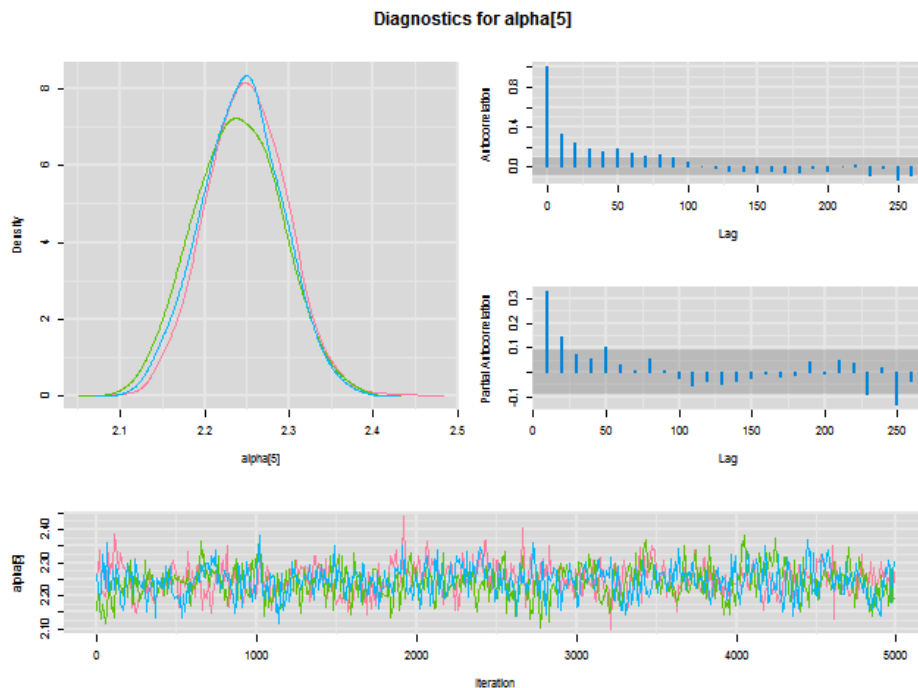


Figura B.3: Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el parámetro de discriminación del Ítem 9, tiempo de atención, estimado con el Modelo 2PL-GRM

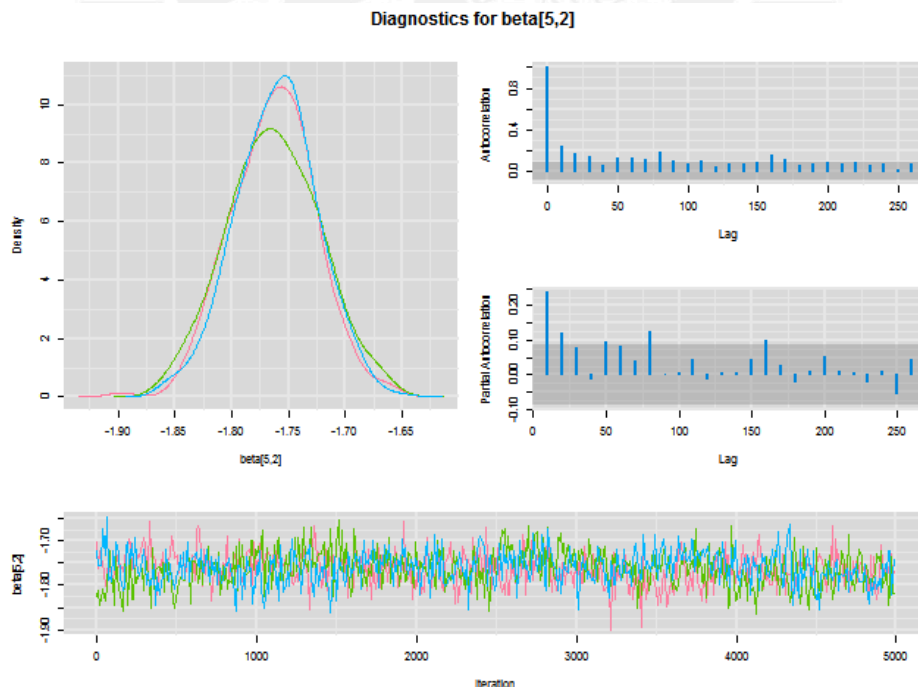


Figura B.4: Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el segundo umbral (parámetro de dificultad) del Ítem 9, tiempo de atención, estimado con el Modelo 2PL-GRM

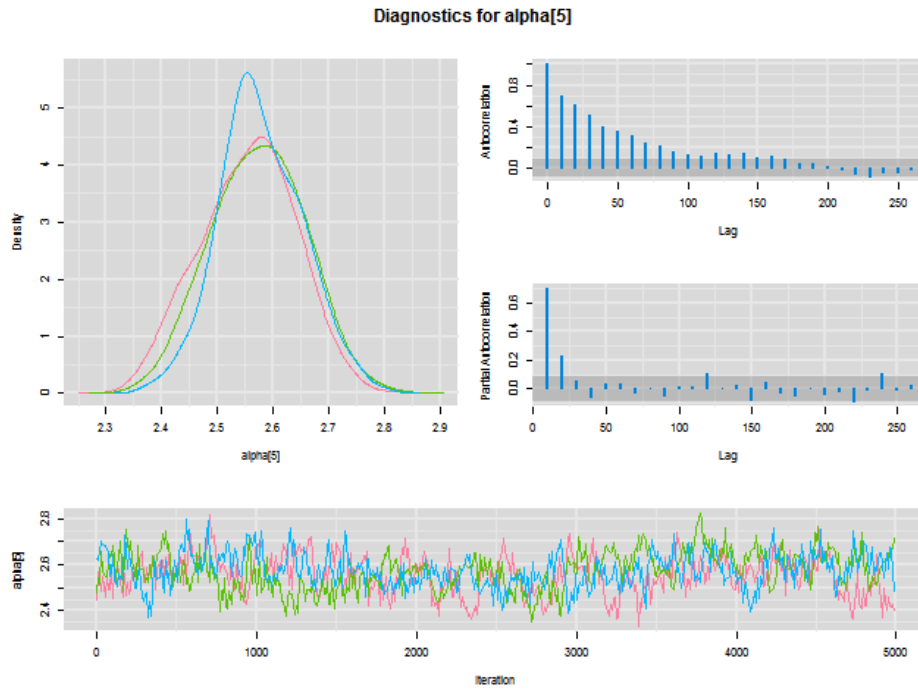


Figura B.5: Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el parámetro de discriminación del Ítem 9, tiempo de atención, estimado con el Modelo LPE-GRM

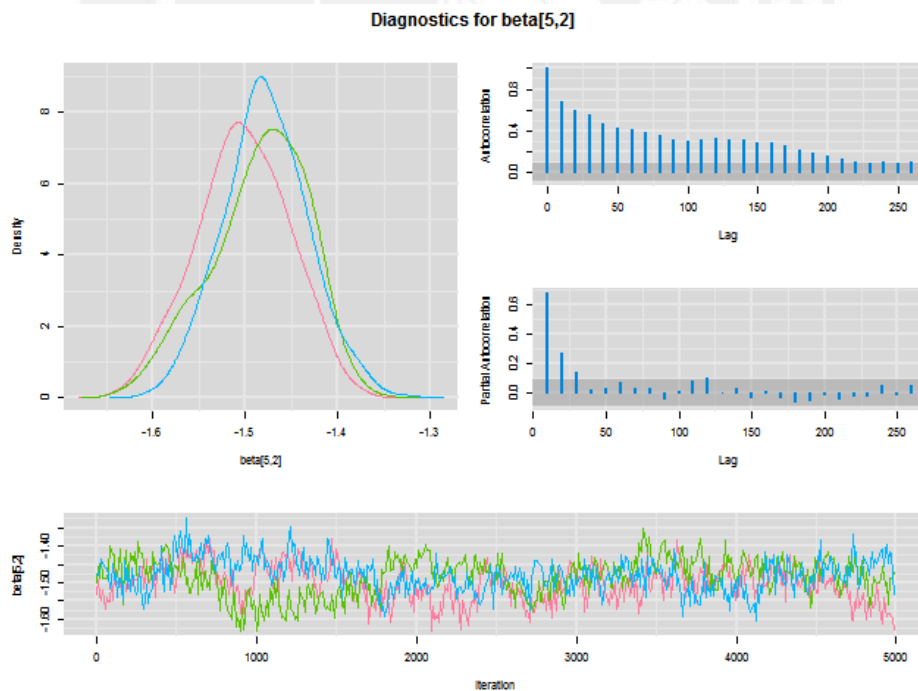


Figura B.6: Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el segundo umbral (parámetro de dificultad) del Ítem 9, tiempo de atención, estimado con el Modelo LPE-GRM

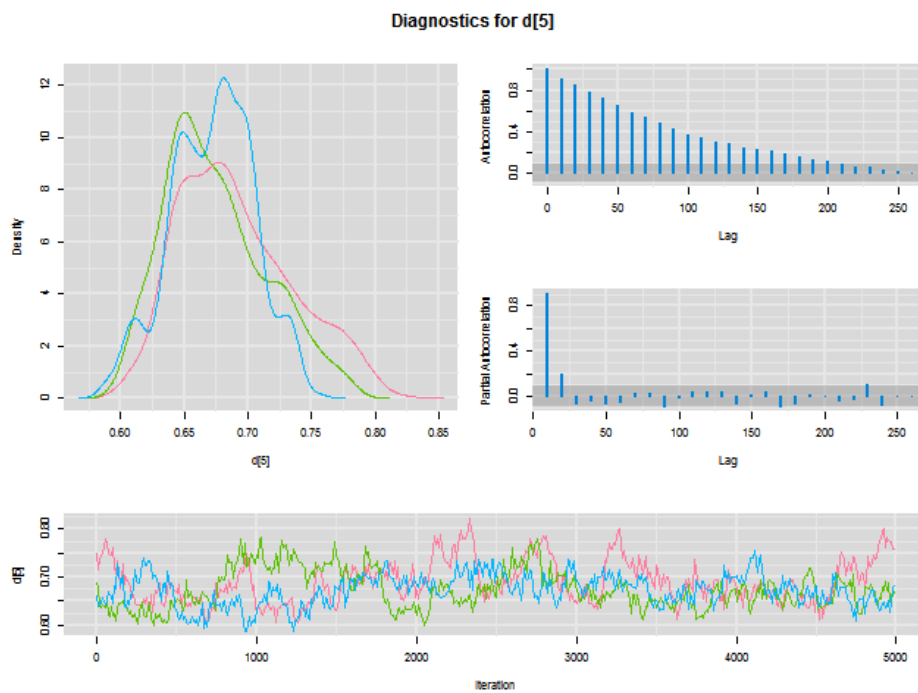


Figura B.7: Histórico de las tres cadenas y gráficas de las densidades y autocorrelaciones para el parámetro de penalización del Ítem 9, tiempo de atención, estimado con el Modelo LPE-GRM

## Apéndice C

### Cuestionario de Satisfacción: Atributos medidos

Ítem	Descripción
1	Facilidad de comunicarse con la central
2	Utilidad de poder seleccionar el idioma
3	Sencillez del menú de opciones
4	Tiempo de espera hasta que se contestó la llamada
5	Amabilidad y cordialidad del asesor
6	Interés y compromiso por atender la llamada
7	Capacidad o nivel de conocimiento del personal para dar una solución
8	Claridad de la información brindada
9	Tiempo de atención
10	Tiempo hasta que terminó la gestión
11	Solución brindada a la gestión

Cuadro C.1: Ítems considerados dentro del cuestionario de satisfacción con la atención de un Call Center

## Bibliografía

- Allen, D. (2004). *Customer Satisfaction Research Management: a comprehensive guide to integrating customer loyalty and satisfaction metrics*, ASQ Quality Press, Milwaukee.
- Azevedo, C. (2003). *Métodos de estimação na teoria de resposta ao item*, Master's thesis, IME-USP.
- Baker, F. B. y Kim, S. (2004). *Item Response Theory: Parameter estimation techniques*, M Dekker, Inc., New York, USA.
- Bartholomew, D. J., Knott, M. y Moustaki, I. (2011). *Latent variable models and factor analysis: a unified approach*, John Wiley & Sons, Ltd, Chichester, UK.
- Bazán, J. (2005). *Uma familia de modelos de resposta ao item normal asimétrica*, Master's thesis, IME-USP.
- Bazán, J. (2011). Modelos assimétricos em tri, Minicurso desarrollado en el 2º Congresso Brasileiro de Teoria de Resposta ao Item. Salvador, BAHIA, de 7 a 9 de Dezembro de 2011.
- Bazán, J., Branco, M. D. y Bolfarine, H. (2006). A skew item response model., *Bayesian Analysis* **1**: 861–892.
- Bazán, J., Calderon, A. y Valdivieso, L. (2010). Enfoque bayesiano en modelos de teoria de respuesta al item, Reporte del Minicurso desarrollado en la I Jornada Internacional de Probabilidad y Estadística. Lima, 3-5 de febrero de 2010.
- Bazán, J. L., Mazzon, J. y Hernani, M. (2012). Clasificação de modelos de resposta ao item policotômicos com aplicação ao marketing, *Revista Brasileira de Estatística* pp. 460–469.
- Bock, R. D. y Moustaki, I. (2007). Item response theory in a general framework, in C. R. Rao y S. Sinharay (eds), *Handbook of Statistics, vol. 26*, Elsevier Science B.V., Amsterdam, The Netherlands, pp. 469–513.
- Bolfarine, H. y Bazán, J. (2010). Bayesian estimation of the logistic positive exponent irt model, *Journal of Educational Behavioral Statistics* pp. 693–713.
- Busacca, B. y Padula, G. (2005). Understanding the relationship between attribute performance and overall satisfaction. theory, measurement and implications, *Marketing Intelligence and Planning* **23**: 543–561.
- Chang, H.-H. y Ying, Z. (1996). A global information approach to computerized adaptive testing, *Applied Psychological Measurement* **20**: 213–230.
- Cronbach, L. (1951). Coeficient alpha and the internal structure of tests, *Psychometrika* **16**: 297–334.
- Curtis, S. M. (2010). BUGS Code for Item Response Theory, *Journal of Statistical Software, Code Snippets* **36**: 1–34.

- Dagostim, J. y Tontini, G. (2008). Análise do contraste da penalidade e da recompensa (PRC): identificando oportunidades de melhoria em um serviço, *RAM-Revista de Administração Mackenzie* **9**: 35–58.
- De Ayala, R. (1994). The influence of dimensionality on the graded response model, *Applied Psychological Measurement* **18**: 155–170.
- De Battisti, F., Nicolini, G. y Salini, S. (2008). Methodological overview of rasch model and application in customer satisfaction survey data, *Departmental Working Papers 2008-04, Department of Economics, Management and Quantitative Methods at Università degli Studi di Milano*.
- De Battisti, F., Nicolini, G. y Salini, S. (2012). The rasch model, in R. Kenett y S. Salini (eds), *Modern Analysis of Customer Surveys: with applications using R*, John Wiley & Sons, Ltd, Chichester, UK, pp. 259–281.
- De Battisti, F., Nicolini, G. y Salnini, S. (2005). The rash model to measure service quality, *Journal of Services Marketing* **3(3)**: 58–80.
- De Battisti, F., Nicolini, G. y Salnini, S. (2010). The rash model in customer satisfaction survey data, *Quality Technology and Quantitative Management* **7(1)**: 15–34.
- Fisher, R. (1925). Theory of statistical estimation, *Proceedings of the Cambridge Philosophical Society* **22**: 700–725.
- Flores, S. (2012). *Modelos testlet logísticos y logísticos de exponente positivo para pruebas de comprensión de textos*, Master's thesis, PUCP.
- Hafsteinsson, L. G., Donovan, J. J. y Breland, B. T. (2007). An item response theory examination of two popular goal orientation measures, *Educational and Psychological Measurement* **67**: 719–739.
- Hambleton, R. K. y Swaminathan, H. (1985). *Item Response Theory: Principles and applications*, Kluwer Nijhoff, Boston, USA.
- Harwell, M., Stone, C., Hsu, T. y Kirisci, L. (1996). Monte carlo studies in item response theory, *Applied Psychological Measurement* **20**: 101–125.
- Hemker, B. (2001). Reversibility revisited and other comparison of tree types of polytomous IRT models, in A. Boomsma, M. Varn Duijn y T. Snijders (eds), *Essays on item response theory*, Springer-Verlag, New York, pp. 277–296.
- Johnson, M. S. (2007). Marginal Maximum Likelihood Estimation of Item Response Models in R, *Journal of Statistical Software* **20(10)**: 1–24.  
**URL:** "<http://www.jstatsoft.org/v20/i10>"
- Kieftenbeld, V. y Natesan, P. (2012). Recovery of graded response model parameters: A comparison of marginal maximum likelihood and markov chain monte carlo estimation, *Applied Psychological Measurement* **36(5)**: 399–419.
- Lautenschlager, G., Meade, A. y Kim, S. (2006). Cautions regarding sample characteristics when using the graded response model, Poster. Presented at the annual conference of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Lawley, D. N. (1943). On problems connected with item selection and test construction, *Proceedings of the Royal Society of Edinburgh, Series A* **23**: 273–287.



- Lewis, R. y Booms, B. (1983). The marketing aspects of quality, *Emerging perspectives of services marketing* pp. 99–107.
- Lord, F. (1980). *Applications of item response theory to practical testing problems*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- Lord, F. y Novick, M. (1968). *Statistical theories of mental test scores (with contributions by Allen Birnbaum)*, Addison-Wesley, Reading, MA.
- Lunn, D., Thomas, A., Best, N. y Spiegelhalter, D. (2000). WinBUGS - a Bayesian modeling framework: concepts, structure, and extensibility, *Statistics and Computing* **10**: 325–337.
- Mikulić, J. y Prebežac (2008). Prioritizing improvement of service attributes using impact range-performance analysis and impact-asymmetry analysis, *Managing Service Quality* **18**: 559–576.
- Molenaar, D., Dolan, C. V. y De Boeck, P. (2012). The heteroscedastic graded response model with a skewed latent trait: testing statistical and substantive hypotheses related to skewed item category functions, *Psychometrika* **3**: 455–478.
- Muraki, E. (1990). Fitting a polytomous item response model to likert-type data, *Applied Psychological Measurement* **14**: 59–71.
- Neyman, J. y Scott, E. L. (1948). Consistent estimates based on partially consistent observations, *Econometrika* **16**(1): 1–32.
- Nicolini, G. y Salini, S. (2006). Customer satisfaction in the airline industry: The case of british airways, *Quality and Reliability Engineering International* **22**: 1–9.
- Oliver, R. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions, *Journal of Marketing Research* **42**: 460–469.
- Ostini, R. y Nering, M. (2006). *Polytomous Item Response Theory Models*, Sage Publications, Thousand Oaks CA.
- Ostini, R. y Nering, M. (2010). New perspectives and applications, in M. Nering y R. Ostini (eds), *Handbook of Polytomous Item Response Theory Models*, Routledge Academic, London, UK, pp. 3–20.
- Pasquali, L. (2009). *Psicometria: teoria dos testes na psicologia e na educaçao*, Editora Vozes Ltda., Brasil.
- Patz, R. J. y Junker, B. W. (1999). A straightforward approach to markov chain monte carlo methods for item response models, *Journal of Educational and Behavioral Statistics* **24**: 146–178.
- Pfanzagl, J., Baumann, V. y Huber, H. (1971). *Theory of Measurement*, Physica-Verl.
- R Core Team (2012). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.  
**URL:** <http://www.R-project.org/>
- Rao, C. R. y Sinharay, S. (2007). *Handbook of Statistics, Vol. 26: Psychometrics*, Elsevier Science B.V., Amsterdam, The Netherlands.
- Reckase, M. (2009). *Multidimensional Item Response Theory*, Statistics for social and behavioral sciences, Springer-Verlag, New York.

- Rizopoulos, D. (2006). ltm: An R Package for Latent Variable Modeling and Item Response Theory Analyses, *Journal of Statistical Software* **17(5)**: 1–25.  
**URL:** "<http://www.jstatsoft.org/v17i05/>"
- Sahu, S. K. (2002). Bayesian estimation and model choice in item response models, *Journal Statistical Computing Simulation* **72**: 217–232.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores, *Psychometrika Monograph* **17**: 1–100.
- Samejima, F. (1972). A general model of free-response data, *Psychometrika Monograph* **18**.
- Samejima, F. (1995). Acceleration model in the heterogeneous case of the general graded response model, *Psychometrika* **60**: 549–572.
- Samejima, F. (1996). Polychotomous responses and the test score, Paper. Presented at the 1996 National Council on Measurement in Education, New York.
- Samejima, F. (1997a). Departure from normal assumptions: A promise for future psychometrics with substantive mathematical modeling, *Psychometrika* **62**: 471–493.
- Samejima, F. (1997b). Graded response model, in W. J. Van Linden y R. K. Hambleton (eds), *Handbook of modern item response theory*, Springer-Verlag, New York, pp. 85–100.
- Samejima, F. (2000). Logistic positive exponent family of models: Virtue of asymmetric item characteristics curves, *Psychometrika* **65**: 319–335.
- Samejima, F. (2004). Graded response model, in K. Kempf-Leonard (ed.), *Encyclopedia of social measurement*, Vol. 2, Elsevier, Amsterdam, pp. 145–153.
- Samejima, F. (2010). The Graded Response Model, in M. Nering y R. Ostini (eds), *Handbook of Polytomous Item Response Theory Models*, Routledge Academic, London, UK, pp. 77–107.
- Savalei, V. (2006). Logistic approximation to the normal: the KL rationale., *Psychometrika* **71**: 763–767.
- Singh, J., Howell, R. D. y Rhoads, G. K. (1990). Adaptive designs for likert-type data: An approach for implementing marketing surveys, *Journal of Marketing Research* **27(3)**: 304–321.
- Spiegelhalter, D. J., Best, N., Carlin, B. y Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion), *Journal of the Royal Statistical Society, Series B* **64**: 583–640.
- Sturtz, S., Ligges, U. y Gelman, A. (2005). R2winbugs: A package for running winbugs from r, *Journal of Statistical Software* **12(3)**: 1–16.  
**URL:** <http://www.jstatsoft.org>
- Tarazona, E. y Bazán, J. L. (2011). Psychometric examination of a customer satisfaction questionnaire using polytomous item response theory models, Poster. Presentado en el 2° Congresso Brasileiro de Teoria de Resposta ao Item (II CONBRATRI). Salvador, Bahia.
- Tarazona, E. y Bazán, J. L. (2012). Evaluación psicométrica de un cuestionario de satisfacción de clientes usando modelos politómicos de la teoría de respuesta al Ítem, Poster. Presentado en la II Jornada Internacional de Probabilidad y Estadística (JIPE II). Pontificia Universidad Católica del Perú, Lima. Febrero, 1-3.

- Wang, X., Bradlow, E. T. y Wainer, H. (2002). A general bayesian model for testlets: Theory and applications, *Applied Psychological Measurement* **26**: 109–128.
- Zhu, X. y Stone, C. (2012). Bayesian comparison of alternative graded response models for performance assessment applications, *Educational and Psychological Measurement* **72**(5): 774–799.

