

**PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ**

**ESCUELA DE POSGRADO**



**MODELO SECUENCIAL CON APLICACIÓN A LA  
MEDICIÓN DEL RENDIMIENTO ESTUDIANTIL**

Tesis para optar el grado de Magíster en Estadística

**AUTOR**

Luis Ángel Mejía Campos

**ASESOR**

Mg. Enver Gerald Tarazona Vargas

**JURADO**

Dr. Luis Hilmar Valdivieso Serrano

Dr. Cristian Luis Bayes Rodríguez

**LIMA - PERÚ**

**2018**

## Dedicatoria

A mi tía por darme acogida en su familia, y a mi madre por darme todo lo demás.



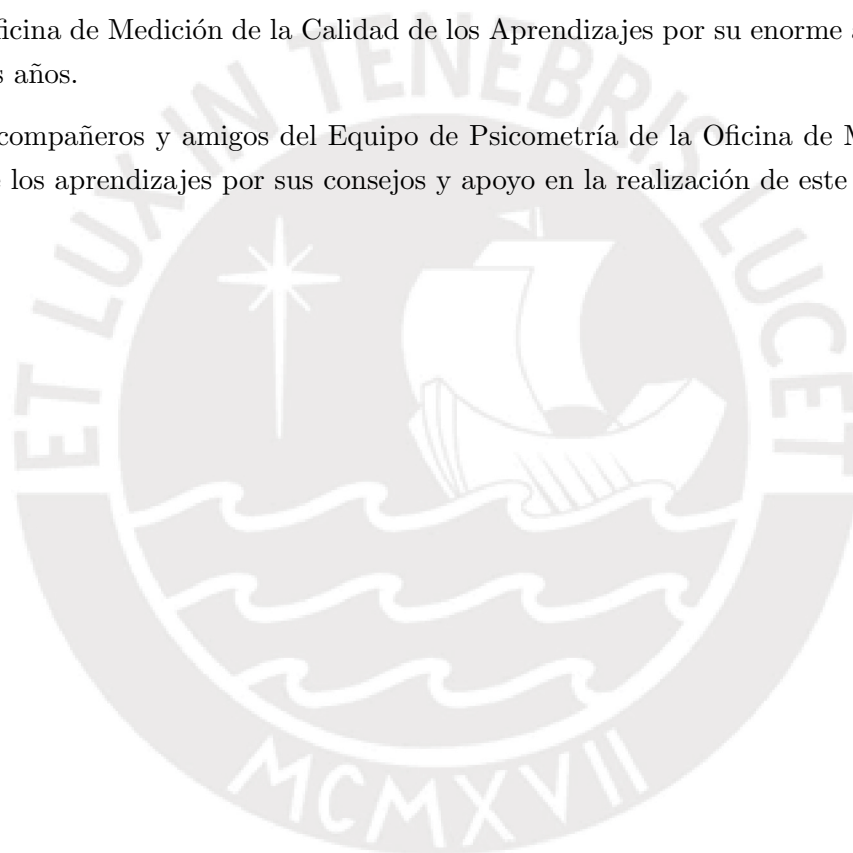
## Agradecimientos

Al profesor Enver Tarazona por todo su apoyo y orientación en el presente trabajo.

A los profesores de la Maestría en Estadística de la Pontificia Universidad Católica del Perú (PUCP) por todo lo aprendido en estos años.

A la Oficina de Medición de la Calidad de los Aprendizajes por su enorme apoyo durante todos estos años.

A mis compañeros y amigos del Equipo de Psicometría de la Oficina de Medición de la Calidad de los aprendizajes por sus consejos y apoyo en la realización de este trabajo.



## Resumen

En este trabajo se presenta el Modelo Secuencial para datos polítomicos ordinales de la teoría de respuesta al ítem y sus características. De forma específica se estudia el Modelo Secuencial Logístico de 2 parámetros (2PL-SM). La estimación de este modelo se realiza utilizando Métodos de Cadenas de Markov de Montecarlo (MCMC), los cuales fueron implementados en R y WinBUGS.

Se realizó un estudio de simulación con el objetivo de estudiar la precisión en la recuperación de parámetros observándose resultados apropiados según los índices de precisión utilizados.

El Modelo Secuencial en estudio fue luego aplicado a la prueba de escritura de la Evaluación Muestral 2013 del Ministerio de Educación, evaluación que fue aplicada a una muestra de 4327 estudiantes de sexto grado de primaria de todo el país. Con la aplicación del modelo a la prueba se pudo determinar que en general esta contiene ítems cuyas dificultades son bajas y que, para los estudiantes, el enfrentarse a esta prueba no debería resultarles complicado.

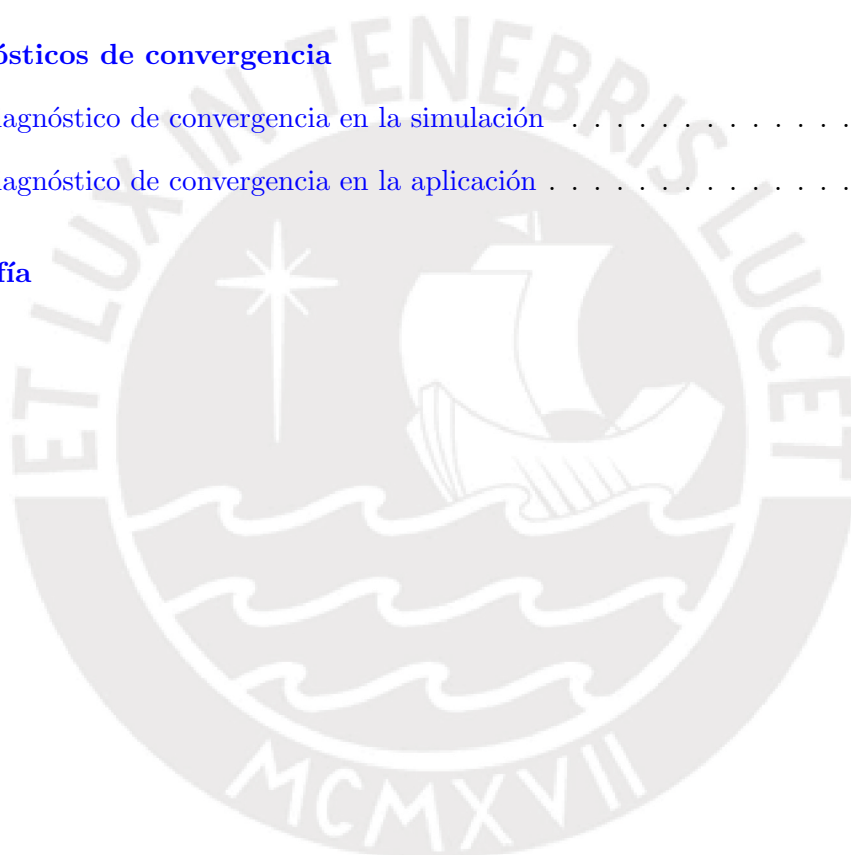
**Palabras-clave:** teoría de respuesta al ítem (TRI), modelos politómicos de teoría de respuesta al ítem (TRIP), variables latentes, inferencia bayesiana, Modelo Secuencial, métodos montecarlo vía cadenas de markov (MCMC).

# Índice general

<b>Lista de Abreviaturas</b>	<b>VIII</b>
<b>Lista de Símbolos</b>	<b>IX</b>
<b>Índice de figuras</b>	<b>X</b>
<b>Índice de tablas</b>	<b>XII</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Consideraciones Preliminares . . . . .	1
1.2. Objetivos . . . . .	3
1.3. Organización del Trabajo . . . . .	3
<b>2. Modelo Secuencial</b>	<b>5</b>
2.1. Modelos de respuesta al ítem politómicos . . . . .	5
2.1.1. Modelo Politómico Nominal . . . . .	5
2.1.2. Modelo Politómico Ordinal . . . . .	6
2.2. Formulación del Modelo Secuencial . . . . .	9
2.2.1. Presentación y racionalidad del modelo . . . . .	9
2.2.2. Observaciones . . . . .	10
2.2.3. Supuestos del Modelo Secuencial . . . . .	12
2.2.4. Características del Modelo Secuencial . . . . .	13
2.2.5. Casos particulares del Modelo Secuencial . . . . .	17
2.3. Modelo Secuencial Logístico . . . . .	18
2.3.1. Definición del Modelo . . . . .	18
2.3.2. Características del Modelo 2PL-SM . . . . .	18

<b>3. Estimación Bayesiana del Modelo Secuencial Logístico</b>	<b>24</b>
3.1. Función de Verosimilitud para el Modelo Secuencial Logístico (2PL-SM) . . . .	24
3.2. Especificación de la distribución a priori en el Modelo Secuencial Logístico (2PL-SM) . . . . .	25
3.3. Distribución a posteriori de los parámetros del Modelo Secuencial Logístico (2PL-SM) . . . . .	25
3.4. Inferencia Bayesiana del Modelo Secuencial Logístico (2PL-SM) . . . . .	26
3.5. Implementación Computacional . . . . .	26
<b>4. Estudio de Simulación</b>	<b>29</b>
4.1. Objetivos . . . . .	29
4.2. Algoritmo para la simulación de datos . . . . .	29
4.3. Metodo para la estimación de parámetros . . . . .	30
4.4. Criterios para evaluar la simulación . . . . .	30
4.5. Escenario . . . . .	31
4.6. Resultados . . . . .	31
4.7. Conclusiones . . . . .	32
<b>5. Aplicación</b>	<b>34</b>
5.1. Instrumento de medición . . . . .	34
5.2. Diseño Muestral . . . . .	36
5.3. Recolección de datos . . . . .	36
5.4. Estimación de los parámetros mediante el Modelo Secuencial . . . . .	36
5.4.1. Modelo Secuencial Logístico (2PL-SM) . . . . .	36
<b>6. Conclusiones</b>	<b>42</b>
6.1. Conclusiones . . . . .	42
6.2. Sugerencias para investigaciones futuras . . . . .	43
<b>A. Códigos en R y BUGS</b>	<b>44</b>
A.1. Código en R para graficar la Función de Respuesta a las Etapas de los Ítems (FREI) del SM . . . . .	44

A.2. Código en R para graficar la Función de Respuesta a las Categorías de los Ítems (FRCI) del SM . . . . .	45
A.3. Código en R para graficar la Función de Información de los Ítems (FII) del SM	46
A.4. Código en R para graficar la Función de Información del Test (FIT) del SM .	49
A.5. Código en R para simular datos provenientes del Modelo Secuencial . . . . .	51
A.6. Función en R para realizar la estimación bayesiana del modelo secuencial usando MCMC . . . . .	52
A.7. Código en BUGS para el Modelo Secuencial Logístico (2PL-SM) . . . . .	54
A.8. Código en R para el estudio de simulación . . . . .	55
<b>B. Diagnósticos de convergencia</b>	<b>63</b>
B.1. Diagnóstico de convergencia en la simulación . . . . .	63
B.2. Diagnóstico de convergencia en la aplicación . . . . .	64
<b>Bibliografía</b>	<b>65</b>

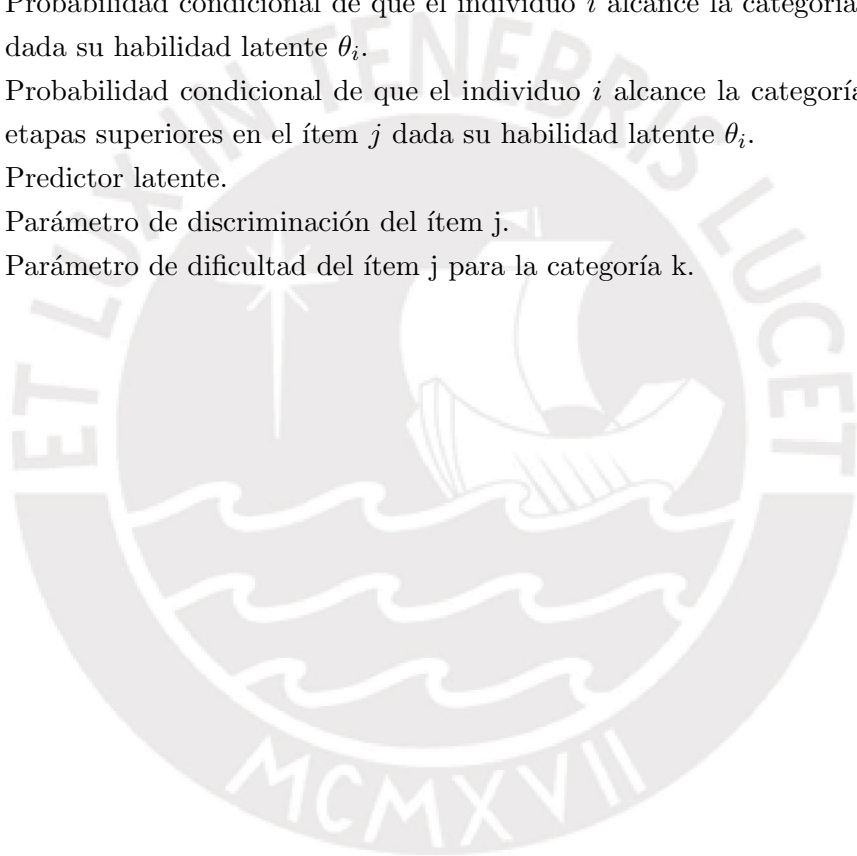


## Lista de Abreviaturas

TRI	Teoría del respuesta al Ítem.
2PL-SM	Modelo Secuencial Logístico de 2 parámetros.
CCI	Curva Característica del Ítem. Similar a FREI.
FREI	Función de Respuesta a las Etapas del Ítem.
FRCI	Función de Respuesta a las Categorías del Ítem.
FII	Función de Información del Ítem.
FIT	Función de Información del Test.
JML	Máxima Verosimilitud Conjunta ( <i>Joint Maximum Likelihood</i> ).
MML	Máxima Verosimilitud Marginal ( <i>Marginal Maximum Likelihood</i> ).
MCMC	Métodos Montecarlo vía Cadenas de Markov.
RMSE	Raíz del Error Cuadrático Medio ( <i>Root Mean Square Deviation</i> ).
MAE	Promedio del error absoluto ( <i>Mean Absolute Error</i> ).
EM 2013	Evaluación Muestral 2013.
UMC	Oficina de Medición de la Calidad de los Aprendizajes.

## Lista de Símbolos

$Y_{ij}$	Variable manifiesta que modela la respuesta del individuo $i$ en el ítem $j$ en un test.
$\theta_i$	Valor de la variable latente $\Theta_i$ (habilidad o rasgo latente) para el individuo $i$ .
$\xi_j$	Vector de parámetros del ítem $j$ .
$P_{ijk}$	Probabilidad condicional de que el individuo $i$ alcance la categoría $k$ en el ítem $j$ dada su habilidad latente $\theta_i$ .
$P_{ijk}^+$	Probabilidad condicional de que el individuo $i$ alcance la categoría $k$ y todas las etapas superiores en el ítem $j$ dada su habilidad latente $\theta_i$ .
$\eta_{ijk}$	Predictor latente.
$a_j$	Parámetro de discriminación del ítem $j$ .
$b_{jk}$	Parámetro de dificultad del ítem $j$ para la categoría $k$ .



## Índice de figuras

2.1. Representación gráfica de las FREI del Modelo 2PL-SM para tres ítems de 4 categorías con parámetros de discriminación $a_1 = 1$ y $a_2 = a_3 = 2$ , y con parámetros de dificultad $b_{1r} = b_{2r} = -2.0, 0.1, 0.8, 1.4$ y $b_{3r} = -2.0, 0.1, -0.5, 1.4$ .	20
2.2. Representación gráfica de las FRCI del Modelo 2PL-SM para seis ítems de cuatro categorías con parámetros de discriminación $a_1 = a_3 = a_5 = 1$ y $a_2 = a_4 = a_6 = 2$ ; y con parámetros de dificultad $b_{1r} = b_{2r} = -3.5, -2.0, -0.5, 1.0, 1$ , $b_{3r} = b_{4r} = -1.0, 0.0, 1.0, 2.0, 1$ y $b_{5r} = b_{6r} = -0.5, 1.0, 2.5, 3.5, 1$ . . . . .	21
2.3. Representación gráfica de las FII del Modelo 2PL-SM para seis ítems de cuatro categorías con parámetros de discriminación $a_1 = a_3 = a_5 = 1$ y $a_2 = a_4 = a_6 = 2$ ; y con parámetros de dificultad $b_{1r} = b_{2r} = -3.5, -2.0, -0.5, 1.0, 1$ , $b_{3r} = b_{4r} = -1.0, 0.0, 1.0, 2.0, 1$ y $b_{5r} = b_{6r} = -0.5, 1.0, 2.5, 3.5, 1$ . . . . .	23
2.4. Representación gráfica de la FIT del Modelo 2PL-SM para un test de seis ítems, de cuatro categorías, con parámetros de discriminación $a_1 = a_3 = a_5 = 1$ y $a_2 = a_4 = a_6 = 2$ ; y con parámetros de dificultad $b_{1r} = b_{2r} = -3.5, -2.0, -0.5, 1.0, 1$ , $b_{3r} = b_{4r} = -1.0, 0.0, 1.0, 2.0, 1$ y $b_{5r} = b_{6r} = -0.5, 1.0, 2.5, 3.5, 1$ . . . . .	23
5.1. FREI del modelo 2PL-SM para los ítems 5 (Mantenimiento del tema) y 11 (Estructura narrativa) . . . . .	37
5.2. FRCI del modelo 2PL-SM para los ítems 5 (Mantenimiento del tema) y 11 (Estructura narrativa) . . . . .	38
5.3. FII del modelo 2PL-SM para la prueba de escritura . . . . .	40
5.4. FIT del modelo 2PL-SM para la prueba de escritura . . . . .	40
5.5. Distribución de la habilidad estimada de los estudiantes de 6to de primaria en la prueba de escritura usando el Modelo 2PL-SM . . . . .	41
B.1. Histórico de dos cadenas para los parámetros del ítem 4 . . . . .	63
B.2. Gráficas de densidades y autocorrelaciones para los parámetros del ítem 4 . . . . .	63

B.3. Histórico de dos cadenas, gráficas de densidades y autocorrelaciones para el parámetro de discriminación del Ítem3, Adecuación al tema, estimado con el Modelo 2PL-SM . . . . . 64

B.4. Histórico de dos cadenas, gráficas de densidades y autocorrelaciones para el segundo umbral del Ítem 9, Referencia, estimado con el Modelo 2PL-SM . . . . . 64



## Índice de tablas

4.1. Parámetros de los ítems usados en la simulación . . . . .	31
4.2. Sesgo en la recuperación de parámetros de los ítems usando estimación bajo MCMC para el modelo 2PL-SM . . . . .	32
4.3. Raíz del error cuadrático medio (RMSE) 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-SM) considerando R=20 réplicas . . . .	32
4.4. 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 Secuencial Logístico (2PL-SM) considerando R= 20 réplicas . . . . .	32
5.1. Estimación de los parámetros de discriminación del Modelo Secuencial 2PL-SM	38
5.2. Estimación de los parámetros de dificultad del Modelo Secuencial 2PL-SM . .	39

# Capítulo 1

## Introducción

### 1.1. Consideraciones Preliminares

La medición del rendimiento escolar utilizando la teoría de respuesta al ítem (TRI) se ha venido desarrollando con gran popularidad a nivel internacional. Pruebas de rendimiento como PISA (Programme for International Student Assessment) (OECD, 2017), TERCE (Tercer Estudio Regional Comparativo y Explicativo) (OREAL, 2016) e ICCS (International Civic and Citizenship Study) (IEA, 2011) vienen utilizando la TRI desde hace ya muchos años. En el ámbito local, la medición del rendimiento estudiantil utilizando esta teoría no es algo nuevo pero tampoco es algo muy generalizado. Generalmente las pruebas de rendimiento utilizan ítems dicotómicos, que son aquellos con solo dos categorías (correcto e incorrecto). El rendimiento es estimado utilizando el modelo Rasch, el cual permite la estimación conjunta de la dificultad de los ítems y la habilidad del estudiante. Por otro lado la teoría de respuesta al ítem considera también otros modelos, llamados de dos parámetros, que además de la habilidad del estudiante y la dificultad del ítem incorpora la discriminación del ítem. Este parámetro permite detectar que ítems son aquellos que diferencian más a los estudiantes en cuanto a sus habilidades.

Según Michael y Ostini (2010) los ítems politómicos, que son aquellos con más de dos categorías, se han vuelto muy populares en pruebas educacionales y psicológicas debido a que ellas ofrecen una más rica experiencia de medición además de mayor información psicométrica acerca del constructo medido. En el Perú las pruebas de rendimiento estudiantil que tienen ítems politómicos no son muy frecuentes y cuando se construyen, los ítems parten de respuestas abiertas que luego son codificadas por especialistas. Esta codificación es imposible de hacer para una evaluación de gran magnitud dados los tiempos que se tienen para presentar resultados, sin embargo hay algunas evaluaciones a menor escala que sí incluyen este tipo de ítems.

Según Van Der Ark (2001) las ventajas de los ítems politómicos frente a los dicotómicos son las siguientes:

- Se necesitan pocos ítems politómicos para obtener el mismo grado de confiabilidad.
- Algunos rasgos latentes son más fáciles de ser medidos en escalas graduadas.

- Para algunas variables, las respuestas a los ítems son mejor expresadas en una escala ordinal.

Dentro de la clasificación de modelos politómicos mostrada por [Bazan et al. \(2012\)](#), el Modelo Secuencial de la TRI, originalmente derivado por [Tutz \(1990\)](#), está diseñado para aplicarse en instrumentos con ítems politómicos ordinales donde cada categoría es alcanzada de manera consecutiva; es decir, una categoría puede ser alcanzada únicamente si la categoría anterior también lo fue. Esta es una diferencia fundamental respecto a otros modelos politómicos como el Modelo de Respuesta Graduada donde el evaluado debe analizar todas las categorías simultáneamente antes de escoger una, o el Modelo de Crédito Parcial donde el evaluado puede desarrollar todas las etapas del ítem y la construcción del puntaje es la suma de todas las etapas bien realizadas.

Denotemos como  $j = 1, \dots, J$  a los ítems de un test o encuesta, como  $i = 1, \dots, n$  a las personas que rinden el test o encuesta de tamaño  $J$ , y como  $k = 0, 1, \dots, m_j$  a la categoría de respuesta posible al ítem  $j$ . El número de categorías para cada ítem, como se aprecia, podrían ser diferentes y se asume que están ordenadas de forma creciente, siendo la probabilidad de que un individuo  $i$  con cierto nivel del rasgo latente  $\theta_i$  alcance la categoría  $k$  del ítem  $j$  dada por:

$$P_{ijk} = \left[ \prod_{r=0}^k \frac{\exp(a_j(\theta_i - b_{jr}))}{1 + \exp(a_j(\theta_i - b_{jr}))} \right] \left[ \frac{1}{1 + \exp(a_j(\theta_i - b_{j(k+1)}))} \right]. \quad (1.1)$$

Se asume aquí la variable aleatoria manifiesta condicional que modela la respuesta politómica del individuo  $i$  para el ítem  $j$ ,  $Y_{ij}$ , tiene distribución

$$Y_{ij|\theta_i, \xi_j} \sim \text{geométrica truncada generalizada}(P_{ij0}, P_{ij1}, \dots, P_{ijm_j}, m_j) \quad (1.2)$$

donde:

- $\theta_i$ : rasgo latente (habilidad) del individuo  $i$ .
- $a_j$ : Parámetro de discriminación del ítem  $j$ .
- $b_{jr}$ : Parámetro de dificultad en la etapa  $r$  para el ítem  $j$ .

El presente trabajo plantea la aplicación de este modelo a una prueba de rendimiento estudiantil, particularmente una prueba de escritura en la que cada estudiante debe redactar un texto, posteriormente este texto es calificado en diversos aspectos por evaluadores especialmente capacitados. Esta prueba deberá tener ítems politómicos que se ajusten a un Modelo Secuencial sobre los cuales se harán las estimaciones de los parámetros desde la perspectiva bayesiana. Según [Swaminathan \(2014\)](#) si bien las estimaciones de máxima verosimilitud, ya sean conjunta, condicional o marginal proveen estimaciones razonables de los parámetros de

los ítems y de las personas, algunas veces para modelos de 2 o 3 parámetros estos procedimientos generan estimaciones fuera de rangos aceptables. Las estimaciones bayesianas, dado que toman en cuenta información a priori y colateral, a menudo superan estos problemas. Por otro lado, el tener un tamaño de muestra pequeño es un problema para la precisión de la estimación de máxima verosimilitud. Según [Matteucci y Veldkamp \(2015\)](#) la estimación bayesiana de los modelos TRI, al incluir prioris informativas, compensa el hecho de tener un tamaño de muestra pequeño.

Es importante mencionar que las estimaciones, desde la perspectiva bayesiana para el Modelo Secuencial, no han sido implementadas antes y que tampoco se ha encontrado implementaciones del Modelo Secuencial en los software estadísticos usados comunmente en psicometría como el R o el Winsteps.

## 1.2. Objetivos

El objetivo general de la tesis es presentar el Modelo Secuencial para datos politómicos ordinales de la TRI propuesto por [Tutz \(1990\)](#). Esto incluye el estudio de las propiedades del modelo, su estimación desde el punto de vista 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 educativa. De manera específica:

- Presentar las diferentes propuestas de modelos de la TRI para atributos politómicos.
- Estudiar propiedades, e implementar la estimación del Modelo Secuencial desde la perspectiva bayesiana.
- Realizar estudios de simulación acerca del Modelo Secuencial considerando computación intensiva sobre diferentes escenarios.
- Aplicar el modelo a un conjunto de datos reales de calidad educativa.

## 1.3. Organización del Trabajo

En el Capítulo 2, presentamos los más importantes modelos de la TRI para atributos politómicos, así como la formulación general del Modelo Secuencial ( $SM$ ) y del Modelo Secuencial Logístico de dos parámetros ( $2PL - SM$ ) como un caso particular del Modelo Secuencial. En el capítulo 3 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 la implementación computacional usando R y WinBUGS. En el capítulo 4 se presenta un estudio de simulación de recuperación de parámetros para el modelo  $2PL - SM$ . En el capítulo 5 se presenta la aplicación del modelo  $2PL - SM$  en el ámbito educativo, específicamente en la base de datos de la prueba de escritura de la Evaluación Muestral 2013 ( $EM2013$ ) del Ministerio de Educación. Finalmente, en el capítulo 6, se presentan conclusiones y sugerencias obtenidas en este trabajo.

En el anexo A presentamos los códigos de los programas usados para la simulación de datos, la estimación bayesiana de los parámetros y en general todos los programas utilizados

para el estudio de simulación y la aplicación del modelo 2PL-SM. En el anexo B se presentan los resultados para el diagnóstico de convergencia de las distribuciones a posteriori de los parámetros estimados en la simulación y la aplicación.



## Capítulo 2

# Modelo Secuencial

En este capítulo se presenta una breve descripción de algunos modelos de respuesta al ítem politómicos, así como la formulación del Modelo Secuencial y del Modelo Secuencial Logístico.

### 2.1. Modelos de respuesta al ítem politómicos

[Ostini y Nering \(2006\)](#) define a los ítems politómicos como ítems categóricos con más de dos posibles categorías de respuesta. En los modelos para ítems binarios el modelamiento suele estar basado en una de las categorías ya que la información sobre la otra categoría puede ser deducida por complemento. Por otro lado en los modelos politómicos esto no es aplicable y es necesario el modelamiento de cada categoría explícitamente como se menciona en [Ostini y Nering \(2006\)](#).

Asímismo en [Ostini y Nering \(2006\)](#) se menciona que una de las ventajas del uso de ítems politómicos frente a los dicotómicos es que, en virtud de su gran número de categorías de respuesta, ellos proveen mayor información sobre una gama más amplia del rasgo latente que se desea medir.

[Mellenbergh \(1995\)](#) clasifica a los ítems politómicos en dos clases, los nominales y los ordinales. Siguiendo esta clasificación se dividirá a los modelos de respuesta politómica en las mismas dos clases.

#### 2.1.1. Modelo Politómico Nominal

El modelo politómico nominal como lo explica [Thissen y Cai \(2016\)](#) ha sido diseñado para ítems con más de dos alternativas, las cuales no tienen un orden predeterminado. Este modelo usa una distribución logística multivariada para modelar la probabilidad de selección de cada alternativa. En [Thissen y Cai \(2016\)](#) además se menciona que uno de los posibles usos para el modelo nominal puede ser el determinar el orden de las categorías de un ítem ordinal cuando se tengan dudas sobre este. Un análisis con el modelo nominal podría demostrar que el orden original de las categorías de un ítem no es el adecuado.

Un ejemplo de ítem para este tipo de modelos es el utilizado en el cuestionario al docente de la Evaluación Censal de Estudiantes 2012 en segundo de primaria del Ministerio de

Educación:

Ítem: ¿Cuál de las siguientes afirmaciones se aproxima mejor a su opinión acerca de como, las niñas y niños, adquieren la noción de número?

- a) Mediante el uso práctico de sus representaciones. Por ejemplo: monedas y billetes.
- b) Practicando muchas veces la simbolización y lectura de los números.
- c) Resolviendo situaciones que involucran acciones y relaciones entre cantidades.

En este caso no existe una respuesta correcta dado que el ítem solo busca recoger las opiniones de los docentes.

### 2.1.2. Modelo Politémico Ordinal

Ostini y Nering (2006) define a los ítems politómicos ordinales como aquellos donde las categorías de respuesta tienen un ordenamiento explícito con respecto al rasgo latente de interés. Esto quiere decir que en estos modelos se asume que las categorías están ordenadas de forma creciente. Las categorías ordenadas están definidas por límites o etapas que las separan, el número de etapas siempre será uno menos que el número de categorías.

En esta sección se presentará brevemente las 3 familias de modelos politómicos ordinales presentados en Bazan et al. (2012). En general en los modelos de la TRI politómicos la probabilidad de pasar de una categoría del ítem a otra, dado un límite o etapa entre estas dos categorías, puede ser escrita en función de un rasgo latente unidimensional  $\theta$ , esta función es llamada Función de Respuesta a las Etapas del Ítem (FREI). De esta manera cada modelo tiene una FREI distinta.

#### Modelo de Respuesta Graduada

Según Bazan et al. (2012) en esta familia de modelos se asume un orden para las etapas del ítem; es decir, si una etapa no es alcanzada, todas las etapas siguientes tampoco lo son y por otro lado si una etapa es alcanzada, todas las etapas anteriores también lo son.

En el modelo de respuesta graduada la FREI es llamada acumulativa y se define como  $P_{ijk}^+$ , que denota la probabilidad acumulada de que el evaluado  $i$  alcance la categoría  $k$  y todas las categorías posteriores a esta en el ítem  $j$ .

Estos modelos también son llamados modelos de diferencia, dado que la probabilidad de un evaluado  $i$  de alcanzar la categoría  $k$  de un ítem  $j$  está dada por  $P_{ijk} = P_{ijk}^+ - P_{ijk+1}^+$ . Por otra parte, los parámetros de dificultad sobre cada probable respuesta a un ítem deben estar perfectamente ordenados; es decir, de la forma  $b_{j1} < b_{j2} < \dots < b_{jm_j}$ , donde  $m_j$  es el máximo valor que se puede alcanzar en el ítem  $j$ .

Los ítems para este tipo de modelos son aquellos en los que el evaluado debe analizar todas las categorías simultáneamente antes de escoger una, de esta manera como menciona Bazan et al. (2012), el evaluado atribuye un juicio completo en una escala. Un ejemplo de este tipo de ítems pueden ser los que tienen como categorías de respuesta a una escala de

Likert, como se presenta a continuación:

Ítem: El profesor fomenta la participación de los alumnos en clase.

- a) Muy de acuerdo.
- b) De acuerdo.
- c) Ni de acuerdo ni en desacuerdo.
- d) En desacuerdo.
- e) Muy en desacuerdo.

### Modelo de Crédito Parcial

En esta familia de modelos se enfatiza en la probabilidad de alcanzar la categoría  $k$  en lugar de la categoría  $k - 1$ , de esta manera a diferencia del modelo de respuesta graduada el análisis se concentra en condicionar un par de categorías adyacentes dejando de lado el resto de categorías. En el modelo de Crédito Parcial la FREI es llamada de crédito parcial y se define como  $\frac{P_{ijk}}{P_{ijk-1} + P_{ijk}}$ , donde  $P_{ijk}$  denota la probabilidad de que el evaluado  $i$  alcance la categoría  $k$  en el ítem  $j$ . Como se menciona en [Bazan et al. \(2012\)](#) solo las categorías  $k$  y  $k - 1$  tienen influencia directa en la FREI y no la evaluación de todas las etapas como en el modelo de respuesta graduada.

Estos modelos también son conocidos como modelos de categorías adyacentes. En este caso, a diferencia del modelo de respuesta graduada, no es necesario que los parámetros de dificultad estén perfectamente ordenados.

Como se explica en [Bazan et al. \(2012\)](#) los ítems para este tipo de modelos son aquellos en los que el evaluado construye la solución independientemente de conocer todas las etapas del ítem. El evaluador es quien califica su respuesta de acuerdo a algún criterio definido. Cabe resaltar que el evaluado puede tentar todas las etapas pero el puntaje que consigue depende de cuantas etapas él completa correctamente. Además, a diferencia del modelo de respuesta graduada, en este tipo de ítems las etapas no son evaluadas de forma simultánea, sino de forma separada. En este caso el evaluado desarrolla todas las etapas del ítem y la construcción del puntaje es la suma de todas las etapas bien realizadas (binarias). Esto es llamado formación de puntaje paralela.

Un ejemplo de este tipo de ítem se puede encontrar en el cuadernillo modelo de la Evaluación Censal de Estudiantes del Ministerio de Educación de segundo de secundaria del 2016:

Ítem: Históricamente, el pueblo de Monte Alegre estaba poblado por familias dedicadas a la agricultura y ganadería de subsistencia, es decir, solo les alcanzaba para la alimentación y el vestido de su familia.

Esta situación cambió con la construcción de fábricas textiles en Monte Alegre. Con la llegada de las fábricas, al poco tiempo, muchas personas se asentaron en Monte Alegre y

ocuparon espacios antes despoblados.

¿Cómo se relaciona la construcción de fábricas con la ocupación de los espacios en Monte Alegre?

La calificación de este ítem se hace de acuerdo a los siguientes criterios:

- Crédito completo (2): Respuestas que señalen que la gente migra porque puede encontrar trabajo en las fábricas o respuestas que consideren que la presencia de las fábricas trae consigo oportunidades para vivir más cómodamente en el lugar.
- Crédito parcial (1): Respuestas que identifican un impacto en el entorno (ya sea tanto contaminando como simplemente ocupando espacio) pero no que la gente va a Monte Alegre por oportunidades.
- Sin crédito (0): Respuestas que mencionen los dos sucesos (construcción de fábricas y ocupación de espacios) pero no establezcan ninguna relación entre ellos además de la temporal. Respuestas sin sentido o que no contestan a la pregunta.

### Modelo Secuencial

En esta familia de modelos se enfatiza en la probabilidad de alcanzar la categoría  $k$  dado que ya se alcanzó la categoría  $k - 1$ . El Modelo Secuencial tiene como característica estar diseñado para instrumentos cuyos ítems deben ser resueltos de manera consecutiva; es decir, pasar de una etapa del ítem a otra implica que todas las anteriores han tenido que ser alcanzadas. Asimismo, el fallar en una etapa implica que las etapas siguientes ya no serán consideradas. Como lo explica [Welchowski \(2014\)](#) esta es una gran diferencia con respecto al modelo de crédito parcial donde no necesariamente todos los pasos previos deben ser completados antes de alcanzar la categoría más alta. [Masters \(2016\)](#) con respecto al modelo de crédito parcial define a la probabilidad de que un individuo alcance la categoría  $k$  como la probabilidad de que una persona obtenga  $k$  en lugar de  $k - 1$  a diferencia del Modelo Secuencial en donde la probabilidad de alcanzar una categoría  $k$  es condicional a que ya alcanzó la categoría  $k - 1$ . Para [Tutz \(2016\)](#), dado que el modelo de crédito parcial se concentra en solo un par de categorías dejando de lado el resto, no es tan apropiado para modelar procesos en los que se solucionan los problemas de forma secuencial.

Por otro lado, si bien el ordenamiento de las etapas del ítem debe ser fijo, no es necesario que los parámetros de dificultad estén perfectamente ordenados como se menciona en [Akkermans \(1998\)](#). Estos modelos también son conocidos como modelos condicionales.

En el Modelo Secuencial la FREI es llamada FREI condicional y se define como  $\frac{P_{ijk}^+}{P_{ijk-1}^+}$ . Como ya se mencionó anteriormente,  $P_{ijk}^+$  denota la probabilidad acumulada de que el evaluado  $i$  alcance la categoría  $k$  y todas las categorías posteriores a esta en el ítem  $j$ .

En los ítems de este tipo de modelos el evaluado construye su solución paso a paso y el evaluador asigna el puntaje dependiendo de si cada etapa es alcanzada por el evaluado.

Como se mencionó anteriormente, si ocurre una falla en alguna de las etapas el evaluador ya no considerará las etapas siguientes. Como menciona Akkermans (1998), es importante señalar que es intrascendente si el evaluado intenta realizar todas las etapas del ítem o si este se detiene al fallar. Es la regla para asignar el puntaje lo que hace al proceso secuencial. De esta manera, el evaluado podría fallar en la etapa  $k$  de un ítem y hacer bien la etapa  $k + 1$ , pero por la regla de asignación del puntaje, este solo obtendrá  $k - 1$ .

Un ejemplo de este tipo de ítem es el siguiente:

Ítem: Diana invierte una cantidad de dinero para la impresión de 40 revistas de arte de 25 hojas cada una, buscando obtener una ganancia del 80 % del costo total de impresión. Si se sabe que el costo de impresión de cada hoja es de S/.0.20 y que con todo el dinero recaudado quiere comprarse una impresora de S/.550, ¿cuánto dinero le falta a Diana para comprarse la impresora?

- Puntaje 1: 40 revistas  $\times$  25 hojas  $\times$  0.20 soles = 200 soles.
- Puntaje 2: 0.8  $\times$  200 + 200 = 360 soles.
- Puntaje 3: 550 - 360=190 soles le falta para comprar la impresora.

## 2.2. Formulación del Modelo Secuencial

En esta sección se presenta una formulación general para el Modelo Secuencial, sus supuestos, casos particulares, características, función de información, y la descripción del Modelo Secuencial Logístico.

### 2.2.1. Presentación y racionalidad del modelo

Consideremos un conjunto de  $i = 1, \dots, n$  sujetos que respondieron un cuestionario o prueba con  $j = 1, \dots, J$  ítems, y sea  $k = 0, 1, \dots, m_j$  las categorías de respuestas posibles al ítem  $j$ . De esta manera cada ítem cuenta con  $m_j + 1$  categorías politómicas ordinales, las cuales pueden ser distintas entre los ítems pero siempre deben estar ordenadas de tal manera que para alcanzar una categoría, necesariamente se tiene que alcanzar todas las anteriores. El Modelo Secuencial es definido a través de la distribución

$$Y_{ij} \mid \theta_i, \xi_j \sim \text{geométrica truncada generalizada} (P_{ij0}, P_{ij1}, \dots, P_{ijm_j}, m_j) \quad (2.1)$$

donde para  $i = 1, \dots, n$ ;  $j = 1, \dots, J$ ;  $k = 0, 1, \dots, m_j$ :

$$P_{ijk} = P(Y_{ij} = k \mid \theta_i, \xi_j) = \left( \prod_{r=0}^k \frac{P_{ijr}^+}{P_{ijr-1}^+} \right) \left( 1 - \frac{P_{ijk+1}^+}{P_{ijk}^+} \right), \quad (2.2)$$

$$\frac{P_{ijk}^+}{P_{ijk-1}^+} = F(\eta_{ijk}) = F(a_j(\theta_i - b_{jk})). \quad (2.3)$$

Aquí consideramos por conveniencia de notación que  $\frac{P_{ij0}^+}{P_{ij0-1}^+} \equiv 1$  y  $\frac{P_{ijm_j+1}^+}{P_{ijm_j}^+} \equiv 0$ .

Donde:

- $Y_{ij}$  es la variable aleatoria manifiesta condicional que modela la respuesta política del individuo  $i$  alcanzando la categoría  $k$  del ítem  $j$ , donde  $k \in \{0, 1, \dots, m_j\}$ .
- $P_{ijk}$ : es la probabilidad condicional de que el individuo  $i$  alcance la categoría  $k$  en el ítem  $j$  y se satisface que  $\sum_{r=0}^{m_j} P_{ijr} = 1$ .
- $P_{ijk}^+ = P(Y_{ij} \geq k \mid \theta_i, \xi_j)$  expresa la probabilidad de que el evaluado  $i$  responda correctamente la etapa  $k$  y todas las etapas superiores del ítem  $j$  dada su habilidad latente  $\theta_i$  y las características del ítem.
- $\eta_{ijk}$  es el predictor latente lineal relacionado con  $\theta_i$ .
- $F$  es una función de distribución acumulada que relaciona el predictor latente  $\eta_{ijk}$  con las probabilidades condicionales.
- $\theta_i$  es el valor de la variable latente  $\Theta_i$  para el individuo  $i$  tal que  $\theta_i \in \mathbb{R}$ .
- $\xi_j = (a_j, b_{j1}, \dots, b_{jm_j})$  es el vector de parámetros del ítem  $j$ . El parámetro de discriminación  $a_j > 0$  está relacionado con la inclinación del ítem y los  $b_{jm_j}$  son parámetros de dificultad que representan la posición de las etapas (también llamados umbrales) entre dos categorías del ítem. Nótese que los parámetros están en función de las etapas y por lo tanto la notación para los subíndices de los parámetros de dificultad empieza en 1.

### 2.2.2. Observaciones

- Si  $X$  es el número de éxitos hasta la ocurrencia de la primera falla en una secuencia de  $m$  ensayos de Bernoulli independientes pero no idénticamente distribuidos con probabilidad de éxito en el ensayo  $r$  de  $P_r$ ,  $X$  sigue una distribución geométrica truncada generalizada con función de probabilidad:

$$P(X = x) = \left( \prod_{r=0}^x P_r \right) (1 - P_{x+1}), \quad x = 0, 1, \dots, m \quad (2.4)$$

donde  $P_r$  es la probabilidad de éxito en el  $r$ -ésimo ensayo de Bernoulli y consideramos por conveniencia la notación  $P_0 \equiv 1$  y  $P_{m+1} \equiv 0$ . En este caso se denota que  $X \sim$  geométrica truncada generalizada( $P_1, \dots, P_m, m$ ).

A esta distribución se le denomina truncada ya que a diferencia de la distribución geométrica, donde  $X$  puede ir hasta el infinito, en este caso  $X$  solo puede alcanzar como máximo el valor de  $m$  que es la cantidad de ensayos de Bernoulli. Además, otra diferencia con la distribución geométrica es que la probabilidad de éxito puede ser distinta en cada ensayo, y por esto es llamada generalizada.

Tal como detalla [Akkermans \(1998\)](#), en el Modelo Secuencial tenemos que el número de aciertos antes de la ocurrencia del primer fallo, en una secuencia de ensayos de Bernoulli, sigue una distribución geométrica truncada en la categoría máxima del ítem ( $m_j$ ). De esta manera si  $Y_{ij}$  se define como el número de éxitos hasta que ocurra el primer fallo en una secuencia de  $m_j$  ensayos Bernoulli, esta sigue una distribución geométrica truncada con parámetros  $\pi_{ijk} = \frac{P_{ijk}^+}{P_{ijk-1}^+}$  y  $m_j$ .

Por otro lado, cada parámetro  $\pi_{ijk}$  varía en cada categoría del ítem. En ese sentido cada  $\pi_{ijk}$  se ha definido como una función de la variable  $\theta_i$ . Además cada  $\pi_{ijk}$  es también llamada Curva Características de los Ítems (CCIs) por [Akkermans \(1998\)](#), sin embargo dado que estamos en un modelo politómico, en el presente trabajo se llamará Función de Respuesta a las Etapas del Ítem (FREI).

- Para [Tutz \(2016\)](#) las respuestas ordinales permiten medir la habilidad de forma más precisa de lo que sería posible con ítems dicotómicos.
- [Tutz \(2016\)](#) llama a los parámetros  $b_{jk}$  como la dificultad de los ítems para el paso  $k$ . De forma general, la dificultad del ítem representa la dificultad para la transición de la categoría  $k-1$  a la categoría  $k$ . Además, a diferencia del modelo de respuesta graduada donde cada parámetro  $b_{jk}$  debe ser menor que el siguiente ( $-\infty = b_{j0} < b_{j1} < \dots < b_{jm_{j-1}} < b_{jm_j} = +\infty$ ), en el Modelo Secuencial esto no es necesario.
- Como se detalla en [Tutz \(2016\)](#), el proceso siempre empieza en cero. Si la transición a la categoría 1 falla, el proceso se detiene y el puntaje de la persona  $i$  en el ítem  $j$  es  $Y_{ij} = 0$ . Si la transición es satisfactoria la persona alcanza por lo menos la categoría 1 ( $Y_{ij} = 1$ ) y podrá enfrentarse a la siguiente categoría, de esta manera tenemos 2 opciones:
  - $Y_{ij} = 1$  dado que  $Y_{ij} \geq 1$  si la persona no alcanza la categoría 2;
  - $Y_{ij} = 2$  dado que  $Y_{ij} \geq 1$  si la persona alcanza la categoría 2;

De esta forma la asunción básica es que cada ítem es resuelto paso a paso.

- Alternativamente el modelo también puede ser presentado de la siguiente manera:

$$P_{ijk} = \prod_{r=0}^k \frac{P_{ijr}^+}{P_{ijr-1}^+} - \prod_{r=0}^{k+1} \frac{P_{ijr}^+}{P_{ijr-1}^+}, \quad k = 0, 1, \dots, m_j \quad (2.5)$$

Considerando por conveniencia de notación que  $\prod_{r=0}^0 \frac{P_{ijr}^+}{P_{ijr-1}^+} \equiv 1$  y  $\prod_{r=0}^{m_j+1} \frac{P_{ijr}^+}{P_{ijr-1}^+} \equiv 0$ .

- Otra forma de presentar el modelo, tomando como referencia la presentada en [Michael y Ostini \(2010\)](#) es la siguiente:

$$P_{ijk} = \begin{cases} 1 - \frac{P_{ijk+1}^+}{P_{ijk}^+}, & \text{si } k = 0 \\ \left(\prod_{r=1}^k \frac{P_{ijr}^+}{P_{ijr-1}^+}\right) \left(1 - \frac{P_{ijk+1}^+}{P_{ijk}^+}\right), & \text{si } 0 < k < m_j \\ \prod_{r=1}^k \frac{P_{ijr}^+}{P_{ijr-1}^+}, & \text{si } k = m_j \end{cases} \quad (2.6)$$

### 2.2.3. Supuestos del Modelo Secuencial

- Según [Tutz \(2016\)](#) el modelo es unidimensional. Hay una única escala latente para medir la habilidad y la dificultad de los ítems. Por otro lado como menciona [Attorresi et al. \(2009\)](#) en la práctica es difícil que este supuesto de unidimensionalidad del rasgo latente se satisfaga totalmente ya que múltiples factores pueden afectar a la respuesta a un ítem. En los modelos unidimensionales sólo se exige un rasgo fundamental (factor dominante) que explique las respuestas de los sujetos.
- Un supuesto general para los modelos TRI es el de la independencia entre las respuestas de distintos individuos. De esta forma si definimos a  $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{ij}, \dots, Y_{iJ})$  como el conjunto de respuestas de un individuo  $i$  a los  $J$  ítems del test, la probabilidad del conjunto global de respuestas de los  $n$  individuos estará dado por

$$P(Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n \mid \theta_1, \dots, \theta_n, \xi_1, \dots, \xi_n) = \prod_{i=1}^n P(\mathbf{Y}_i = \mathbf{y}_i \mid \theta_i, \xi_i) \quad (2.7)$$

- Otro supuesto general para los modelos TRI es el de la independencia local de los ítems. Según [Hidalgo \(2016\)](#) este supuesto asume que las respuestas de un individuo para cualquier pareja de ítems en el test no están relacionadas cuando la habilidad se mantiene constante. En [Attorresi et al. \(2009\)](#) se menciona que este supuesto iría de la mano con el de unidimensionalidad ya que si dos ítems no son independientes significa que otro factor, ajeno al que se pretende medir incide en la estimación del rasgo latente violando el supuesto de unidimensionalidad.

De esta manera, la probabilidad del patrón de respuesta  $\mathbf{y}_i = (y_{i1}, \dots, y_{iJ})$  para los  $J$  ítems estará definida por

$$P(\mathbf{Y}_i = \mathbf{y}_i \mid \theta_i, \xi_j) = \prod_{j=1}^J P(Y_{ij} = y_{ij} \mid \theta_i, \xi_j) \quad (2.8)$$

- En [Tutz \(2016\)](#) se define que las respuestas multicatóricas se pueden codificar en variables dummy de la siguiente forma:

$$W_{ijr} = \begin{cases} 1, & Y_{ij} = r \\ 0, & \text{otros casos} \end{cases} \quad (2.9)$$

donde el vector  $\mathbf{W}_{ij}^T = (W_{ij0}, W_{ij1}, \dots, W_{ijm_j})$  es un vector con solo un "1", el cual indica la categoría alcanzada por la persona  $i$ .

De esta manera la log verosimilitud de las respuestas de una persona  $i$  a un ítem  $j$  tiene la siguiente forma:

$$l_{ij} = \log(L_{ij}) = \sum_{r=0}^{m_j} w_{ijr} \log(P_{ijr}) \quad (2.10)$$

donde  $L_{ij}$  es la verosimilitud de la respuesta de la persona  $i$  a un ítem  $j$  la cual puede definirse como:

$$L_{ij} = P(Y_{ij} = y_{ij} \mid \theta_i, \xi_j) = P(W_{ij} = w_{ij} \mid \theta_i, \xi_j) = \prod_{r=0}^{m_j} [P_{ijr}]^{w_{ijr}}. \quad (2.11)$$

- **Akkermans (1998)** detalla la propiedad de invarianza de la CCI (aquí llamada FREI) como esencial en un proceso de calificación secuencial. Esta propiedad dice que para cualquiera, sean dos procesos de calificación secuencial A y B, donde las etapas del proceso B son idénticas a las del A, excepto por una etapa  $k$  la cual no está en el proceso B, se debe cumplir que:

$$P(Y_{jk+1} = 1 \mid Y_{jk} = 1; \theta_i, \xi_j) = P(Y_{jk+1} = 1 \mid Y_{jk-1} = 1; \theta_i, \xi_j)$$

La parte izquierda de la ecuación es el proceso de calificación A y la parte derecha el proceso B. Cualquier modelo de calificación secuencial necesita poder cumplir con este supuesto. **Akkermans (1998)** explica que el Modelo de Respuesta Graduada solo cumple con este supuesto si es que la etapa eliminada es la última y por el contrario, el Modelo de Crédito Parcial solo cumple con este supuesto si la etapa eliminada es la primera.

#### 2.2.4. Características del Modelo Secuencial

En **Ostini y Nering (2006)** se explica que comparados con los modelos dicotómicos, una característica de los modelos politómicos es que la distinción entre categorías de respuesta y etapas reflejan dos diferentes tipos de probabilidades condicionales: (a) La probabilidad de alcanzar una categoría y (b) la probabilidad de responder positivamente en lugar de negativamente en una etapa dada entre dos categorías. La primera probabilidad se encuentra definida por la Función de Respuesta a las Categorías del Ítem (FRCI), mientras que la segunda se encuentra definida por la Función de Respuesta a las Etapas del Ítem (FREI). En el caso dicotómico ambas probabilidades son lo mismo y dicha probabilidad se encuentra definida por la Función de Respuesta al Ítem (FRI).

En esta sección se presentará la definición de la FREI y de la FRCI, además de la Función de Información que también es llamada Información de Fisher.

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

La FREI para un ítem  $j$  en el Modelo Secuencial se encuentra definida por:

$$\frac{P_{ijk}^+}{P_{ijk-1}^+} = F(\eta_{ijk}) \quad (2.12)$$

donde

$$\eta_{ijk} = a_j(\theta_i - b_{jk}) \quad (2.13)$$

con  $k = 0, 1, \dots, m_j$

Aquí

- $P_{ijk}^+$  es la probabilidad condicional acumulada de que un individuo con habilidad  $\theta_i$  logre superar alcanzar la categoría  $k$  y todas las posteriores a esta en el ítem  $j$ .
- $F$  es una función de distribución que relaciona el predictor latente  $\eta_{ijk}$  con las probabilidades condicionales, donde:

$F : \mathbb{R} \rightarrow [0, 1]$  es continua por la derecha y monótona creciente con

$$\lim_{x \rightarrow -\infty} F(x) = 0$$

$$\lim_{x \rightarrow +\infty} F(x) = 1$$

$F$  también es llamada función de enlace y podría ser la función acumulada de cualquier distribución dado que cumpliría con los requisitos presentados anteriormente.

De acuerdo a lo mencionado en [Bazan et al. \(2012\)](#) estas funciones corresponden a las del tipo condicional.

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

La FRCI para un ítem  $j$  en el Modelo Secuencial fue definida en (2.2), siendo expresada esta en términos de la FREI como:

$$P_{ijk} = P(Y_{ij} = k \mid \theta_i, \xi_j) = \prod_{r=1}^k F(\eta_{ijr})(1 - F(\eta_{ijk+1})) \quad (2.14)$$

con  $k = 0, 1, \dots, m_j$ .

### Información de Fisher en el Modelo Secuencial

De acuerdo a [Hidalgo \(2016\)](#) la Función de Información representa la cantidad de información que proporciona un ítem en cada nivel del continuo de habilidad (rasgo latente). Esta función de información en la TRI está relacionada con la fiabilidad de la medida. Cuanta más información tiene un ítem o un test, más precisa es la estimación de la habilidad del sujeto.

En [Reckase \(2009\)](#) se menciona que el estimador de máxima verosimilitud de un parámetro, en este caso la habilidad  $\Theta$ , se distribuye asintóticamente con una media igual a  $\theta$  y una varianza igual a:

$$\sigma^2(\hat{\Theta} | \Theta) = \frac{1}{E[(\frac{\partial \ln L}{\partial \theta})^2 | \Theta]} \quad (2.15)$$

donde  $L$  es la verosimilitud de los datos dado el parámetro, en este caso, la verosimilitud de las respuestas de los ítems dado  $\Theta$ .

En [Reckase \(2009\)](#) también se menciona que el término en el denominador en (2.15) es igual a la información del test cuando los ítems han sido ajustados a un modelo dicotómico; sin embargo, los conceptos de información acerca de  $\theta$  han sido generalizados a los test compuestos por ítems politómicos. Por otro lado, el recíproco de la varianza del estimador de máxima verosimilitud fue denominado como información por Fisher en 1925, de aquí es que la formulación general de información presentada en [Reckase \(2009\)](#) es llamada Información de Fisher.

Consideremos un sujeto con habilidad  $\theta$ , y sea  $Y_{j|\theta, \xi_j} \sim \text{geométrica truncada generalizada}$  ( $P_{j0}, P_{j1}, \dots, P_{jm_j}, m_j$ ) la respuesta del sujeto a un ítem  $j$ , con  $k$  categorías donde  $k = 0, 1, \dots, m_j$  y  $f(y, \theta)$  la función de verosimilitud de la variable latente  $\theta$ , la cual dados los supuestos presentados en (2.8) y (2.11) puede ser expresada como:

$$f(y; \theta) = \prod_{j=1}^J \prod_{r=0}^{m_j} [P_{jr}]^{w_{jr}} \quad (2.16)$$

donde  $P_{jr}$  es la probabilidad de que el evaluado responda a la categoría  $r$  del ítem  $j$ , y  $w_{jr}$  la parametrización alternativa para  $y_{jk}$  planteada en (2.9).

El logaritmo de la verosimilitud estará definido por:

$$\ln f(y; \theta) = \sum_{j=1}^J \sum_{r=0}^{m_j} w_{jr} \ln P_{jr}$$

Tomando la primera derivada con respecto a  $\theta$

$$\ln f(y; \theta)' = \sum_{j=1}^J \sum_{r=0}^{m_j} \frac{w_{jr} P_{jr}'}{P_{jr}} \quad (2.17)$$

Si partimos de (2.15) la Información de Fisher queda definida como:

$$I(\Theta) = E \left[ \left( \frac{\partial \ln f(y; \theta)}{\partial \theta} \right)^2 \mid \Theta \right] \quad (2.18)$$

Alternativamente (2.18) también puede ser expresado como:

$$I(\Theta) = -E \left[ \left( \frac{\partial^2 \ln f(y; \theta)}{\partial^2 \theta} \right) \mid \Theta \right] \quad (2.19)$$

Reemplazando 2.17 en 2.18

$$I(\theta) = E \left[ \left( \sum_{j=1}^J \sum_{r=0}^{m_j} \frac{W_{jr} P'_{jr}}{P_{jr}} \right)^2 \mid \theta \right] = \sum_{j=1}^J \sum_{r=0}^{m_j} \frac{[P'_{jr}]^2}{P_{jr}^2} E[W_{jr}^2]$$

Dado que  $W_{jr} \sim \text{Bernoulli}(P_{jr})$ ;  $E[W_{jr}] = P_{jr}$  y  $\text{Var}[W_{jr}] = E[W_{jr}^2] - E[W_{jr}]^2$ .

$$I(\theta) = \sum_{j=1}^J \sum_{r=0}^{m_j} \frac{[P'_{jr}]^2}{P_{jr}^2} E[P_{jr}(1 - P_{jr}) + P_{jr}^2] = \sum_{j=1}^J \sum_{r=0}^{m_j} \frac{[P'_{jr}]^2}{P_{jr}}$$

Partiendo de la ecuación anterior, la Función de Información de un Ítem (FII) queda definida como:

$$I_j(\theta) = \sum_{r=0}^{m_j} \frac{[P'_{jr}]^2}{P_{jr}}$$

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

$$I(\theta) = \sum_{j=1}^J I_j(\theta)$$

En Hambleton (1985) se presentan las siguientes características para la Función de Información del Test:

- La cantidad de información es influenciada por la calidad y número de ítems en el test. Mientras más empinada la pendiente, mayor cantidad de información. Mientras más pequeña la varianza del ítem, mayor información.
- $I(\theta)$  no depende de una combinación particular de ítems. La contribución de cada ítem es independiente de los otros ítems en el test.
- La cantidad de información provista por un conjunto de ítems en un nivel de habilidad está inversamente relacionado al error asociado con el estimado de la habilidad en dicho

nivel de habilidad.

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (2.20)$$

donde  $SE(\theta)$  es el error estándar del estimado de la habilidad en el nivel de habilidad  $\theta$ .

### 2.2.5. Casos particulares del Modelo Secuencial

#### Modelo de Respuesta Graduada

Tutz (2016) detalla que solo para el caso en el que para el Modelo Secuencial se utilice la función de distribución  $F$ , definida en (2.12), como de valor extremo con densidad  $f(x) = \exp(-\exp(-x))$  entonces el modelo respuesta graduada y el secuencial serán equivalentes.

#### Modelo Rasch Secuencial

En Tutz (1990) se menciona que si  $F$  es la distribución logística (considerando al parámetro de discriminación como 1 para todos los casos), el paso de la categoría  $r$  a la categoría  $r + 1$  dado que la categoría  $r$  ya fue alcanzada, es modelado por el modelo Rasch dicotómico. De la misma forma, el parámetro  $b_{j,r}$  será la dificultad de alcanzar la etapa  $r + 1$  del ítem  $j$  si la etapa  $r$  ya fue alcanzada. Este modelo es llamado Modelo Rasch Secuencial (en inglés Sequential Rasch Model o Stepwise Rasch Model).

#### Modelo Secuencial Multidimensional

En Deng y Bolt (2016) se discute una extensión multidimensional para el Modelo Secuencial. En este artículo se plantea como instrumento una prueba de corrección de oraciones donde al evaluado se le presenta una oración la cual podría tener un error gramatical. Cada ítem tiene 5 alternativas donde la primera opción es una réplica de la oración, esto indicaría que el evaluado no detecta el error gramatical. Las 4 alternativas restantes son opciones para corregir el error gramatical. Como primer paso para resolver uno de estos ítems el evaluado tendrá que ser capaz de detectar el error y con esto obtener un score de 1 como mínimo. Una vez hecho esto, el evaluado deberá escoger entre las 4 opciones restantes la alternativa correcta, en cuyo caso obtendrá un score de 2.

De esta manera, el proceso secuencial viene dado por dos etapas asociadas cada una a un rasgo latente distinto. La primera etapa asociada al rasgo latente que permite la detección del error y denominada como  $\eta$ ; y la segunda etapa asociada al rasgo latente que permite la corrección del error, denominada  $\theta$ .

La primera etapa es modelada por un modelo logístico de dos parámetros mientras que la segunda etapa es modelada por un modelo de respuesta nominal. La probabilidad general de que el evaluado  $i$  seleccione la categoría  $k$  del ítem  $j$  viene modelada por una combinación de las dos etapas anteriores.

### 2.3. Modelo Secuencial Logístico

En esta sección se definen y presentan las propiedades del Modelo Secuencial Logístico (2PL – SM).

#### 2.3.1. Definición del Modelo

Dada una matriz de datos  $Y = [Y_{ij}]_{n \times J}$  que representa las respuestas ordinales de los  $n$  individuos evaluados en un instrumento de  $J$  ítems con  $Y_{ij}$  definida como en (2.1). De esta manera el Modelo Secuencial Logístico de dos Parámetros (2PL-SM) se puede considerar como un caso particular del Modelo Secuencial (SM) en el cual  $F$  es la función de distribución acumulada de la distribución logística estándar. Considerando la ecuación (2.2) el modelo 2PL-SM queda definido como:

$$Y_{ij} \mid \theta_i, \xi_j \sim \text{geométrica truncada generalizada} (P_{ij0}, P_{ij1}, \dots, P_{ijm_j}, m_j) \quad (2.21)$$

donde para  $i = 1, \dots, n$ ;  $j = 1, \dots, J$ ;  $k = 0, 1, \dots, m_j$ :

$$P_{ijk} = P(Y_{ij} = k \mid \theta_i, \xi_j) = \left( \prod_{r=0}^k \frac{P_{ijr}^+}{P_{ijr-1}^+} \right) \left( 1 - \frac{P_{ijk+1}^+}{P_{ijk}^+} \right), \quad (2.22)$$

y

$$\frac{P_{ijk}^+}{P_{ijk-1}^+} = \Psi(\eta_{ijk}) = \Psi(a_j(\theta_i - b_{jk})). \quad (2.23)$$

Aquí consideramos por conveniencia de notación que  $\frac{P_{ij0}^+}{P_{ij0-1}^+} \equiv 1$  y  $\frac{P_{ijm_j+1}^+}{P_{ijm_j}^+} \equiv 0$ . Note que  $P_{ijk}$  representa la probabilidad condicional de responder a la categoría  $k$  en el ítem  $j$ ,  $P_{ijk}^+ = P(Y_{ij} \geq k \mid \theta_i, \xi_j)$  la probabilidad condicional de que el evaluado  $i$  alcance la categoría  $k$  y todas las superiores del ítem  $j$  dada su habilidad latente  $\theta_i$  y las características del ítem,  $\eta_{ijk} = a_j(\theta_i - b_{jk})$  es el predictor latente lineal relacionado con  $\theta_i \in \mathbb{R}$ ,  $\xi_j = (a_j, b_{j1}, \dots, b_{jm_j})$  es el vector de parámetros del ítem  $j$  y  $\Psi$  la función de distribución acumulada de la distribución logística estándar:

$$\Psi(\eta_{ijk}) = \frac{\exp(\eta_{ijk})}{1 + \exp(\eta_{ijk})} = \frac{1}{1 + \exp(-\eta_{ijk})}, \quad -\infty < \eta_{ijk} < \infty.$$

#### 2.3.2. Características del Modelo 2PL-SM

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

Las FREI para el modelo 2PL – SM estarán definidas como:

$$\frac{P_{ijk}^+}{P_{ijk-1}^+} = \Psi(\eta_{ijk}) = \frac{1}{1 + \exp(-\eta_{ijk})}, \quad (2.24)$$

donde

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

$$i = 1, \dots, n; \quad j = 1, \dots, J; \quad k = 0, 1, \dots, m_j .$$

En la figura 2.1 se muestra las FREI para 3 ítems de 5 categorías con umbrales de  $b = (-2.0, 0.1, 0.8, 1.4)$  para los dos primeros y con  $b = (-2.0, 0.1, -0.5, 1.4)$  para el tercero. Además el primer ítem tiene un parámetro de discriminación de 1 y los dos últimos ítems tienen un parámetro de discriminación de 2.

Si comparamos los dos primeros ítems donde la diferencia radica en el parámetro de discriminación podemos observar que a mayor valor de este la curva es más empinada. De esta manera, un mayor valor de discriminación permitirá diferenciar mejor a los evaluados.

Si comparamos los ítems 2 y 3 se puede observar que la diferencia se encuentra en el tercer umbral. En el tercer ítem el valor del tercer umbral es menor que el del segundo umbral. El Modelo Secuencial no necesita que los parámetros de dificultad estén ordenados de acuerdo a las categorías ordinales del ítem.

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

En el modelo  $2PL - SM$  las FRCI estarán definidas como:

$$P_{ijk} = \prod_{r=0}^k \left[ \frac{1}{1 + \exp(-\eta_{ijr})} \right] \left[ \frac{\exp(-\eta_{ijk+1})}{1 + \exp(-\eta_{ijk+1})} \right], \quad (2.25)$$

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

donde

$$i = 1, \dots, n; \quad j = 1, \dots, J; \quad k = 0, 1, \dots, m_j .$$

Así, como en el caso de las FREI, la forma que toman las FRCI también depende de los parámetros  $a_j$  y  $b_{jk}$ . En la figura 2.2 se puede observar el comportamiento de de la FRCI de 6 ítems con 4 categorías de respuesta. Los ítems de la derecha (ítems 2, 4 y 6) presentan parámetros de discriminación altos comparados con los de la izquierda (ítems 1, 3 y 5) y es por esto que las curvas de la derecha son mucho más empinadas. Por otro lado, si comparamos los parámetros de dificultad, los ítems 1 y 2 presentan los índices de dificultad más bajos,

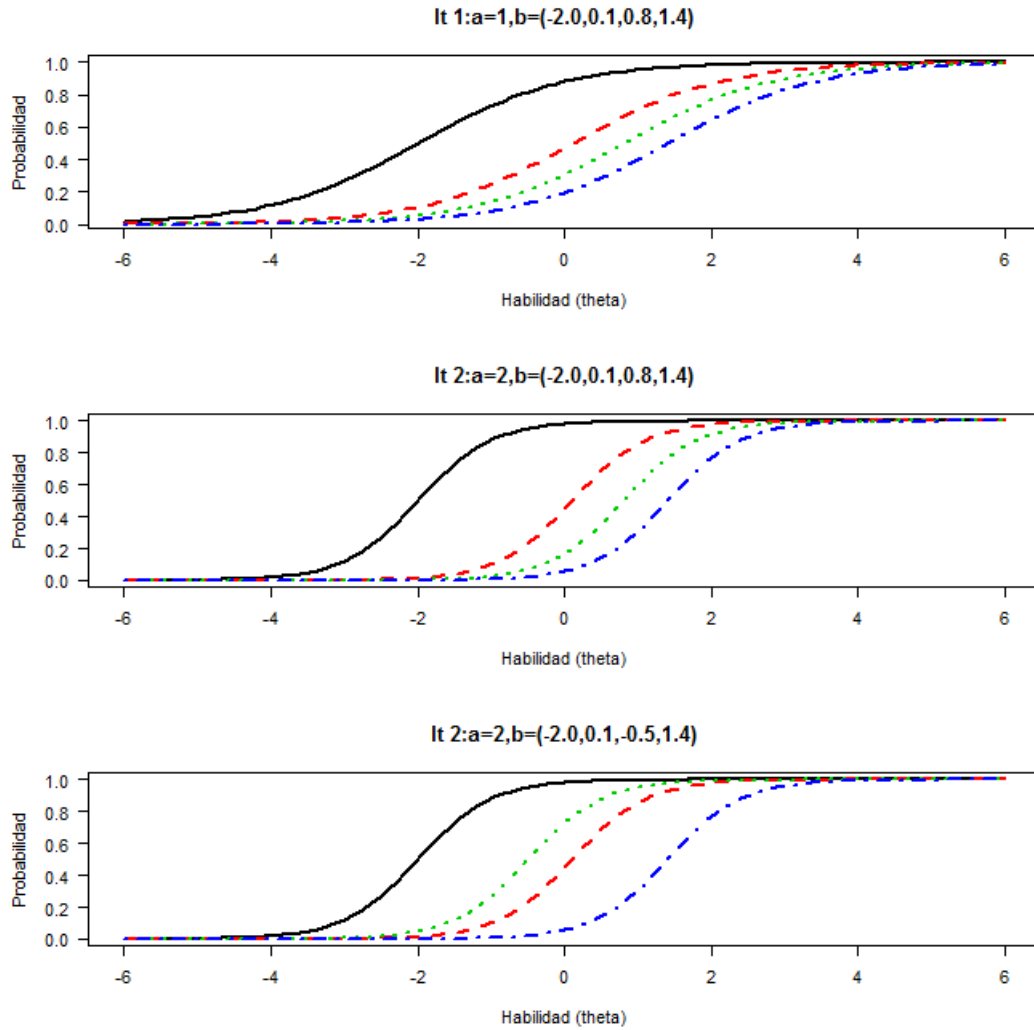


Figura 2.1: Representación gráfica de las FREI del Modelo 2PL-SM para tres ítems de 4 categorías con parámetros de discriminación  $a_1 = 1$  y  $a_2 = a_3 = 2$ , y con parámetros de dificultad  $b_{1r} = b_{2r} = -2.0, 0.1, 0.8, 1.4$  y  $b_{3r} = -2.0, 0.1, -0.5, 1.4$

los ítems 3 y 4 presentan índices de dificultad que en este caso podrían ser intermedios y por último los ítems 5 y 6 presentan los parámetros de dificultad más altos.

### Información de Fisher

Como se vio en la sección 2.2.4, la Función de Información de un Ítem (FII) para el Modelo Secuencial está dada por:

$$I_j(\theta) = \sum_{r=0}^{m_j} \frac{[P'_{jr}(\theta)]^2}{P_{jr}}, \quad (2.26)$$

donde:

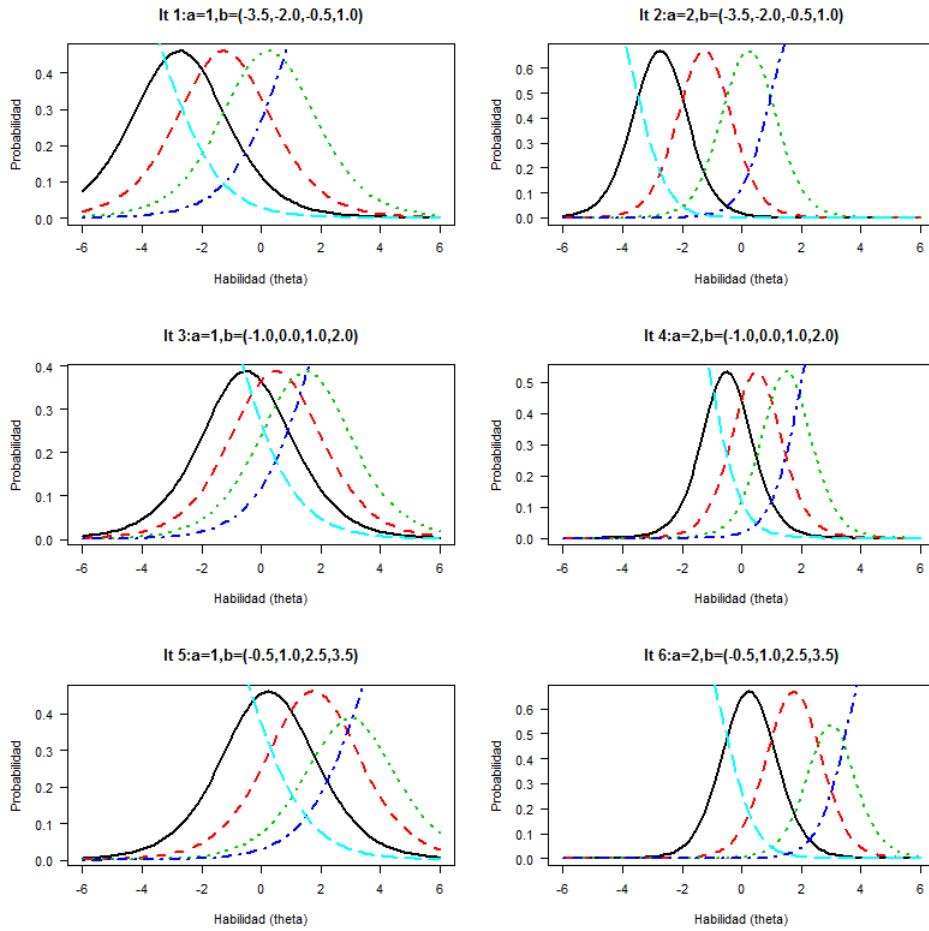


Figura 2.2: Representación gráfica de las FRCI del Modelo 2PL-SM para seis ítems de cuatro categorías con parámetros de discriminación  $a_1 = a_3 = a_5 = 1$  y  $a_2 = a_4 = a_6 = 2$ ; y con parámetros de dificultad  $b_{1r} = b_{2r} = -3.5, -2.0, -0.5, 1.0, 1$ ,  $b_{3r} = b_{4r} = -1.0, 0.0, 1.0, 2.0, 1$  y  $b_{5r} = b_{6r} = -0.5, 1.0, 2.5, 3.5, 1$

$$P_{jr} = \prod_{q=0}^r F(\eta_{ijq})(1 - F(\eta_{ijr+1})) \quad (2.27)$$

Considerando por conveniencia de notación que  $F(\eta_{ij0}) \equiv 1$  y  $F(\eta_{ijm_j+1}) \equiv 0$

Dado que

$$P'_{jr} = (\ln P_{jr})' P_{jr} \quad (2.28)$$

y

$$\begin{aligned}
(\ln P_{jr})' &= \sum_{r=0}^{m_j} (\ln F(\eta_{ijr}))' + (\ln(1 - F(\eta_{ijr+1})))' = \sum_{r=0}^{m_j} \frac{F'(\eta_{ijr})}{F(\eta_{ijr})} + \frac{(1 - F(\eta_{ijr+1}))'}{(1 - F(\eta_{ijr+1}))} \\
&= \sum_{r=0}^{m_j} \frac{F'(\eta_{ijr})}{F(\eta_{ijr})} - \frac{(F'(\eta_{ijr+1}))}{(1 - F(\eta_{ijr+1}))} = \sum_{r=0}^{m_j} \frac{\frac{a_j \exp(-a_j(\theta - b_{jr}))}{[1 + \exp(-a_j(\theta - b_{jr}))]^2}}{\frac{1}{1 + \exp(-a_j(\theta - b_{jr}))}} - \frac{\frac{a_j \exp(-a_j(\theta - b_{jr+1}))}{[1 + \exp(-a_j(\theta - b_{jr+1}))]^2}}{\frac{\exp(-a_j(\theta - b_{jr+1}))}{1 + \exp(-a_j(\theta - b_{jr+1}))}} \\
&= \sum_{r=0}^{m_j} \frac{a_j \exp(-a_j(\theta - b_{jr}))}{1 + \exp(-a_j(\theta - b_{jr}))} - \frac{a_j}{1 + \exp(-a_j(\theta - b_{jr+1}))} \\
&= \sum_{r=0}^{m_j} a_j(1 - F(\eta_{ijr})) - a_j F(\eta_{ijr+1}) \tag{2.29}
\end{aligned}$$

Se tiene que reemplazando (2.28) y (2.29) en (2.26) la FII para el modelo  $2PL - SM$  estará dada por:

$$I_j(\theta) = \sum_{r=0}^{m_j} \frac{[(\ln P_{jr})' P_{jr}]^2}{P_{jr}} = \sum_{r=0}^{m_j} [(\ln P_{jr})']^2 P_{jr} \tag{2.30}$$

$$I_j(\theta) = \sum_{r=0}^{m_j} \left[ \left[ \sum_{r=0}^{m_j} a_j(1 - F(\eta_{ijr})) - a_j F(\eta_{ijr+1}) \right]^2 \prod_{r=0}^{m_j} F(\eta_{ijr})(1 - F(\eta_{ijr+1})) \right] . \tag{2.31}$$

Así, la Función de Información del Test estará definida por:

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

Así mismo, la función para el error estándar estará dada por la siguiente ecuación:

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

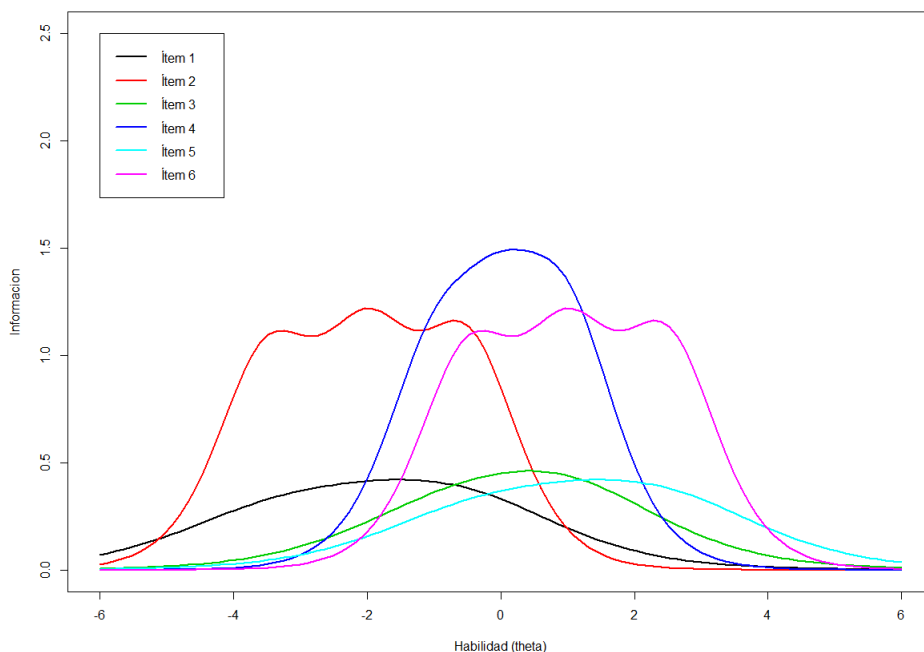


Figura 2.3: Representación gráfica de las FII del Modelo 2PL-SM para seis ítems de cuatro categorías con parámetros de discriminación  $a_1 = a_3 = a_5 = 1$  y  $a_2 = a_4 = a_6 = 2$ ; y con parámetros de dificultad  $b_{1r} = b_{2r} = -3.5, -2.0, -0.5, 1.0, 1$ ,  $b_{3r} = b_{4r} = -1.0, 0.0, 1.0, 2.0, 1$  y  $b_{5r} = b_{6r} = -0.5, 1.0, 2.5, 3.5, 1$

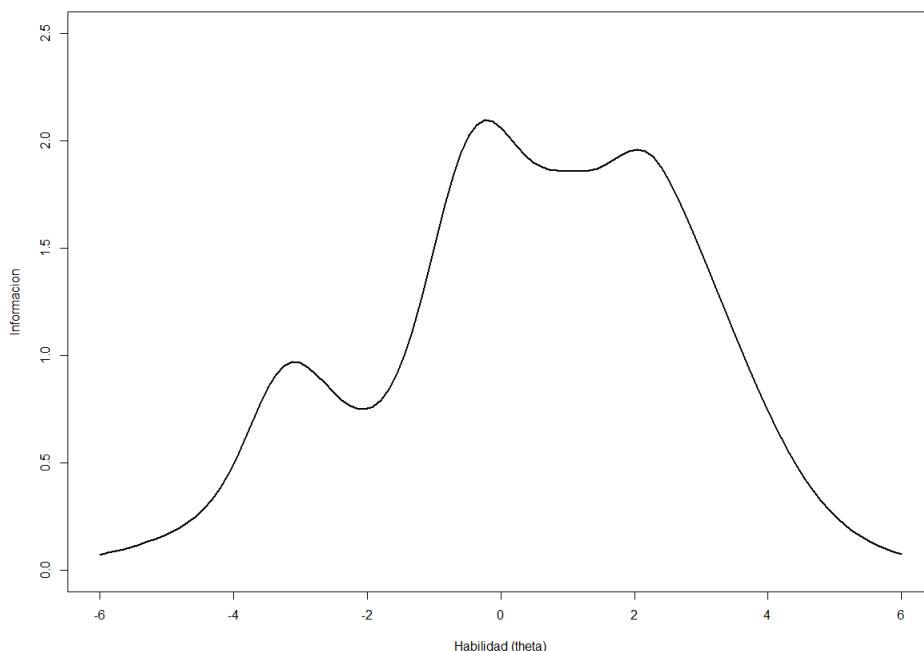


Figura 2.4: Representación gráfica de la FIT del Modelo 2PL-SM para un test de seis ítems, de cuatro categorías, con parámetros de discriminación  $a_1 = a_3 = a_5 = 1$  y  $a_2 = a_4 = a_6 = 2$ ; y con parámetros de dificultad  $b_{1r} = b_{2r} = -3.5, -2.0, -0.5, 1.0, 1$ ,  $b_{3r} = b_{4r} = -1.0, 0.0, 1.0, 2.0, 1$  y  $b_{5r} = b_{6r} = -0.5, 1.0, 2.5, 3.5, 1$

## Capítulo 3

# Estimación Bayesiana del Modelo Secuencial Logístico

En [Tutz \(2016\)](#) se explican dos métodos de estimación clásica para la estimación de los parámetros del Modelo Secuencial, estos métodos son el de máxima verosimilitud conjunta (JML) y el de máxima verosimilitud marginal (MML). En este capítulo se desarrollará la inferencia desde la perspectiva bayesiana usando los métodos MCMC. Para su implementación se utilizarán los programas R y WinBUGS.

### 3.1. Función de Verosimilitud para el Modelo Secuencial Logístico (2PL-SM)

Dada una matriz de datos  $Y = [Y_{ij}]_{n \times J}$  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),  $W_{ijr}$  la reparametrización alternativa de estas variables dada en (2.9),  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)^T$  el vector de parámetros de habilidades para las personas y  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_J)^T$  el vector de parámetros de los ítems, con  $\boldsymbol{\xi}_j = (a_j, b_{j1}, \dots, b_{jm_j})^T$ . Asumiendo los supuestos (2.7), (2.8) y (2.11), la función de verosimilitud puede ser escrita como:

$$f(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\xi}) = \prod_{i=1}^n \prod_{j=1}^J \prod_{k=0}^{m_j} [P_{ijk}]^{w_{ijk}} \quad (3.1)$$

$$= \prod_{i=1}^n \prod_{j=1}^J \prod_{k=0}^{m_j} \left[ \left( \prod_{r=0}^k \Psi(\eta_{ijr}) \right) (1 - \Psi(\eta_{ijk+1})) \right]^{w_{ijk}} \quad (3.2)$$

$$= \prod_{i=1}^n \prod_{j=1}^J \prod_{k=0}^{m_j} \left[ \left( \prod_{r=0}^k \frac{1}{1 + \exp(-a_j(\theta_i - b_{jr}))} \right) \left( \frac{\exp(-a_j(\theta_i - b_{jk+1}))}{1 + \exp(-a_j(\theta_i - b_{jk+1}))} \right) \right]^{w_{ijk}} \quad (3.3)$$

### 3.2. Especificación de la distribución a priori en el Modelo Secuencial Logístico (2PL-SM)

Se asumirá que la distribución a priori conjunta de los parámetros de las personas y de los ítems sigue una estructura de prioris independientes de la siguiente manera:

$$P(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}) = P(\boldsymbol{\theta})P(\mathbf{a}, \mathbf{b}) = \prod_{i=1}^n P(\theta_i) \prod_{j=1}^J P(a_j) \prod_{k=0}^{m_j} P(b_{jk}) \quad (3.4)$$

donde  $\mathbf{a} = (a_1, a_2, \dots, a_J)$  y  $\mathbf{b} = (b_1, b_2, \dots, b_J)$  con  $\mathbf{b}_j = (b_{j1}, \dots, b_{jm_j})$  contienen los parámetros de discriminación y dificultad respectivamente.

Se definen prioris informativas dado que utilizar prioris impropias no garantiza la existencia de la distribución conjunta a posteriori. Las distribuciones a priori propuestas se detallan a continuación:

- Para el vector de parámetros de habilidad de las personas se consideró una distribución normal estándar  $\theta_i \sim N(0, 1)$  con lo que se espera obtener  $E[\theta_i] = 0$  y  $V[\theta_i] = 1$ .
- Para el parámetro de discriminación Tarazona (2013) considera una distribución normal truncada positiva  $a_j \sim N_+(\mu_a, \sigma_a^2)$  con parámetros de locación  $\mu_a = 1$  y escala  $\sigma_a^2 = 0.5$  con lo que se espera obtener  $E[a_j] = 1 + \sqrt{0.5} \times \frac{\phi(-\sqrt{2})}{1-\phi(-\sqrt{2})} = 1.1126$  y  $V[a_j] = 0.5 \times \left\{ 1 - \frac{\phi(-\sqrt{2})}{1-\phi(-\sqrt{2})} \left[ \frac{\phi(-\sqrt{2})}{1-\phi(-\sqrt{2})} + \sqrt{2} \right] \right\} = 0.3747$
- Para los parámetros de dificultad de los ítems Tarazona (2013) considera una distribución normal  $b_{jk} \sim N(0, 2)$  con media cero y varianza dos con lo que se espera obtener  $E[b_j] = 0$  y  $V[b_j] = 2$ .

### 3.3. Distribución a posteriori de los parámetros del Modelo Secuencial Logístico (2PL-SM)

Considerando la función de verosimilitud dada en (3.3) y las prioris consideradas en (3.4), definimos la distribución a posteriori de los parámetros dado los datos como:

$$P(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b} | y) \propto P(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}) f(y | \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}) \quad (3.5)$$

$$\propto P(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}) \prod_{i=1}^n \prod_{j=1}^J \prod_{k=0}^{m_j} \left[ \left( \prod_{r=0}^k \Psi(\eta_{ijr}) \right) (1 - \Psi(\eta_{ijk+1})) \right]^{w_{ijk}} \quad (3.6)$$

$$\propto P(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}) \prod_{i=1}^n \prod_{j=1}^J \prod_{k=0}^{m_j} \left[ \left( \prod_{r=0}^k \frac{1}{1 + \exp(-a_j(\theta_i - b_{jr}))} \right) \left( \frac{\exp(-a_j(\theta_i - b_{jk+1}))}{1 + \exp(-a_j(\theta_i - b_{jk+1}))} \right) \right]^{w_{ijk}} \quad (3.7)$$

### 3.4. Inferencia Bayesiana del Modelo Secuencial Logístico (2PL-SM)

La distribución a posteriori dada en (3.7) no es identificable. Por lo tanto, para la estimación bayesiana se considera métodos de Montecarlo vía Cadenas de Markov (MCMC), utilizando el algoritmo de Gibbs, que es implementado en el software WinBUGS, para lo cual se escribe el código del modelo, en base a la especificación de la verosimilitud jerárquica del modelo general:

$$y_{ij} \mid \theta_i, a_j, b_{jk} \sim \text{geométrica truncada generalizada} (P_{ij0}, P_{ij1}, \dots, P_{ijm_j}, m_j) \quad (3.8)$$

$$P_{ijk} = \left( \prod_{r=0}^k \frac{P_{ijr}^+}{P_{ijr-1}^+} \right) \left( 1 - \frac{P_{ijk+1}^+}{P_{ijk}^+} \right), \quad (3.9)$$

$$\frac{P_{ijk}^+}{P_{ijk-1}^+} = \Psi(a_j(\theta_i - b_{jk})), \quad (3.10)$$

$$\theta_i \sim N(0, 1) \quad (3.11)$$

$$a_j \sim N_+(1, 0.5) \quad (3.12)$$

$$b_{jk} \sim N(0, 2) \quad (3.13)$$

### 3.5. Implementación Computacional

Para realizar la inferencia bayesiana del Modelo Secuencial logístico de dos parámetros, se programaron códigos en los software R (R Core Team, 2017) y WinBUGS14 (Lunn et al., 2007), este último como ya se mencionó en el punto anterior, utiliza el algoritmo de Gibbs y los métodos de MCMC para la estimación de los parámetros. El algoritmo de Gibbs, explicado en Hoff (2009), es utilizado cuando la distribución a posteriori conjunta de los parámetros no es conocida y es difícil de muestrear directamente. En este caso se muestrea de las distribuciones condicionales completas de cada parámetro. La aproximación de Gibbs es un algoritmo iterativo que construye una secuencia dependiente de los valores de los parámetros cuya distribución converge a la distribución posterior conjunta. Al final se obtiene una secuencia de vectores dependientes  $\phi^{(1)}, \dots, \phi^{(s)}$  donde cada vector contiene los parámetros del modelo.

En esta secuencia,  $\phi^{(s)}$  depende de  $\phi^{(0)}, \dots, \phi^{(s-1)}$  solo a través de  $\phi^{(s-1)}$ , esto es,  $\phi^{(s)}$  es independiente condicionalmente de  $\phi^{(0)}, \dots, \phi^{(s-2)}$  dado  $\phi^{(s-1)}$ . Esto es llamado propiedad de Markov, y la secuencia es llamada Cadena de Markov.

Además se tiene que,

$$P(\phi^{(s)} \in A) \rightarrow \int_A p(\phi) \partial\phi, \quad \text{cuando } s \rightarrow \infty. \quad (3.14)$$

siendo  $A$  un subconjunto arbitrario.

En otras palabras, la distribución empírica de  $\phi^{(s)}$  aproxima a la distribución objetivo cuando  $s \rightarrow \infty$ , sin importar el valor de inicio de  $\phi^{(0)}$ . De esta manera, para la mayoría de funciones  $g$  de interés,

$$\frac{1}{S} \sum_{s=1}^S g(\phi^{(s)}) \rightarrow E[g(\phi)] = \int g(\phi) p(\phi) \partial\phi, \quad \text{cuando } s \rightarrow \infty. \quad (3.15)$$

Esto significa que  $E[g(\phi)]$  puede ser aproximado con el promedio muestral de  $\{g(\phi^{(1)}), \dots, g(\phi^{(s)})\}$ , justo como en una aproximación Montecarlo. Por esta razón, estas aproximaciones son llamadas Métodos de Montecarlo vía Cadenas de Markov (MCMC).

La convergencia de las simulaciones vía MCMC, en principio, fue evaluada mediante la observación de los gráficos de diagnóstico proporcionados por el WinBUGS. En estos gráficos se presentan las densidades a posteriori de los parámetros estimados y las iteraciones de forma sucesiva para cada cadena implementada, esperando obtener una sobreposición de las mismas que deje en claro que hay una convergencia a una distribución común. Además se espera observar estacionariedad en las gráficas de iteraciones.

La convergencia también fue evaluada mediante el indicador Rhat propuesto en [Gelman et al. \(2014\)](#) el cual estima la disminución potencial en la variabilidad entre cadenas con respecto a la variabilidad dentro de las cadenas. Si el valor de este indicador fuera muy alto, entonces se tendrían razones para creer que es necesario un mayor número de simulaciones para mejorar nuestra inferencia sobre el parámetro objetivo. Un valor cercano a 1 indicaría que la convergencia ha sido alcanzada.

$$Rhat = \sqrt{\frac{\frac{N-1}{N}W + \frac{M+1}{MN}B}{W}} \quad (3.16)$$

donde  $M$  representa el número de cadenas,  $N$  representa el número de simulaciones,  $B$  representa la varianza entre cadenas y  $W$  representa la varianza dentro de las cadenas.

El código en R mostrado en el Anexo [A.6](#) contiene la función smMCMC 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 los capítulos 4 y 5 respectivamente, así como también para que permita configurar los parámetros para las simulaciones y los valores de los parámetros de las distribuciones a priori consideradas en el modelo. 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.

El código en BUGS para realizar la inferencia bayesiana bajo MCMC en WinBUGS del modelo 2PL-SM fue adaptado de [Tarazona \(2013\)](#) y se muestra en el Anexo [A.7](#).

Adicionalmente, en el Anexo [A.5](#) se muestra un código para simular datos provenientes del Modelo Secuencial, mientras que en los Anexos [A.1](#), [A.2](#), [A.3](#) y [A.4](#) 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), la Función de Información de los Ítems (FII) y la Función de Información del Test (FIT) respectivamente.



## Capítulo 4

# Estudio de Simulación

### 4.1. Objetivos

En este capítulo buscamos simular un conjunto de datos politómicos ordinales generado a partir del Modelo Secuencial (2PL-SM), con el objetivo de evaluar la precisión del método MCMC en la recuperación de los parámetros de los ítems y las habilidades del Modelo Secuencial (2PL-SM). Ello también nos proporciona sustento a los resultados obtenidos en el capítulo 5.

### 4.2. Algoritmo para la simulación de datos

Para la simulación de datos provenientes de ítems politómicos ordinales aplicables a un Modelo Secuencial (2PL-SM) se realizó el siguiente procedimiento:

- Definir la cantidad de sujetos  $n$ , cantidad de ítems  $J$  y número de categorías por cada ítem  $M$ .
- Generar los parámetros del modelo. Se simulan las habilidades para las personas considerando  $\theta_i \sim N(0, 1)$ . Los parámetros de dificultad y discriminación de los ítems pueden ser generados considerando  $b_{jk} \sim N_+(0, 2)$  y  $a_j \sim N_+(1, 0.5)$  tomando como referencia lo presentado en Tarazona (2013) para el modelo de respuesta graduada. En el caso del presente trabajo, las dificultades de los ítems son tomadas del ejemplo presentado en Tutz (2016) en el cual se plantean 6 ítems. La discriminación se estableció en 1 para los dos primeros ítems, 1.5 para los dos siguientes y 2 para los últimos. Ver cuadro 4.1.
- Calcular para cada ítem la probabilidad de alcanzar cada categoría  $k$ ,  $P_{ijk}$ , de acuerdo a (2.2).
- Simular para cada uno de los  $n$  sujetos su respuesta en cada ítem a partir de un número aleatorio sobre una población de  $k = 0, 1, \dots, M$  elementos, cada uno con probabilidad  $P_{ijk}$  que fue calculada en el paso anterior.

Para este proceso de simulación, los parámetros  $a_j$ ,  $b_{jk}$  y habilidades  $\theta_i$  se mantuvieron fijos a lo largo de todas las réplicas de la simulación. Todo estos insumos son utilizados para

la simulación de las respuestas en cada réplica. Como ya se mencionó anteriormente, el código para simular respuestas politómicas ordinales basadas en el Modelo Secuencial 2PL-SM se presenta en el Anexo A.5.

### 4.3. Metodo para la estimación de parámetros

La estimación de los parámetros del Modelo Secuencial (2PL-SM) se realizó mediante inferencia bayesiana, se consideró como estimador a la media posteriori, la cual fue calculada utilizando el método MCMC y fue implementado en el programa WinBUGS14. Con este método se estiman  $M \times J + n$  parámetros en cada réplica de la simulación. Estos parámetros son estimados calculando las medias estimadas de sus distribuciones a posteriori.

En el Anexo A.7 se presenta la implementación computacional para la inferencia bayesiana del Modelo Secuencial (2PL-SM) y en el Anexo A.6 se presenta la función en R usando la librería R2WinBUGS que permite ejecutar la inferencia bayesiana en WinBUGS.

### 4.4. Criterios para evaluar la simulación

Dos indicadores son utilizados para evaluar la precisión en la recuperación de los parámetros, la raíz del error cuadrático medio (RMSE) y el promedio del error absoluto (MAE).

- El RMSE se define como la raíz cuadrada del promedio de las diferencias entre los valores reales y los estimados a lo largo de todas las réplicas:

$$RMSE(\lambda) = \sqrt{\frac{1}{R} \sum_{s=1}^R (\hat{\lambda}_s - \lambda_s)^2} . \quad (4.1)$$

donde  $\lambda$  representa los diferentes parámetros del modelo,  $\hat{\lambda}$  representa el estimado de dichos parámetros y R el número de réplicas en la simulación.

- El MAE es definido como el promedio de las diferencias en valor absoluto entre los valores reales y los estimados a lo largo de todas las réplicas:

$$MAE(\lambda) = \frac{1}{R} \sum_{s=1}^R |\hat{\lambda}_s - \lambda_s| . \quad (4.2)$$

Además de los criterios mencionados, también son calculadas, dentro de cada réplica, las correlaciones entre los valores estimados y los parámetros generados.

$$corr = \frac{\sum_{s=1}^R (\hat{\lambda}_s - \lambda_s)(\hat{\lambda}_s - \lambda_s)}{(\sqrt{\sum_{s=1}^R (\hat{\lambda}_s - \lambda_s)^2})(\sqrt{\sum_{s=1}^k (\hat{\lambda}_s - \lambda_s)^2})} . \quad (4.3)$$

#### 4.5. Escenario

Se consideraron  $n = 1500$  individuos,  $J = 6$  ítems con  $M = 3$  categorías cada uno. Los parámetros de dificultad y discriminación de los ítems son mostrados en el cuadro 4.1. También se consideraron  $R=20$  réplicas, cada una con un conjunto diferente de datos. La implementación computacional del presente estudio de simulación se encuentra en el Anexo A.8.

Ítem	$b1_j$	$b2_j$	$a_j$
1	-0.527	-0.279	1.000
2	-0.365	-1.448	1.000
3	-0.365	-0.421	1.500
4	-0.692	0.080	1.500
5	-0.748	0.958	2.000
6	-0.979	0.233	2.000

Tabla 4.1: Parámetros de los ítems usados en la simulación

#### 4.6. Resultados

Se realizaron algunos estudios preliminares con el objetivo de determinar el número de iteraciones necesarias para asegurar la convergencia de las estimaciones de los parámetros con WinBUGS en cada réplica. De acuerdo a dichos estudios se utilizaron 2 cadenas de 40000 iteraciones cada una, de las cuales las 10000 primeras fueron descartadas mientras que para las 30000 restantes se consideraron saltos de 20 en cada cadena con el objetivo de reducir la autocorrelación entre iteraciones. De esta manera quedaron finalmente 1500 iteraciones efectivas por cadena.

Se evaluó la convergencia de la distribución a posteriori de los parámetros esperando una distribución estacionaria. Esta convergencia fue evaluada de forma visual a través de los gráficos de diagnóstico proporcionados por WinBUGS a través de las librerías mcmcplots y R2WinBUGS del R. Dichos gráficos son presentados en el Anexo B.

En la Figura B.1 se presenta el histórico para la discriminación y el primer umbral del cuarto ítem utilizando dos cadenas. Se puede apreciar que ambas cadenas están sobrepuestas y que además convergen a una distribución estacionaria.

En la Figura B.2 se presentan las gráficas de densidades y autocorrelaciones para los parámetros del cuarto ítem. Ambas cadenas están sobrepuestas y tampoco existe un gran problema de autocorrelación.

Cabe resaltar que cada criterio de convergencia fue evaluado con una réplica distinta y que además el resto de ítems mostraron un comportamiento similar.

Por último, para complementar el análisis se obtuvieron valores de Rhat, todos ellos cercanos a 1. Esto nos indicaría una buena convergencia en los resultados.

En el cuadro 4.2 se presentan los resultados del sesgo de las estimaciones de los parámetros de los ítems. Las medidas RMSE y MAE se presentan en el cuadro 4.3. Los valores de estas

medidas nos indican que la recuperación de parámetros se da de manera satisfactoria.

Ítem	$b1_j$	$b2_j$	$a_j$
1	0.014	-0.017	-0.028
2	0.041	0.015	0.007
3	0.001	0.026	0.031
4	0.001	0.006	-0.037
5	-0.006	0.028	-0.037
6	-0.019	0.015	-0.095

Tabla 4.2: Sesgo en la recuperación de parámetros de los ítems usando estimación bajo MCMC para el modelo 2PL-SM

Medida de Precisión	Ítem	$b1_j$	$b2_j$	$a_j$
RSME	1	0.067	0.104	0.073
	2	0.074	0.157	0.084
	3	0.040	0.058	0.076
	4	0.044	0.046	0.103
	5	0.038	0.048	0.104
	6	0.056	0.044	0.167
MAE	1	0.055	0.078	0.058
	2	0.064	0.126	0.074
	3	0.033	0.046	0.058
	4	0.037	0.039	0.089
	5	0.033	0.039	0.094
	6	0.045	0.040	0.145

Tabla 4.3: Raíz del error cuadrático medio (RMSE) 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-SM) considerando R=20 réplicas

Parámetro	RMSE	MAE	r
$b1$	0.053	0.044	0.983
$b2$	0.076	0.061	0.997
$a$	0.101	0.086	0.976
$\theta$	0.459	0.376	0.877

Tabla 4.4: 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 Secuencial Logístico (2PL-SM) considerando R= 20 réplicas

#### 4.7. Conclusiones

- Los parámetros de discriminación, dificultad y las habilidades de las personas para el Modelo Secuencial (2PL-SM) en una muestra de 1500 individuos con 6 ítems de 3 categorías cada uno fueron recuperados de manera satisfactoria aplicando inferencia bayesiana vía MCMC con el programa WinBUGS.
- Dado que la recuperación de parámetros fue exitosa, se puede concluir que la simulación de los datos politómicos ordinales generados a partir del Modelo Secuencial (2PL-SM) se llevó a cabo de manera adecuada. Las probabilidades de alcanzar cada categoría  $k$  fueron calculadas a partir de (2.2).

- Se necesitó un gran número de iteraciones para lograr la convergencia de las estimaciones. En cada réplica, en total se utilizaron 2 cadenas de 40000 iteraciones cada una, desechando las 10000 primeras y para las 30000 restantes se consideraron saltos de 20 en cada cadena para reducir la autocorrelación entre iteraciones.
- El proceso de estimación de parámetros utilizando el programa Winbugs en una computadora con un procesador Intel(R) Core (TM) i7-3770 CPU 3.40 GHz para una muestra de 1500 individuos con 6 ítems de 3 categorías cada uno es lento (en promedio 1 hora y 2 minutos por cada réplica).



## Capítulo 5

### Aplicación

La aplicación en el presente trabajo se encuentra en el ámbito educativo, específicamente en la medición de rendimiento estudiantil en el área de escritura. En [Minedu-UMC \(2016\)](#) se define la escritura como un proceso cognitivo y social que involucra el uso de diversos conocimientos, estrategias y recursos con el propósito de producir textos escritos en diferentes situaciones de comunicación.

En el 2013 el Ministerio de Educación (*MINEDU*) llevó a cabo la Evaluación Muestral 2013 (*EM2013*) en sexto grado de primaria, la cual incluyó una prueba de escritura con el propósito de brindar resultados sobre cuánto lograron aprender los estudiantes respecto a la producción de textos narrativos y argumentativos.

El objetivo de esta aplicación es la estimación bayesiana con el método MCMC de la habilidad latente y los parámetros de discriminación y dificultad en la EM2013 de los ítems con el Modelo Secuencial Logístico ( $2PL - SM$ ) en una muestra de 4327 estudiantes de 6to de primaria.

#### 5.1. Instrumento de medición

En el 2013 se realizó la Evaluación Muestral de Estudiantes (EM 2013) en sexto de primaria. Este estudio evaluó el rendimiento estudiantil en áreas como lectura, matemática, ciudadanía y escritura. Para el presente trabajo se utilizó la prueba de escritura, específicamente la prueba de escritura en su versión narrativa.

La prueba de escritura consistía en pedirle al estudiante que redacte un texto con características narrativas (cuento), con un destinatario implícito (los jueces de un concurso) y una estructura definida (el estudiante debía contar como empieza, que sucede luego y cómo termina). Posteriormente los textos de todos los estudiantes pasan a un proceso de codificación conducido por expertos, docentes y lingüistas especialmente entrenados para esta tarea, quienes codificaron el texto según una serie de criterios. Finalmente, los ítems que conformaron la prueba de escritura son los siguientes:

1. Secuencia textual: Este criterio evalúa en qué medida el texto presenta una secuencia textual (descriptiva, narrativa o argumentativa) completa.

2. Género cuento: Este criterio evalúa en qué medida el texto presenta las características propias del género solicitado.
3. Adecuación al tema: Este criterio evalúa en qué medida el texto presenta el tema planteado por la consigna.
4. Registro lingüístico: Este criterio evalúa en qué medida el texto presenta un registro (modo de expresarse) apropiado según la situación comunicativa.
5. Mantenimiento del tema: Este criterio evalúa en qué medida el texto presenta un tema que se desarrolla sin digresiones (sin desviarse del tema).
6. Progresión temática: Este criterio evalúa en qué medida el texto presenta ideas que no se reiteran innecesariamente.
7. Suficiencia comunicativa: Este criterio evalúa en qué medida el texto presenta toda la información necesaria para su comprensión.
8. Conexión: Este criterio evalúa en qué medida el texto presenta un uso adecuado de los conectores al relacionar proposiciones.
9. Referencia: Este criterio evalúa en qué medida el texto presenta un uso adecuado de los mecanismos de referencia (anáfora, catáfora, elipsis).
10. Desarrollo: Este criterio evalúa en qué medida el texto presenta enunciados que expanden, precisan o amplían ideas mencionadas previamente en el texto.
11. Estructura narrativa: Este criterio evalúa en qué medida el texto presenta una argumentación eficaz, convincente.

Cada uno de estos criterios tiene un número distinto de categorías ordinales que el estudiante puede alcanzar, donde las categorías más altas implican un mayor nivel de habilidad. La asignación de estas categorías se realizó de manera secuencial; es decir, para alcanzar una categoría el estudiante debe haber pasado todas las anteriores. Además, si el estudiante no alcanzaba una categoría el proceso de codificación se detenía.

A manera de ejemplo se presenta la rúbrica utilizada para calificar al ítem 1:

3	2	1
El texto narra mediante una secuencia en la que hay una situación inicial, una situación central y una situación final, independientemente del orden en que se presenten.	El texto narra mediante una secuencia narrativa de al menos dos situaciones relacionadas entre las que está presente la situación central.	el texto solo presenta un evento sin articularlo con otros momentos narrativos. O, el texto articula otra secuencia (describe, argumenta, etc).

## 5.2. Diseño Muestral

La prueba de escritura de la EM 2013 fue aplicada a un total de 4327 estudiantes de sexto grado de primaria de distantes escuelas alrededor del país. El muestreo utilizado fue estratificado y por conglomerados bietápico. Los estratos fueron la combinación área geográfica (urbana y rural) y gestión (estatal y no estatal). Como conglomerado se consideró a la Institución Educativa y dentro de esta se seleccionó una sección en el caso de tener 2 y 2 secciones en el caso de tener 3 o más.

## 5.3. Recolección de datos

Los datos fueron proporcionados por la Oficina de Medición de la Calidad de los Aprendizajes (UMC) del Ministerio de Educación. Cabe señalar que estos datos no fueron analizados mediante algún modelo de la TRI por la UMC.

## 5.4. Estimación de los parámetros mediante el Modelo Secuencial

Como ya se mencionó en el capítulo anterior, la implementación del modelo se realizó con la librería R2WinBUGS del software R y el programa WinBUGS. Los códigos desarrollados en WinBUGS se basaron en lo presentado por Tarazona (2013) siendo adaptados al Modelo Secuencial. La aplicación fue ejecutada para 1500 iteraciones efectivas en cada 2 cadenas (40000 iteraciones por cadena donde se quemaron las 10000 primeras y utilizando saltos de 20), es decir, un total de 3000 iteraciones efectivas.

### 5.4.1. Modelo Secuencial Logístico (2PL-SM)

En el Modelo Secuencial Logístico (2PL-SM), la probabilidad  $P_{ijk}$  de alcanzar la categoría  $k$  en el ítem  $j$  depende del nivel de habilidad que tiene el estudiante  $i$  en escritura argumentativa y de los parámetros de discriminación  $a_j$  y dificultad  $b_{jk}$ , y está definida como en 2.22. El código para estimar el Modelo 2PL-SM se encuentra en el Anexo A.7.

Los valores estimados de los parámetros de discriminación  $a_j$ , sus errores estándar ( $ee_a$ ), sus medianas y otros estadísticos de orden son presentados en el Cuadro 5.1. Los ítems Mantenimiento del tema, Progresión temática y Suficiencia comunicativa (ítems 5, 6 y 7 respectivamente) son aquellos que presentan mayor discriminación, esto indicaría que en estos ítems un cambio en la habilidad del estudiante tiene un mayor impacto en la probabilidad de acertarlo con respecto a otros ítems de menor discriminación. Los ítems Registro lingüístico y Conexión (ítems 4 y 8) son aquellos que menos discriminan entre estudiantes.

Con respecto a los valores estimados de los parámetros de dificultad  $b_{jk}$ , estos se muestran en el Cuadro 5.2. Cabe señalar que no todos los ítems tenían el mismo número de categorías, además también se puede observar que todos los parámetros de dificultad resultaron negativos lo que nos indicaría que la prueba no resultó del todo difícil.

Utilizando las librerías mcmcplots (Curtis, 2018) y R2WinBUGS del R se examinó la convergencia de forma visual, algunos gráficos son presentados en el Anexo B.2. En la Figura

B.3 se muestra el histórico de dos cadenas, gráficas de densidades y autocorrelaciones para el parámetro de discriminación del ítem 3 (Adecuación al tema). Por otro lado, en la Figura B.4 se muestran gráficas de diagnóstico similares pero para el segundo umbral del ítem 9 (Referencia). En general se puede observar que no existe un problema grave de autocorrelación y que las dos cadenas convergen. El resto de ítems mostraron un comportamiento similar. Además del diagnóstico visual se obtuvieron valores de Rhat muy cercanos a 1 para cada uno de los parámetros, esto nos indicaría una buena precisión en los resultados.

En las Gráficas 5.1 y 5.2 se muestran, a manera de ejemplo, las FREI y FRCI para los ítems 5 y 11. El código en R para el gráfico de estas funciones se encuentra en A.1 y A.2 respectivamente.

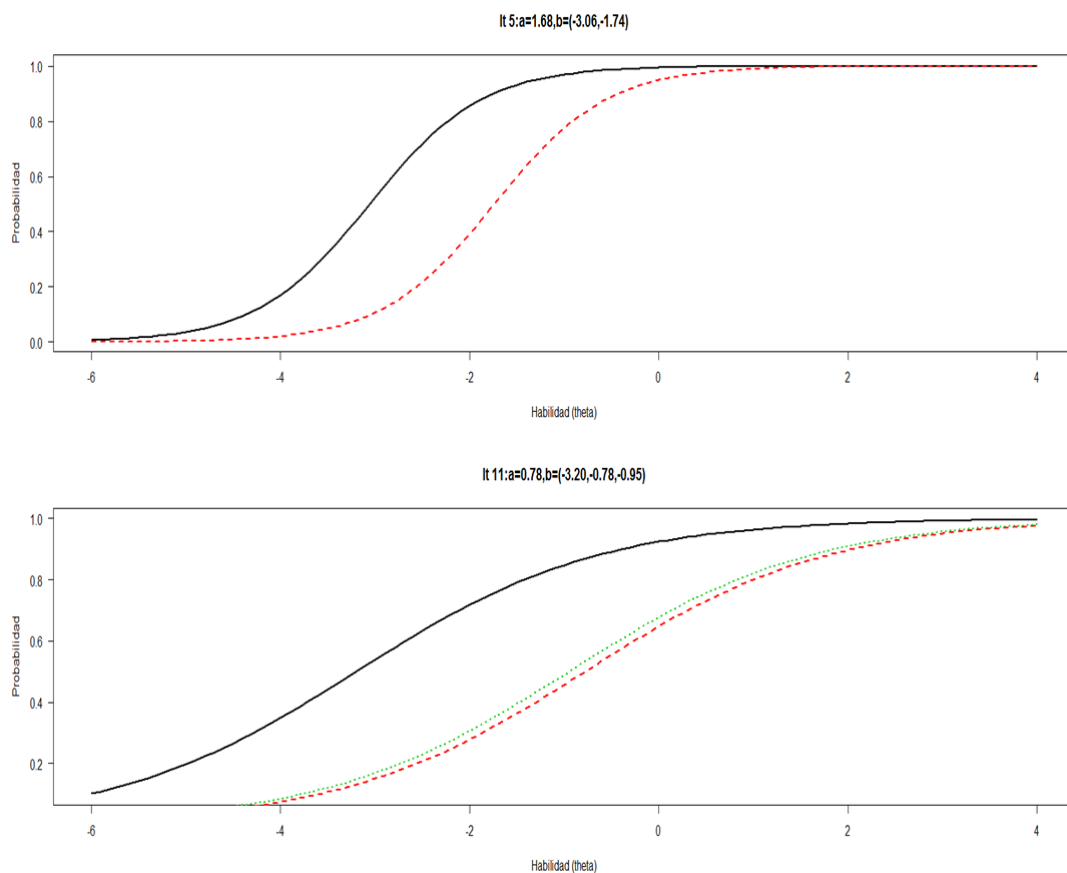


Figura 5.1: FREI del modelo 2PL-SM para los ítems 5 (Mantenimiento del tema) y 11 (Estructura narrativa)

En el Gráfico 5.3 se presenta las funciones de información para todos los ítems. El ítem 8 (Conexión) proporciona mayor información para valores de habilidad entre -8 y -5 mientras que el resto de ítems proporcionan mayor información para valores de habilidad entre -5 y 1. Mayor información indica una mayor precisión para la estimación  $\theta$  en ese rango de valores.

En el Gráfico 5.4 se presenta la función de información del test. En general el test de escritura aporta mayor información para valores de habilidad entre -4 y 1. En otras palabras, valores de  $\theta$  serán estimados con mayor precisión si se encuentran entre -4 y 1.

Ítem	$a_j$	$ee_{-a_j}$	2.5 %	Mediana	97.5 %
1	1.22	0.08	1.07	1.22	1.39
2	1.47	0.13	1.23	1.46	1.73
3	1.11	0.07	0.97	1.11	1.25
4	0.65	0.05	0.55	0.64	0.75
5	1.68	0.11	1.48	1.68	1.90
6	1.53	0.10	1.34	1.53	1.74
7	1.55	0.10	1.36	1.55	1.76
8	0.62	0.06	0.51	0.62	0.75
9	0.98	0.06	0.87	0.98	1.11
10	0.95	0.07	0.82	0.95	1.09
11	0.78	0.04	0.69	0.78	0.87

Tabla 5.1: Estimación de los parámetros de discriminación del Modelo Secuencial 2PL-SM

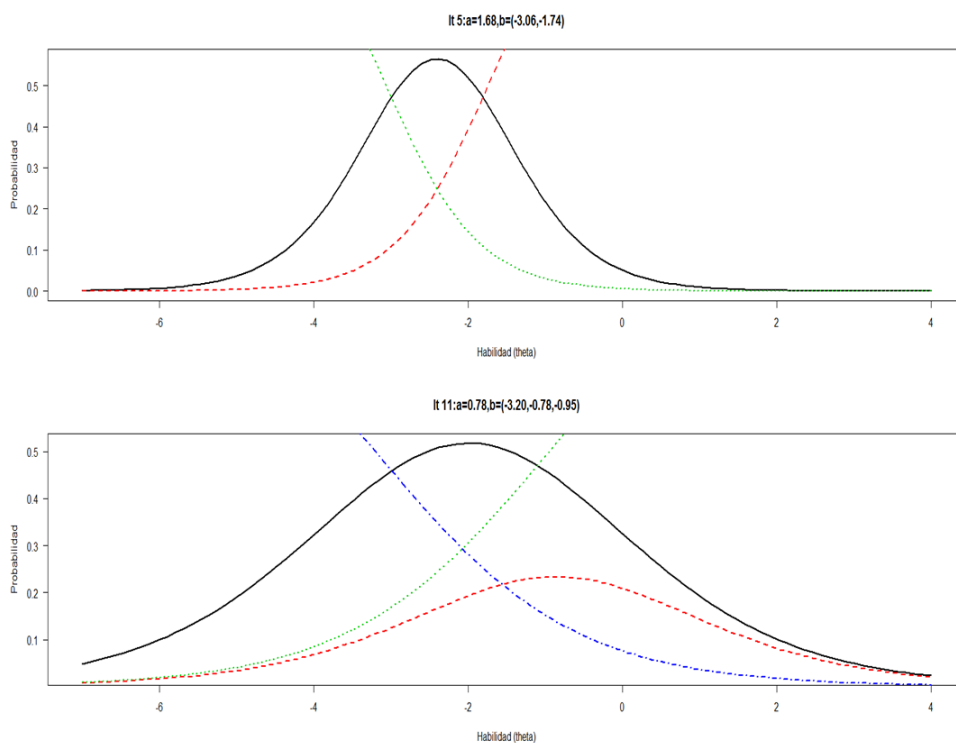


Figura 5.2: FRCI del modelo 2PL-SM para los ítems 5 (Mantenimiento del tema) y 11 (Estructura narrativa)

Ítem	Estadístico	b1	b2	b3	número de categorías
1	Media	-6.74	-2.28	—	3
	ee_b	0.57	0.12	—	
	2.5%	-7.94	-2.53	—	
	Mediana	-6.70	-2.28	—	
	97.5%	-5.74	-2.07	—	
2	Media	-2.80	—	—	2
	ee_b	0.17	—	—	
	2.5%	-3.15	—	—	
	Mediana	-2.78	—	—	
	97.5%	-2.51	—	—	
3	Media	-3.26	-2.20	—	3
	ee_b	0.17	0.12	—	
	2.5%	-3.64	-2.45	—	
	Mediana	-3.25	-2.19	—	
	97.5%	-2.95	-1.99	—	
4	Media	-8.60	-4.47	—	3
	ee_b	0.63	0.32	—	
	2.5%	-9.88	-5.13	—	
	Mediana	-8.60	-4.46	—	
	97.5%	-7.38	-3.88	—	
5	Media	-3.06	-1.74	—	3
	ee_b	0.13	0.07	—	
	2.5%	-3.34	-1.89	—	
	Mediana	-3.05	-1.73	—	
	97.5%	-2.81	-1.60	—	
6	Media	-1.88	—	—	2
	ee_b	0.08	—	—	
	2.5%	-2.04	—	—	
	Mediana	-1.87	—	—	
	97.5%	-1.73	—	—	
7	Media	-1.68	—	—	2
	ee_b	0.07	—	—	
	2.5%	-1.83	—	—	
	Mediana	-1.68	—	—	
	97.5%	-1.55	—	—	
8	Media	-2.99	—	—	2
	ee_b	0.27	—	—	
	2.5%	-3.59	—	—	
	Mediana	-2.97	—	—	
	97.5%	-2.53	—	—	
9	Media	-3.14	-1.04	—	3
	ee_b	0.17	0.07	—	
	2.5%	-3.49	-1.17	—	
	Mediana	-3.14	-1.04	—	
	97.5%	-2.84	-0.91	—	
10	Media	-2.05	—	—	2
	ee_b	0.12	—	—	
	2.5%	-2.31	—	—	
	Mediana	-2.04	—	—	
	97.5%	-1.82	—	—	
11	Media	-3.20	-0.78	-0.95	4
	ee_b	0.17	0.06	0.09	
	2.5%	-3.55	-0.92	-1.14	
	Mediana	-3.19	-0.78	-0.94	
	97.5%	-2.88	-0.67	-0.78	

Tabla 5.2: Estimación de los parámetros de dificultad del Modelo Secuencial 2PL-SM

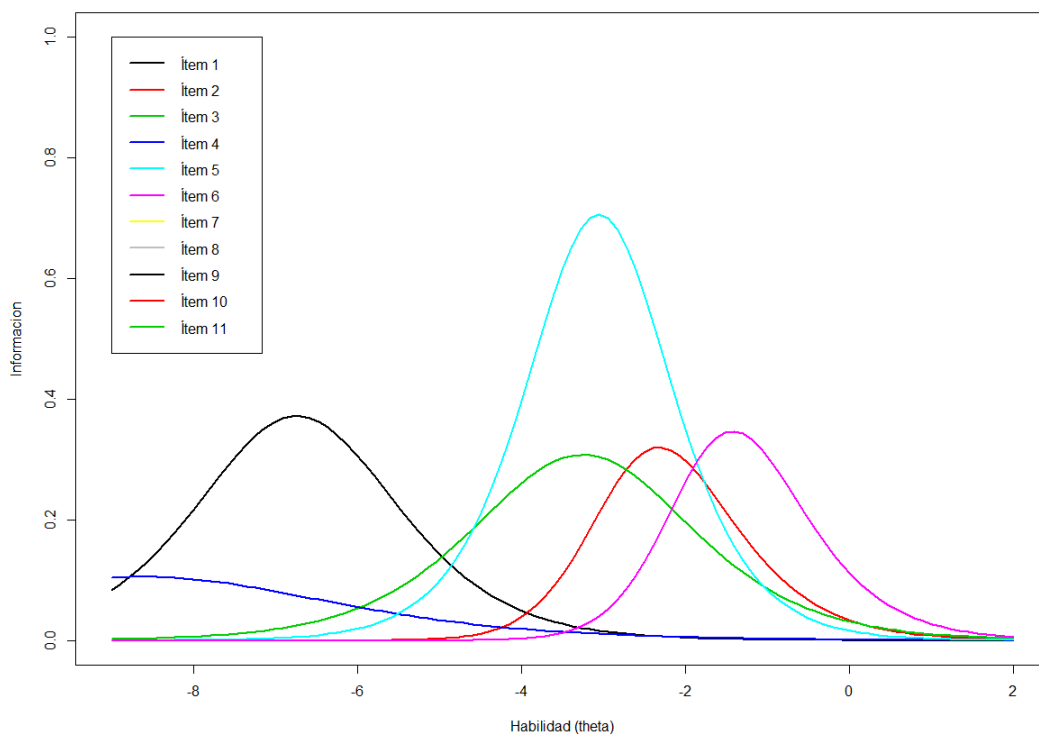


Figura 5.3: FII del modelo 2PL-SM para la prueba de escritura

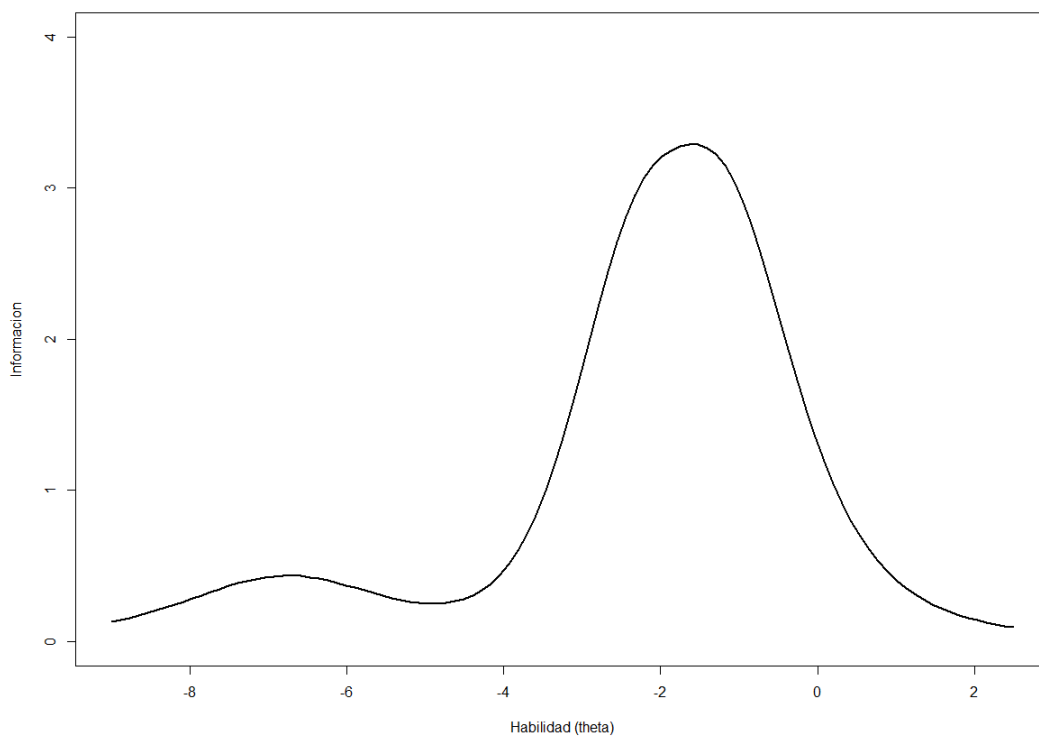


Figura 5.4: FIT del modelo 2PL-SM para la prueba de escritura

Adicionalmente, la distribución de las habilidades en escritura de los estudiantes  $\theta$  se presenta en el Gráfico 5.5, aquí se puede observar que dichos valores se encuentran principalmente entre -1 y 1.

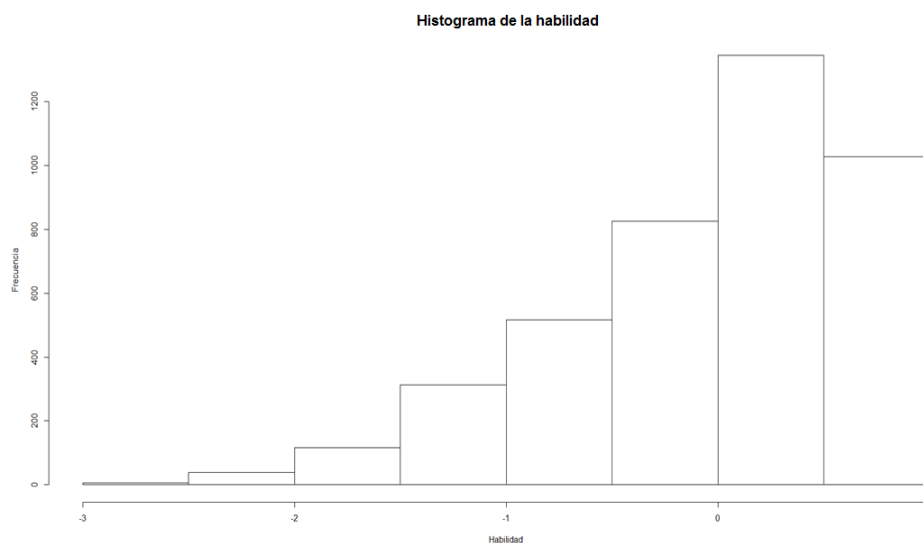


Figura 5.5: Distribución de la habilidad estimada de los estudiantes de 6to de primaria en la prueba de escritura usando el Modelo 2PL-SM

Finalmente, en el gráfico 5.4 se observa claramente que el test aporta información principalmente para valores de habilidad entre -4 y 1 mientras que la distribución de las habilidades estimadas presentadas en el gráfico 5.5 presenta una mayor frecuencia de valores entre -1 y 1. Esto indicaría que en general la habilidad de los estudiantes es superior a las dificultades de los ítems y por lo tanto el enfrentarse a esta prueba no debería resultarles del todo complicado.

## Capítulo 6

### Conclusiones

#### 6.1. Conclusiones

- En el presente trabajo se presenta una breve descripción de los principales modelos TRI para ítems politómicos. Se tiene como finalidad presentar algunas características de dichos modelos, así como la identificación del tipo de ítem apropiado para cada modelo.
- Se presenta el Modelo Secuencial Logístico (2PL-SM), el cual está diseñado para ser aplicado en ítems que deben ser resueltos de manera consecutiva. Este modelo tiene como función de enlace la función logística y considera la estimación de los parámetros de discriminación y dificultad para los ítems.
- Se implementa la estimación de los parámetros del Modelo (2PL-SM) a través de la inferencia bayesiana con MCMC utilizando el software WinBUGS debido a la facilidad para su implementación. Cabe resaltar que el proceso de estimación suele tomar mucho tiempo.
- Se realizó un estudio de simulación de 6 ítems de 3 categorías con el objetivo de evaluar la precisión del método MCMC en la recuperación de los parámetros del Modelo (2PL-SM). Dados los resultados del sesgo, RMSE y MAE se puede concluir que los parámetros fueron recuperados adecuadamente.
- La aplicación fue realizada con datos de la Prueba de escritura de sexto de primaria de la Evaluación Muestral 2013 del Ministerio de Educación. De los resultados obtenidos se concluye que:
  - El Modelo Secuencial Logístico fue aplicado de manera satisfactoria en el conjunto de datos estudiados.
  - La prueba contiene ítems cuyas dificultades se encuentran en la parte baja de la escala, comparados con las estimaciones de habilidad de los estudiantes.
  - Con respecto a la estimación de los parámetros se puede concluir que los ítems más difíciles son el 9 (Referencia) y el 11 (Estructura Narrativa) mientras que los más fáciles fueron el 1, el 3 y el 4 (Secuencia textual, Adecuación al tema y Registro

lingüístico respectivamente). Los que más discriminan son el 5 (Mantenimiento del tema), el 6 (Progresión temática) y el 7 (Suficiencia comunicativa) mientras que los de menor discriminación son el 4 (Registro lingüístico) y el 8 (Conexión).

## 6.2. Sugerencias para investigaciones futuras

- Realizar estudios con una mayor diversidad de escenarios, utilizando diferentes números de individuos, ítems y categorías.
- Extender el estudio del Modelo Secuencial para situaciones de multidimensionalidad en donde la habilidad o rasgo latente del entrevistado tenga más de un factor dominante.
- Realizar el estudio del Modelo Secuencial con funciones de enlace distintas a la logística.
- Optimizar el tiempo de estimación de los parámetros del modelo considerando otras programas como Stan.



## Apéndice A

### Códigos en R y BUGS

#### A.1. Código en R para graficar la Función de Respuesta a las Etapas de los Ítems (FREI) del SM

```
#####  
# Autor : Luis Ángel Mejía Campos #  
# # #  
# Función: frei.sm.pirt() #  
# # #  
# Fecha de inicio : 16-08-2017 #  
# # #  
# Función : Grafica la Función de Respuesta a las Etapas de los Ítems #  
# (FREI) para el Modelo Secuencial (SM) #  
# # #  
# Asesor :Enver Gerald Tarazona Vargas #  
# # #  
# Versión : 2.0 #  
# # #  
# Última Modificación : 29-01-2018 #  
#####  
  
frei.sm.pirt<- function(item,  
rangex=c(-6,6),  
main=""  
)  
{  
m<-length(item)  
for(k in 1:(m-1)){  
Pac<-function(x) plogis(item[m] * (x - item[k]))  
plot(Pac,add=k>1,xlim=rangex,col=k,xlab="Habilidad (theta)",  
ylab="Probabilidad",lty=k, lwd=2, main=main,las=1)  
}  
}
```

```
}

```

```
item1<-c(-2.0,0.1,0.8,1.4,1.0)
item2<-c(-2.0,0.1,0.8,1.4,2.0)
item3<-c(-2.0,0.1,-0.5,1.4,2.0)
par(mfrow = c(3, 1))
frei.sm.pirt(item1,rangex=c(-6,6), main="It 1:a=1,b=(-2.0,0.1,0.8,1.4)")
frei.sm.pirt(item2,rangex=c(-6,6), main="It 2:a=2,b=(-2.0,0.1,0.8,1.4)")
frei.sm.pirt(item3,rangex=c(-6,6), main="It 2:a=2,b=(-2.0,0.1,-0.5,1.4)")
```

## A.2. Código en R para graficar la Función de Respuesta a las Categorías de los Ítems (FRCI) del SM

```
#####
# Autor : Luis Ángel Mejía Campos #
# # #
# Función: frci.sm.pirt() #
# # #
# Fecha de inicio : 16-08-2017 #
# # #
# Función : Grafica la Función de Respuesta a las Categorías de los Ítems #
# (FRCI) para el Modelo Secuencial (SM) #
# # #
# Asesor : Enver Gerald Tarazona Vargas #
# # #
# Versión : 2.0 #
# # #
# Última Modificación : 29-01-2018 #
#####
temp=1

frci.sm.pirt<-function(item,
rangex=c(-6,6),
main=""
)
{
m<-length(item)
for(k in 1:m){
P<-function(x) {

temp=temp*plogis(item[m] * (x - item[k]))
```

```

if(k < (m-1)){
temp*(1-plogis(item[m] * (x - item[k+1])))
}

else{

if(k ==(m-1)){
temp
}
else{
1-plogis(item[m] * (x - item[1]))
}
}
}

plot(P,add=k>1,xlim=rangex,col=k,xlab="Habilidad (theta)",
ylab="Probabilidad",lty=k, lwd=2,main=main,las=1)
}
}

item1<-c(-3.5,-2.0,-0.5,1.0,1)
item2<-c(-3.5,-2.0,-0.5,1.0,2)
item3<-c(-1.0,0.0,1.0,2.0,1)
item4<-c(-1.0,0.0,1.0,2.0,2)
item5<-c(-0.5,1.0,2.5,3.5,1)
item6<-c(-0.5,1.0,2.5,3.5,2)

par(mfrow = c(3, 2))
frci.sm.pirt(item1,rangex=c(-6,6), main="It 1:a=1,b=(-3.5,-2.0,-0.5,1.0)")
frci.sm.pirt(item2,rangex=c(-6,6), main="It 2:a=2,b=(-3.5,-2.0,-0.5,1.0)")
frci.sm.pirt(item3,rangex=c(-6,6), main="It 3:a=1,b=(-1.0,0.0,1.0,2.0)")
frci.sm.pirt(item4,rangex=c(-6,6), main="It 4:a=2,b=(-1.0,0.0,1.0,2.0)")
frci.sm.pirt(item5,rangex=c(-6,6), main="It 5:a=1,b=(-0.5,1.0,2.5,3.5)")
frci.sm.pirt(item6,rangex=c(-6,6), main="It 6:a=2,b=(-0.5,1.0,2.5,3.5)")

```

### A.3. Código en R para graficar la Función de Información de los Ítems (FII) del SM

```

#####
# Autor : Luis Ángel Mejía Campos #
# # #
# Función: fii.sm.pirt() #
# # #
# Fecha de inicio : 23-08-2017 #

```

```

# #
# Función : Graficar la Función de Información de los Ítems (FII) #
# para el Modelo Secuencial (SM) #
# #
# Asesor : Enver Gerald Tarazona Vargas #
# #
# Versión : 2.0 #
# #
# Última Modificación :29-01-2018 #
#####

temp=1
temp2=0
I=0

fii.sm.pirt<-function(item,
rangex=c(-6,6),
ylim=c(0,3.5),
col=1,lty=1,add=T,
main=""
)
{

m<-length(item)
n<-length(item)-1

fii<-function(x) {
for(k in 1:n){

temp=temp*plogis(item[m] * (x - item[k]))
temp2=temp2+item[m]*(1-plogis(item[m] * (x - item[k])))

if(k < (n-1)){
p=temp*(1-plogis(item[m] * (x - item[k+1])))
L=temp2-item[m]*(plogis(item[m] * (x - item[k+1])))
I=I+L^2*p
}

else{

if(k ==(n-1)){
p=temp

```

```

L=temp2
I=I+L^2*p
}
else{
p=1-plogis(item[m] * (x - item[1]))
L=-item[m]*plogis(item[m] * (x - item[1]))
I=I+L^2*p
}
}
}
I
}

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

item1<-c(-3.5,-2.0,-0.5,1.0,1)
item2<-c(-3.5,-2.0,-0.5,1.0,2)
item3<-c(-1.0,0.0,1.0,2.0,1)
item4<-c(-1.0,0.0,1.0,2.0,2)
item5<-c(-0.5,1.0,2.5,3.5,1)
item6<-c(-0.5,1.0,2.5,3.5,2)

fii.sm.pirt(item=item1,rangex=c(-6,6),ylim=c(0,3.5),col=1,lty=1,add=F,
main="")
fii.sm.pirt(item=item2,rangex=c(-6,6),ylim=c(0,3.5),col=2,lty=1,add=T,
main="")
fii.sm.pirt(item=item3,rangex=c(-6,6),ylim=c(0,3.5),col=3,lty=1,add=T,
main="")
fii.sm.pirt(item=item4,rangex=c(-6,6),ylim=c(0,3.5),col=4,lty=1,add=T,
main="")
fii.sm.pirt(item=item5,rangex=c(-6,6),ylim=c(0,3.5),col=5,lty=1,add=T,
main="")
fii.sm.pirt(item=item6,rangex=c(-6,6),ylim=c(0,3.5),col=6,lty=1,add=T,
main="")

legend(-6,3.5, c("Ítem 1","Ítem 2","Ítem 3","Ítem 4","Ítem 5","Ítem 6"),
lty=rep(1,6),lwd=rep(2.5,6),col=c(1,2,3,4,5,6))

```

#### A.4. Código en R para graficar la Función de Información del Test (FIT) del SM

```
#####
# Autor : Luis Ángel Mejía Campos                                     #
#                                                                 #
# Función: fit.sm.pirt()                                           #
#                                                                 #
# Fecha de inicio : 30-08-2017                                     #
#                                                                 #
# Función : Graficar la Función de Información del Test (FIT)    #
# para el Modelo Secuencial (SM)                                  #
#                                                                 #
# Asesor : Enver Gerald Tarazona Vargas                           #
#                                                                 #
# Versión : 2.0                                                    #
#                                                                 #
# Última Modificación :29-01-2018                                 #
#####

temp=1
temp2=0
I=0
IT=0

fit.sm.pirt<-function(test,
rangex=c(-6,6),
ylim=c(0,3.5),
col=1,lty=1,add=T,
main=""
)
{

fii<-function(x) {

for(z in 1:6){

item=test[z,]
m<-length(item)
n<-length(item)-1
for(k in 1:n){
```

```

temp=temp*plogis(item[m] * (x - item[k]))
temp2=temp2+item[m]*(1-plogis(item[m] * (x - item[k])))

if(k < (n-1)){
p=temp*(1-plogis(item[m] * (x - item[k+1])))
L=temp2-item[m]*(plogis(item[m] * (x - item[k+1])))
I=I+L^2*p
}

else{

if(k ==(n-1)){
p=temp
L=temp2
I=I+L^2*p
}
else{
p=1-plogis(item[m] * (x - item[1]))
L=-item[m]*plogis(item[m] * (x - item[1]))
I=I+L^2*p
}
}
}
I
}
IT=IT+I
}

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

item1<-c(-3.5,-2.0,-0.5,1.0,1)
item2<-c(-3.5,-2.0,-0.5,1.0,2)
item3<-c(-1.0,0.0,1.0,2.0,1)
item4<-c(-1.0,0.0,1.0,2.0,2)
item5<-c(-0.5,1.0,2.5,3.5,1)
item6<-c(-0.5,1.0,2.5,3.5,2)

test=rbind(item1,item2,item3,item4,item5,item6)

fit.sm.pirt(test=test,rangex=c(-6,6),ylim=c(0,3.5),col=1,lty=1,add=F

```

```
,main="")
```

### A.5. Código en R para simular datos provenientes del Modelo Secuencial

```
#####
# Autor : Luis Ángel Mejía Campos #
# # #
# Función: sm.pirt() #
# # #
# Fecha de inicio : 01-12-2017 #
# # #
# Función : Simular respuestas aleatorias basadas en el #
# Modelo Secuencial (2PL-SM) #
# # #
# Asesor : Enver Gerald Tarazona Vargas #
# # #
# Versión : 2.0 #
# # #
# Última Modificación :29-01-2018 #
#####

sm.pirt<- function(n, p, M, param,theta)

{
#Función para simular respuestas aleatorias
#basadas en el modelo secuencial
#Argumentos:
#n=número de patrones de respuesta a simular
#p=número de ítems
#M=número de categorías de los ítems
#param=lista de parámetros
# las primeras columnas corresponden las dificultades
# la última columna corresponde a la discriminación
#theta=vector de habilidades

P=matrix(data=NA,nrow=p,ncol=M) #matriz de probabilidades por categoría
cat=matrix(data=NA,nrow=n,ncol=p) #matrix que almacena las respuestas simuladas

#Simulación de respuestas
for (i in 1:n) {

for (j in 1:p) {
acum=1
```

```

for (k in 1:(M-1)){
P[j,k] = acum*(1-inv.logit( param[j,M]*(theta[i]-param[j,k]) ) )
acum = acum*inv.logit(param[j,M]*(theta[i]-param[j,k]))
}
P[j, M] = acum
}

for (j in 1:p) {
cat[i, j] <- sample(1:M, 1, prob = P[j,])
}
}

cat=as.data.frame(cat)
cat
}

```

#### A.6. Función en R para realizar la estimación bayesiana del modelo secuencial usando MCMC

```

#####
# Autor : Luis Ángel Mejía Campos #
# # #
# Función: smMCMC() #
# # #
# Fecha de inicio : 01-12-2017 #
# # #
# Función : Función en R para realizar la estimación bayesiana usando #
# MCMC en winBUGS #
# # #
# Asesor : Enver Gerald Tarazona Vargas #
# # #
# Versión : 2.0 #
# # #
# Última Modificación :29-01-2018 #
#####

smMCMC <- function (data,K = NULL,
bugs.dir="D:/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 máximo de categorías por ítem
K=apply(Y, 2, max)

#Definición de Prioris

#Discriminación
m.alpha <- 1.0

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

#Dificultad
m.beta <- 0.0

#s.beta <- sqrt(2)
pr.beta <- 0.5

data<- list("Y", "n", "p", "K",
"m.alpha" , "pr.alpha",
"m.beta" , "pr.beta")

#Monitor en BUGS
monitor <- c("alpha","beta","theta")

#Modelo usado en BUGS
bugs.file <- file.path(getwd(), "bugs/sm_2pl.bug")

## Inferencia en WinBUGS
n.burn <- 10000
n.thin <- 20
n.sim <- 1500*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,
debug=FALSE)
}

```

### A.7. Código en BUGS para el Modelo Secuencial Logístico (2PL-SM)

```

#####
# Autor : Luis Ángel Mejía Campos                                     #
#                                                                 #
# Función: sm_2pl.bug                                             #
#                                                                 #
# Fecha de inicio : 01-12-2017                                     #
#                                                                 #
# Función : Función en BUGS para realizar la inferencia del Modelo 2PL-SM #
#                                                                 #
# Asesor : Enver Gerald Tarazona Vargas                           #
#                                                                 #
# Versión : 2.0                                                  #
#                                                                 #
# Última Modificación :29-01-2018                                 #
#####

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)

    for (j in 1:p){
      Feta[i, j, 1] <- 1.0
      for (k in 2:K[j]){
        Feta[i, j, k] <- 1/(1+exp(-alpha[j]*(theta[i]-beta[j, k])))
      }
    }
  }
}

```

```

for (j in 1:p){
  acum[i,j,1] <- 1.0
  for (k in 2:K[j]){
    acum[i,j,k] <- acum[i,j,k-1]*Feta[i,j,k]
  }
}

for (j in 1:p){
  for (k in 1:(K[j]-1)){
    prob[i,j,k] <- acum[i,j,k]*(1-Feta[i,j,k+1])
  }
  prob[i,j,K[j]] <- acum[i,j,K[j]]
}
}

## Prioris para las discriminaciones
for (j in 1:p){
  alpha[j] ~ dnorm(m.alpha, pr.alpha) I(0, )
}

## Prioris para las dificultades
for (j in 1:p){
  for (k in 2:K[j] ){
    beta[j, k] ~ dnorm(m.beta, pr.beta)
    ## kappa[j, k] <- ranked(kappa.star[j, 2:K[j]], k)
  }
}
}

```

#### A.8. Código en R para el estudio de simulación

```

#####
# Autor : Luis Ángel Mejía Campos #
# # #
# Programa: Simulación replicada #
# # #
# Fecha de inicio : 15-12-2017 #
# # #
# Función : Programa en R para realizar el estudio de simulación #
# # #
# Asesor : Enver Gerald Tarazona Vargas #
# # #

```

```

# Versión : 2.0 #
# #
# Última Modificación :14-03-2018 #
#####

library(coda)
library(boot)
library(R2WinBUGS)
library(xtable)
library(taRifx)
library(mcmcplots)
library(truncnorm)
setwd("D:/WinBUGS14/")

#Carga los scripts de funciones auxiliares
source("/sm_mcmc.R") #Inferencia Bayesiana del Modelo secuencial usando MCMC
source("/sm_sim.R") #Simulación de n respuestas ordinales para el Modelo secuencial

#Directorio de WinBUGS
bugs.dir="D:/WinBUGS14/"

#####
#Escenario: Utilizando dificultades y discriminaciones de Tutz #
#####

#Número de individuos
n<-1500

#Número de ítems
p<- 6

#Número de categorías por ítem
M<- 3

#Número de replicas
R<- 20

#Semilla para replicar la simulación
set.seed(222)

#####
#Fijación de los Parámetros del Modelo Secuencial #

```

```
#####

#Fijación de los parámetros de lo ítems

betas=matrix(data=NA,nrow=p,ncol=M-1, byrow=TRUE) #matriz de dificultades

betas[1,1]=-0.527
betas[1,2]=-0.279
betas[2,1]=-0.365
betas[2,2]=-1.448
betas[3,1]=-0.365
betas[3,2]=-0.421
betas[4,1]=-0.692
betas[4,2]= 0.080
betas[5,1]=-0.748
betas[5,2]=0.958
betas[6,1]=-0.979
betas[6,2]=0.233

alfas=matrix(data=NA,nrow=p,ncol=1, byrow=TRUE) #matriz de discriminaciones

alfas[1,1]=1.000
alfas[2,1]=1.000
alfas[3,1]=1.500
alfas[4,1]=1.500
alfas[5,1]=2.000
alfas[6,1]=2.000

param=cbind(betas,alfas)#matriz de parámetros

#Simulación de las habilidades asumiendo distribución normal estandar
theta <- rnorm(n)

#####
#Estimación de los parámetros con el Modelo secuencial para cada réplica #
#####

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

for(s in 1:R){
```

```

#Simulación de las respuestas de n individuos usando el Modelo secuencial
sim<-sm.pirt(n=n,p=p,M=M,param=param,theta=theta)
sim
#Estimación de los parámetros usando MCMC
Y<- as.matrix(sim)
sm.bugs.time<-system.time(sm.bugs <- smMCMC(data=Y))

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

str(results.2plsm)

#Evaluación de la precisión de la simulación

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

#Crea una matriz con los valores ajustados de los parámetros de los ítems
items.fit <- lapply(results.2plsm,
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-param)^2
}
)
items.fit.rmse<-sqrt(Reduce("+",items.fit.rmse)/R)
items.fit.rmse

u.fit.rmse<-lapply(results.2plsm,

```

```

function (x){
(x$model$mean$theta-theta)^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-param)
}
)
items.fit.mae<-Reduce("+",items.fit.mae)/R
items.fit.mae

u.fit.mae<-lapply(results.2plsm,
function (x){
abs(x$model$mean$theta-theta)
}
)
u.fit.mae<-Reduce("+",u.fit.mae)/R
u.fit.mae

#Cálculo de las Correlaciones y Resumen del Modelo

items.fit.cor<-matrix(0,R,M)
u.fit.cor<-matrix(0,R)
model.fit.atime<-0
model.fit.DIC<-0
model.fit.pD<-0
for(s in 1:R){
for(k in 1:M){
items.fit.cor[s,k]<-cor(items.fit[[s]][,k],param[,k])
}
u.fit.cor[s]<-cor(results.2plsm[[s]]$model$mean$theta,theta)
model.fit.atime<-model.fit.atime+results.2plsm[[s]]$time[3]/R
model.fit.DIC<-model.fit.DIC+results.2plsm[[s]]$model$DIC/R
model.fit.pD<-model.fit.pD+results.2plsm[[s]]$model$pD/R
}
items.fit.cor
u.fit.cor

```

```

model.fit.atime
model.fit.DIC
model.fit.pD
model.fit.Dbar<-model.fit.DIC-model.fit.pD
Q<-p*M + n
model.fit.EAIC<-model.fit.Dbar+2*Q
model.fit.EBIC<-model.fit.Dbar+Q*log(n)
# número de datos n*p
#Visualizar Resultados

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

#Tabla de resumen con los resultados del sesgo en la recuperación de parámetros

items.fit.bias<-items.fit.means-param
items.fit.bias<-data.frame(c(1:p),items.fit.bias)
dimnames(items.fit.bias)[[2]]<-c("Ítem","b1","b2","a")
items.fit.bias<-xtable(items.fit.bias,
caption=c("Sesgo en la recuperación de parámetros de los ítems usando estimación
bajo MCMC para el Modelo 2PL-SM"),
digits=3,
label="sim1-biasparitems")
print(items.fit.bias, include.rownames=F)

items.fit.resume<-data.frame(
c("b1","b2","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 cuadratico medio (RSME),
la Media absoluta del error (MAE) y la correlación (r) de los parámetros
estimados en la simulación del Modelo secuencial Logístico (2PL-SM)
considerando R= 20 réplicas"),

```

```

digits=3,
label="sim1-fitresume")
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=p),
item=rep(1:p,2),
rbind(items.fit.rmse,items.fit.mae))
dimnames(items.fit.table)[[2]]<-c("Medida de Precisión","Ítem","b1","b2","b3","a")
items.fit.table<-latex.table.by(items.fit.table,
caption=c("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 Secuencial Logístico (2PL-SM) considerando
R= 20 réplicas"),
digits=3,
label="sim1-fititems")
print(items.fit.table, include.rownames=F,sanitize.text.function = force)

#Tabla con el resumen de las medidas de ajuste de la simulación
model.fit.table<-data.frame(model.fit.DIC,
model.fit.pD,
model.fit.atime,
model.fit.Dbar,
model.fit.EAIC,
model.fit.EBIC)
rownames(model.fit.table)<-NULL
colnames(model.fit.table)<-c("DIC","pD", "Tiempo (seg.)","Dbar","EAIC","EBIC")
model.fit.table<-xtable(model.fit.table,
caption=c("Resultados de medidas de ajuste de la simulación del Modelo
Secuencial Logístico (2PL-SM) considerando
R = 20 réplicas"),
digits=3,
label="sim1-fitmodel")
print(model.fit.table, include.rownames=F)

```

```
#tempdir()
#plot(results.2plgrm[[1]]$model)
#results.2plgrm[[1]]$model
#mcmcplot(results.2plgrm[[1]]$model, random=20)
mcmcplot(results.2plsm[[1]]$model, parms=c("alpha","beta"))
```



## Apéndice B

### Diagnósticos de convergencia

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

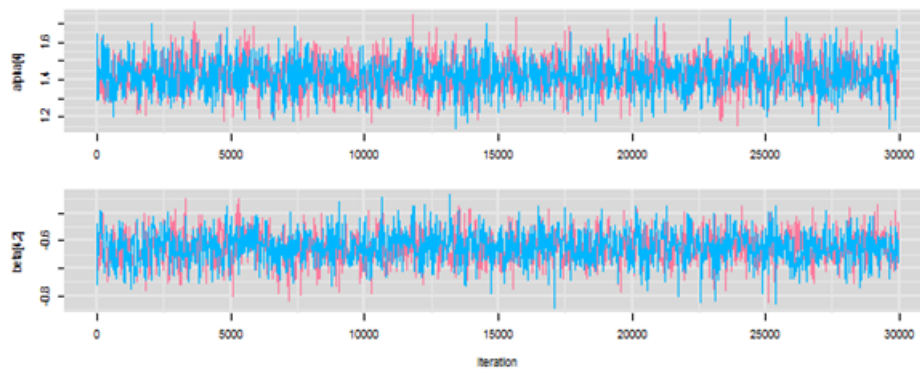


Figura B.1: Histórico de dos cadenas para los parámetros del ítem 4

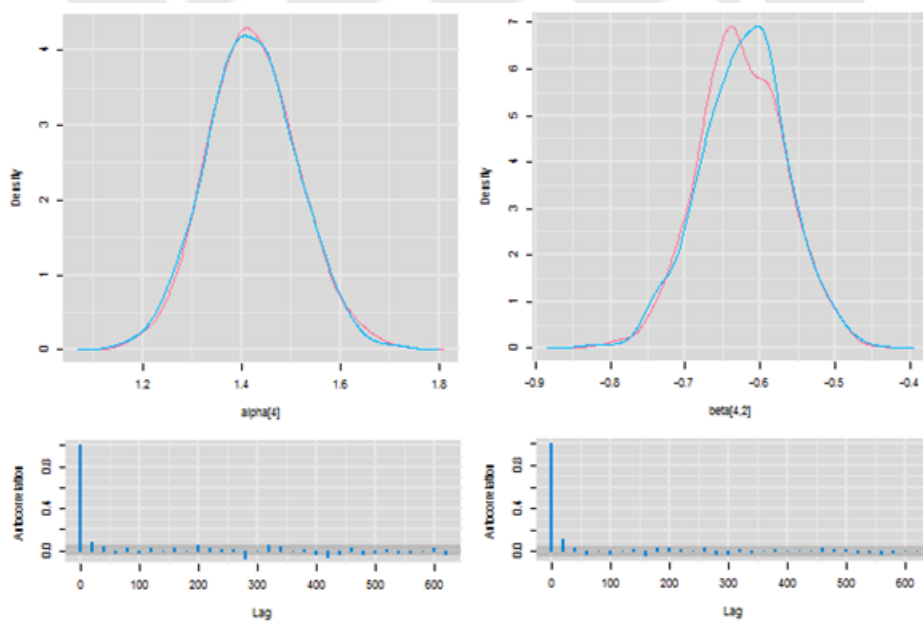


Figura B.2: Gráficas de densidades y autocorrelaciones para los parámetros del ítem 4

B.2. Diagnóstico de convergencia en la aplicación

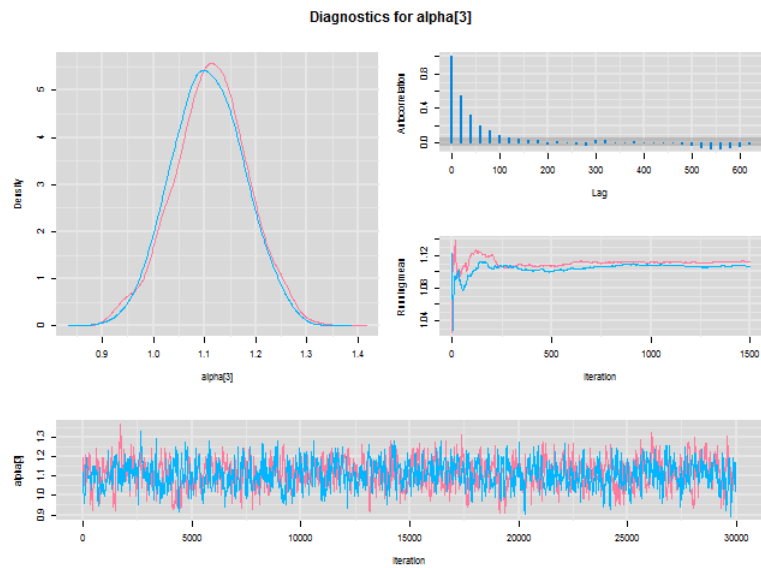


Figura B.3: Histórico de dos cadenas, gráficas de densidades y autocorrelaciones para el parámetro de discriminación del Ítem3, Adecuación al tema, estimado con el Modelo 2PL-SM

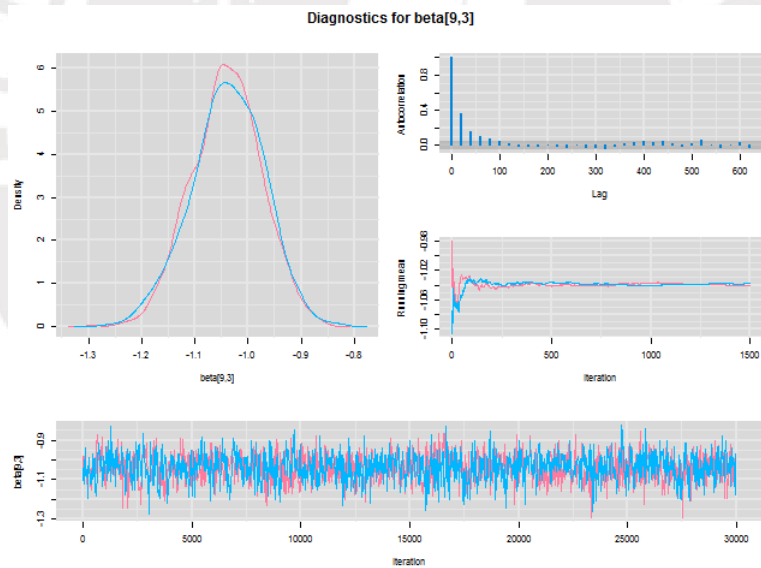


Figura B.4: Histórico de dos cadenas, gráficas de densidades y autocorrelaciones para el segundo umbral del Ítem 9, Referencia, estimado con el Modelo 2PL-SM

## Bibliografía

- Akkermans, L. M. (1998). *Studies on Statistical Models for Polytomously Scored Test Items*, PhD thesis, University of Twente.
- Attorresi, H. F., Lozzia, G. S., Abal, F. J. P., Galibert, M. S. y Aguerri, M. E. (2009). Teoría de respuesta al ítem. conceptos básicos y aplicaciones para la medición de constructos psicológicos, *Revista Argentina de Clínica Psicológica* .
- Bazan, J. L., Mazzon, J. y Hernani, M. (2012). Clasificacao de modelos de resposta ao item policotomico com aplicacao ao marketing, *Revista Brasileira de Estatística* pp. 460–469.
- Curtis, S. M. (2018). *mcmcplots: Create Plots from MCMC Output*. R package version 0.4.3. **URL:** <https://CRAN.R-project.org/package=mcmcplots>
- Deng, S. y Bolt, D. M. (2016). A sequential irt model for multiple-choice items and a multidimensional extension, *Applied Psychological Measurement* **40**(4): 243–257.
- Flores, S. (2012). *Modelos teslet logísticos y logísticos de exponente positivo para pruebas de comprensión de textos*, Master's thesis, Pontificia Universidad Católica del Perú.
- Gelman, A., Carlin, J., Stern, H., Bunson, D., A., V. y Rubin, D. (2014). *Bayesian Data Analysis, Third Edition*, Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis.
- Hambleton, R K. Swaminathan, H. (1985). *Item Response Theory*, 01 edn, Springer Netherlands.
- Hidalgo, M D. French, B. F. (2016). Una introducción didáctica a la teoría de respuesta al ítem para comprender la construcción de escalas, *Revista de Psicología Clínica con Niños y Adolescentes* .
- Hoff, P. D. (2009). *A First Course in Bayesian Statistical Methods*, Springer texts in statistics, 1 edn, Springer-Verlag New York. **URL:** <http://gen.lib.rus.ec/book/index.php?md5=25832249FBBBC256907213CD0025EA0E>
- IEA (2011). ICCS 2009 Technical Report, *Technical report*. **URL:** [http://pub.iea.nl/fileadmin/user\\_upload/Publications/Electronic\\_versions/ICCS\\_2009\\_Technical\\_Report.pdf](http://pub.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICCS_2009_Technical_Report.pdf)
- Lunn, D. J., Thomas, A., Best, N. y Spiegelhalter, D. J. (2007). *WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility*, Hingham, MA, USA. **URL:** <http://dx.doi.org/10.1023/a:1008929526011>
- Masters, G. A. (2016). Handbook of modern item response theory, Vol. Volume One Models, Chapman and Hall/CRC, chapter Partial Credit Model, pp. 109–126. editor Wim J. van der Linden.

- Matteucci, M. y Veldkamp, B. P. (2015). The approach of power priors for ability estimation in irt models, *Quality & Quantity* **49**(3): 917–926.  
**URL:** <https://doi.org/10.1007/s11135-014-0059-y>
- Mellenbergh, G. J. (1995). Conceptual notes on models for discrete polytomous item responses, *Applied Psychological Measurement* **19**(1): 91–100.  
**URL:** <http://dx.doi.org/10.1177/014662169501900110>
- Michael, L. N. y Ostini, R. (eds) (2010). *Handbook of Polytomous Item Response Theory Models*, Vol. 1, Routledge.
- Minedu-UMC (2016). Informe de logros de aprendizaje y sus factores asociados en la.  
**URL:** <http://http://umc.minedu.gob.pe/wp-content/uploads/2016/07/Informe-EM-2013-20-07-1.pdf>
- OECD (2017). PISA 2015 Technical Report, *Technical report*.  
**URL:** <http://www.oecd.org/pisa>
- OREAL (2016). Reporte Técnico TERCE, *Technical report*.  
**URL:** <http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/pdf/Reporte-tecnico-TERCE.pdf>
- Ostini, R. y Nering, M. (2006). *Polytomous Item Response Theory Models*, number no. 144 in *Online access: Sage Sage Research Methods Online*, SAGE Publications.  
**URL:** <https://books.google.com.pe/books?id=wS8VEMtJ3UYC>
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.  
**URL:** <https://www.R-project.org/>
- Reckase, M. D. (2009). *Multidimensional Item Response Theory*, 01 edn, Springer-Verlag New York.
- 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>
- Swaminathan, H. (2014). *Bayesian Item Response Theory Estimation*, John Wiley and Sons, Ltd.  
**URL:** <http://dx.doi.org/10.1002/9781118445112.stat06132>
- Tarazona, E. (2013). *Modelos alternativos de respuesta graduada con aplicaciones en la calidad de servicios*, Master's thesis, Pontificia Universidad Católica del Perú.
- Thissen, D. y Cai, L. (2016). Handbook of modern item response theory, Vol. Volume One Models, Chapman and Hall/CRC, chapter Nominal Categories Models, pp. 51–71. editor Wim J. van der Linden.
- Tutz, G. (1990). Sequential item response models with an ordered response, *British Journal of Mathematical and Statistical Psychology* **43**(1): 39–55.  
**URL:** <http://dx.doi.org/10.1111/j.2044-8317.1990.tb00925.x>
- Tutz, G. (2016). Handbook of modern item response theory, Vol. Volume One Models, Chapman and Hall/CRC, chapter Sequential Models for Ordered Responses, pp. 139–151. editor Wim J. van der Linden.

- Van Der Ark, L. A. (2001). Relationships and properties of polytomous item response theory models, *Applied psychological Measurement* **25**: 273–82.
- Welchowski, T. (2014). *Person parameter estimation in the polytomous rasch model*, Master's thesis, Ludwig Maximilians University Munich.

