Modelo TRI Logístico Para Un Trazo Latente Que Sigue Una Distribución Asimétrica

Leidy Natalia León Carvajal
Matemático

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Estadística
Bogotá, D.C.
Mayo, 2017

Modelo TRI Logístico Para Un Trazo Latente Que Sigue Una Distribución Asimétrica

Leidy Natalia León Carvajal
Matemático

Disertación presentada para optar al título de
Magíster en Ciencias Estadística

Director
Álvaro Mauricio Montenegro Díaz, Ph.D.

Línea de investigación
Teoría de Respuesta al Ítem

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Estadística
Bogotá, D.C.
Mayo, 2017

Título en español

Modelo TRI Logístico Para Un Trazo Latente Que Sigue Una Distribución Asimétrica

Title in English

Logistic IRT Model for a Latent Trait Following an Asymmetric Distribution

Resumen: Al revisar los procedimientos de estimación implementados en la Teoría de Respuesta al Ítem es común encontrarse con la suposición de que el trazo latente subyacente a un test dicotómico sigue una distribución normal estándar, sin embargo, variables latentes que describen conductas o rasgos de personalidad de un conjunto de individuos presentan comportamientos asimétricos. Por esta razón, en este documento se propone una metodología de estimación en la que el trazo latente sigue una distribución Gamma de parámetros α y β. La estimación se basa en la propuesta de Bock & Aitkin (1981), en la que se obtienen estimaciones marginales para los parámetros de los ítems con el algoritmo EM y, posteriormente, se estima el trazo latente usando máxima verosimilitud. La metodología se implementa en un test simulado con trazo asimétrico y en uno de los constructos del Inventario PIHEMA R-2 aplicado durante el proceso de admisión a la Facultad de Humanidades y Educación de la Universidad Central de Venezuela durante 2006, al comparar las estimaciones con las obtenidas mediante el paquete `ltm` del software estadístico `R`, se observa que la suposición de un trazo asimétrico proporciona un mejor ajuste.

Abstract: In reviewing the estimation procedures implemented in the Item Response Theory, we usually find the assumption that the latent trait underlying a dichotomous test follows a standard normal distribution, however, latent variables that describe behaviors or personality traits in a Set of individuals, present asymmetric distributions. For this reason, this paper proposes an estimation methodology with a latent trait following a Gamma distribution of parameters α and β. The estimation is based on the proposal of Bock & Aitkin (1981), where we obtain item parameter's marginal estimates with the EM and, later, we estimate the latent trait using maximum likelihood.This methodology is implemented in a simulated test with asymmetric latent trait and in one of the constructs of the PIHEMA R-2 Inventory applied during the admission process to the Faculty of Humanities and Education of the Central University at Venezuela in 2006, when we compare the estimates with those obtained by the `ltm` package of the statistical software `R`, we note that the assumption of an asymmetric trait provides a better fit.

Palabras clave: Modelo logístico, algoritmo EM, trazo latente, distribución gamma, cuadratura, estimación marginal.

Keywords: Logistic model, EM algorithm, latent trait, gamma distribution, quadrature, marginal estimate.

Dedicado a

A mi familia y Juan Pablo por su apoyo y paciencia.

Agradecimientos

Siempre que se gesta una idea, un nuevo sueño, aparecen inherentemente obstáculos que en principio parecen insuperables. Es cuestión de decisión, de asumir un riesgo y llegar hasta el final. Con miedos, con incertidumbre, y muchas veces sin mayores motivos, pero avanzamos, seguimos luchando por alcanzar la meta, conservamos la esperanza de que cada esfuerzo valdrá la pena, de que cada mañana traerá consigo una nueva oportunidad. Quiero agradecer a mis padres por su apoyo constante y paciencia insuperable durante este recorrido, agradezco también a mis profesores, especialmente a mi director Álvaro Mauricio Montenegro por su acompañamiento y guía.

Agradezco además al grupo de investigación **SICS Research Group** del Departamento de Estadística dentro del cual se desarrolló esta tesis en el marco del proyecto "*Multidimensional item response theory models for practical application in large tests designed to measure multiple constructs*", código Colciencias 110156933802.

Índice general

Índice general **I**

Índice de tablas **III**

Índice de figuras **V**

Introducción **VII**

1. Marco Teórico **1**

1.1. Teoría de Respuesta al Ítem 1

1.1.1. Modelo Logístico de Dos Parámetros (2PL) 2

1.2. Métodos de Estimación para los Modelos TRI Logísticos 3

1.2.1. Estimación Máximo–Verosímil Conjunta (JMLE) 5

1.2.2. Estimación Máximo–Verosímil Marginal (MMLE) 5

1.2.3. Estimación Máximo–Verosímil Marginal Implementando el Algoritmo EM (MMLE–EM) 7

1.2.4. Estimación Bayesiana 9

2. Modelo 2PL con Trazo Latente Asimétrico **10**

2.1. Modelo Propuesto 10

2.2. Estimación 11

2.2.1. Estimación de los Parámetros de los Ítems 11

2.2.2. Estimación del Trazo Latente 16

2.3. Pruebas de Bondad de Ajuste 17

2.4. Implementación 18

2.4.1. Análisis Preliminar de los Datos 18

2.4.2. Estimación 19

2.4.3. Análisis gráfico 19

3. Simulación **20**

3.1. Generación del Test 20

3.2. Recuperación de los Parámetros 21

3.3. Valores Simulados Vs. Recuperados 22

3.4. Pruebas de Bondad de Ajuste 24

4. Aplicación a Datos Reales **27**

4.1. Descripción de los Datos 27

4.2. Selección del Constructo y Análisis Preliminar 28

4.3. Estimación 32

4.3.1. Estimación de los Parámetros de los Ítems 32

4.3.2. Estimación del Trazo 33

4.4. Comparación con otros Paquetes 34

A. Código R **38**

A.1. Estimación Parámetros de los Ítems 38

A.2. Estimación del Trazo Latente 40

B. Resultados Simulación **42**

B.1. Análisis Preliminar 42

B.2. Recuperación 43

B.3. Bondad de Ajuste 45

C. Resultados Datos Reales **49**

C.1. Selección de Datos 49

C.2. Estimación 51

Discusión y Conclusiones **54**

Trabajo Futuro **56**

Bibliografía **57**

Índice de tablas

3.1. Medidas Descriptivas de los Parámetros Simulados 20

3.2. Valores simulados y recuperados de los parámetros de discriminación y dificultad 23

3.3. Resumen Descriptivo de $\widehat{\theta}_i$ 23

3.4. Medidas Descriptivas para $|\theta_i - \widehat{\theta}_i|$ 24

3.5. Medidas Descriptivas de los Errores Absolutos 24

3.6. Correlaciones entre frecuencias esperadas y observadas para cada uno de los ítems del test 26

4.1. Posibles Distribuciones para Modelar el Trazo Latente 31

4.2. Estadísticas de resumen para el modelo GAMLSS ajustado con la distribución Gamma 32

4.3. Estimaciones y errores estándar para los parámetros de los ítems 33

4.4. Estadísticas de resumen para $\widehat{\theta}_i$ obtenido con máxima verosimilitud 34

4.5. Estadísticas de resumen para $\widehat{\theta}_i$ obtenido con EAP 34

4.6. Estadísticas de resumen para $SE(\widehat{\theta}_i)$ con MV 34

4.7. Estadísticas de resumen para $SE(\widehat{\theta}_i)$ con EAP 34

4.8. Estimaciones de los parámetros de los ítems con la metodología propuesta y usando el paquete ltm 35

4.9. Resumen estadístico para el trazo estimado con MV y ltm 35

B.1. Pesos y nodos del procedimiento de cuadratura 43

B.2. Valores iniciales para los parámetros de los ítems 43

B.3. Resumen Descriptivo del primer componente del ACP 44

B.4. Resumen Descriptivo del Trazo Recuperado con EAP 44

B.5. Valores del estadístico calculado y p-valores para la prueba de Bondad de Ajuste 48

C.1. Número de Ítems por Constructo 49
C.2. Pesos y nodos del procedimiento de cuadratura 51
C.3. Valores iniciales para la estimación de los parámetros de los ítems 52

Índice de figuras

1.1. CCI con distintos valores de los parámetros de dificultad y discriminación . 3

1.2. Cuadratura Gaussiana 6

2.1. Nodos de Cuadratura con Distribución de Pesos Gamma 13

2.2. Frecuencias esperadas Vs. frecuencias observadas 18

3.1. Función de Verosimilitud para algunos individuos seleccionados aleatoriamente 21

3.2. Valores simulados de los parámetros de discriminación Vs. los valores recuperados 22

3.3. Valores simulados de los parámetros de dificultad Vs. los valores recuperados 22

3.4. Valores simulados del trazo Vs. valores obtenidos en la estimación 24

3.5. Frecuencias Esperadas Vs. Observadas para algunos ítems seleccionados aleatoriamente 25

3.6. Residuales Estandarizados para algunos ítems 25

3.7. Frecuencias Esperadas Vs. Observadas en el Test 26

4.1. Scores por Constructo 28

4.2. Proporción de varianza representada por autovalor en cada componente . . . 29

4.3. Curva de Cronbach-Mesbah 29

4.4. Proporción de varianza representada por autovalor en cada componente . . . 30

4.5. Curva de Cronbach-Mesbah 30

4.6. Gráficas de Diagnóstico para el modelo GAMLSS ajustado con la distribución Gamma 32

4.7. (a) Densidad $\widehat{a}_j$ (b) Densidad $\widehat{b}_j$ 33

4.8. Comparación ltm vs. metodología propuesta 36

4.9. Frecuencias esperadas Vs. observadas del test 36

B.1. Densidades de los parámetros simulados 42
B.2. (a) Histograma Trazo Simulado (b) Histograma Score del Test 43
B.3. Densidades de los parámetros simulados 44
B.4. Curvas Características del Test y de los Ítems respectivamente 45
B.5. Frecuencias esperadas vs. observadas: la gráfica del lado derecho corresponde al trazo obtenido con MV, la gráfica izquierda al trazo recuperado con EAP. 45
B.6. Frecuencias esperadas vs. observadas 46
B.7. Residuales estandarizados para cada uno de los ítems del test 47

C.1. Box-Plot para los scores de cada constructo. 50
C.2. (a) Resultados ACP Test Completo (b) Resultados ACP Test sin el Ítem 12 50
C.3. Densidades Primer Componente y Distribución Gamma 51
C.4. Curvas características del test y de los ítems 52
C.5. Frecuencias Esperadas Vs. Observadas 53

Introducción

¿Cómo medir el nivel de satisfacción de un individuo frente a una determinada circunstancia? ¿Cómo cuantificar la habilidad de un sujeto en un área específica? Son numerosas las características que no son observables de manera directa y para las que resulta indispensable crear una escala o métrica que facilite su interpretación. Tales variables se denominan *trazos o variables latentes* y su caracterización constituye un objetivo fundamental de la Teoría de Respuesta al Ítem (TRI), en la que se han formulado diversos modelos con la finalidad de explicar la relación entre la habilidad de un individuo y la probabilidad de que responda de manera correcta a un ítem de una prueba determinada.

Dentro de los modelos TRI más empleados, se encuentran aquellos en los que los ítems tienen dos categorías de respuesta (test dicotómicos), Rasch (1960) y Birnbaum (1968) por ejemplo, propusieron modelos en los que la probabilidad de acertar un ítem del test está dada por la distribución logística y en los cuales los parámetros de los ítems caracterizan completamente dicha probabilidad.

Pese a que el trazo latente es una variable no observable de manera directa, resulta necesario crear una escala que permita ubicar a los individuos y que esté relacionada con la probabilidad de responder correctamente a un ítem de un test determinado. En medio de los esfuerzos por explicar el comportamiento del trazo de un individuo, ha surgido, como un criterio generalizado, la presunción de que una distribución normal estándar es el mejor modelo para describirlo; si bien es cierto que tal suposición no constituye una base inquebrantable de la TRI, son escasos los esfuerzos por incluir otras distribuciones en los procesos de estimación.

Woods (2008) y Páyer (2013) por ejemplo, hacen alusión al hecho de que variables como los rasgos de personalidad o de conducta de los individuos, no siguen un comportamiento normal estándar y, por el contrario, presentan características de asimetría.

El presente trabajo tiene como finalidad, realizar el proceso de estimación en un modelo TRI logístico de dos parámetros en el que la variable latente sigue una distribución de este etilo, específicamente, una distribución Gamma.

Los resultados obtenidos se aplicarán posteriormente a un conjunto de datos reales que corresponden al Inventario PIHEMA R-2 aplicado durante el proceso de admisión a la Facultad de Humanidades y Educación de la Universidad Central de Venezuela durante 2006, que contiene además de variables cognitivas y socioeconómicas, variables relacionadas con la personalidad de los aspirantes.

CAPÍTULO 1

Marco Teórico

1.1. Teoría de Respuesta al Ítem

La Teoría de Respuesta al Ítem, TRI, surge como una alternativa para superar las limitaciones de la Teoría Clásica del Test (TCT), en la que la confiabilidad del instrumento de medición se ve influenciada por las características de los individuos y el contexto de aplicación.

Pese a trabajos previos, es a partir de las publicaciones de Lord (1952), Rasch (1960) y Birnbaum (1968) que la TRI ha tenido una mayor difusión y con base en los mismos, se han establecido los conceptos y supuestos básicos para estructurarla y dar lugar así a incontables modelos en función del tipo de ítems en el test, de las características del trazo latente o del número de poblaciones consideradas, entre otros.

Los principales supuestos a considerar en un modelo TRI son los siguientes:

- **Unidimensionalidad:** Se hablará de un test unidimensional si los ítems de la prueba aportan información sobre un único trazo latente.
- **Independencia Condicional:** Como se define en Hambleton et al. (1991), la respuesta del individuo a un ítem del test es independiente de su respuesta a cualquier otro ítem.
- **Independencia entre individuos:** La respuesta de un evaluado al test es independiente de la respuesta de cualquier otro individuo.

Pese a la diversidad de instrumentos que pueden analizarse por medio de la TRI, el presente documento se enfoca en aquellos cuyos ítems solo tienen dos categorías de respuesta (test dicotómicos), aplicados a una sola población y en los que el trazo latente es unidimensional.

En particular, se considerará el modelo en el que la relación entre la respuesta a un ítem y el trazo latente está dada por una función logística de dos parámetros.

1.1.1. Modelo Logístico de Dos Parámetros (2PL)

En un test dicotómico con p ítems aplicado a N individuos, la probabilidad que el individuo i responda correctamente al ítem j puede describirse mediante,

$$P(Y_{ij} = 1|\theta_i) = \frac{e^{a_j(\theta_i - b_j)}}{1 + e^{a_j(\theta_i - b_j)}} = \frac{1}{1 + e^{-a_j(\theta_i - b_j)}} \tag{1.1}$$

para $i = 1, 2, ..., N$, $j = 1, 2, ..., p$, y donde

- Y_{ij}: es igual a 1 si el individuo i acierta al ítem j, y es 0 en caso contrario.
- θ_i: es el trazo latente del individuo i.
- a_j: es el parámetro de discriminación del ítem j.
- b_j: es el parámetro de dificultad del ítem j.

Como puede observarse en la Ecuación (1.1), además del trazo latente de cada individuo aparecen dos parámetros que caracterizan cada uno de los ítems:

- **Parámetro de Discriminación (a_j)**: indica qué tanto cambia la probabilidad de responder acertadamente el ítem j a medida que varía el nivel del trazo latente. Cuando $a_j = 1$, el modelo obtenido se denomina *Modelo 1PL o Modelo de Rasch.*
- **Parámetro de Dificultad (b_j)**: como se indica en Hambleton et al. (1991) y Attorresi et al. (2009), el parámetro de dificultad es el punto sobre la escala del trazo latente donde la probabilidad de responder correctamente el ítem j es igual a 1/2.

Cada uno de los ítems de un test queda completamente caracterizado por su parámetro de discriminación y de dificultad, y su relación con la probabilidad de acierto de acuerdo con el nivel trazo latente, puede describirse mediante la *curva característica del ítem (CCI).* En la Gráfica 1.1 se presentan las curvas características para un ítem con distintos valores de a_j y b_j.

Como se observa en la Gráfica 1.1, el parámetro a_j puede interpretarse como la pendiente de la CCI en la abscisa b_j, de manera que permite clasificar los individuos en distintos niveles del trazo latente teniendo en cuenta que entre más alto sea el valor del parámetro, menor será la discriminación. Nótese además, que el parámetro de dificultad b_j ubica la CCI con relación a la escala del trazo. Es decir, ítems con valores mayores de b_j indican que requieren un mayor nivel del trazo para ser respondidos de manera correcta.

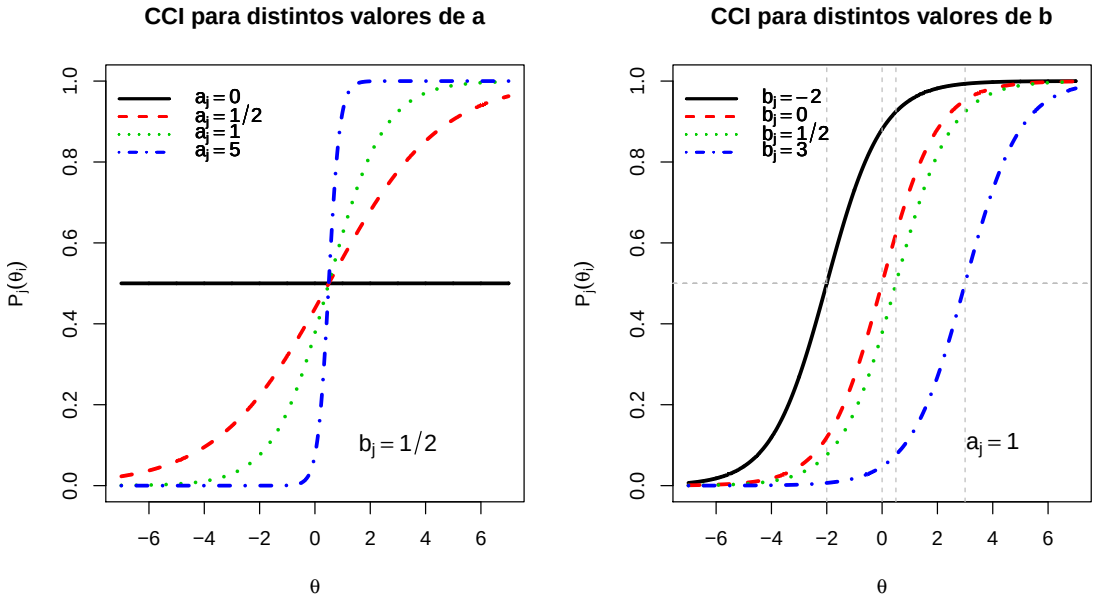


FIGURA 1.1. CCI con distintos valores de los parámetros de dificultad y discriminación

1.2. Métodos de Estimación para los Modelos TRI Logísticos

En los modelos de la TRI, la probabilidad de que un individuo acierte un ítem en un test determinado está caracterizada tanto por el nivel del trazo latente como por los parámetros del ítem, en el caso del modelo 2PL, los parámetros de discriminación y dificultad. Andrade et al. (2000) refiere tres posibles escenarios de estimación:

- Se conocen los parámetros de los ítems y solo es necesario estimar el trazo.
- Se conocen el trazo y se desean estimar los parámetros de los ítems.
- Se necesita estimar tanto el trazo como los parámetros de los ítems.

Esta última situación es la que generalmente se presenta, y es la que se abordará en este documento. Para estimar los parámetros de los ítems y el trazo latente existen a su vez dos alternativas: estimarlos de manera conjunta o llevar a cabo la estimación en dos etapas.

Como se planteó en secciones anteriores, se considera un test dicotómico de p ítems aplicado a N individuos, donde θ_i representa el trazo del individuo i y la probabilidad de respuesta correcta, $P(\theta)$, está dada por el modelo 2PL de la Ecuación (1.1), de modo que la probabilidad de errar la respuesta es $Q(\theta) = 1 - P(\theta)$. Así mismo, sea y_{ij} la variable que indica la respuesta del evaluado i al ítem j, de manera que

$$y_{ij} = \begin{cases} 1, & \text{El inividuo } i \text{ acierta el ítem } j \\ 0, & \text{El inividuo } i \text{ falla el ítem } j \end{cases}$$

Bajo el supuesto de independencia local de los ítems dado θ_i, la probabilidad conjunta de los ítems para el individuo i queda determinada por

$$P(\boldsymbol{Y}_i|\theta_i; a_j, b_j) = \prod_{j=1}^{p} P_j(\theta_i)^{y_{ij}} Q_j(\theta_i)^{1-y_{ij}} \tag{1.2}$$

donde $\boldsymbol{Y}_i = (y_{i1}, y_{i2}, ..., y_{ip})$ es el vector de respuestas del individuo i a los p ítems del test. Si además se asume la independencia entre individuos, la función de verosimilitud será entonces

$$L = \prod_{i=1}^{N} \prod_{j=1}^{p} P_j(\theta_i)^{y_{ij}} Q_j(\theta_i)^{1-y_{ij}} \tag{1.3}$$

De manera que, el objetivo es determinar los valores de los parámetros que maximicen la expresión (1.3) o bien la función de log-verosimilitud que está dada por

$$\mathcal{L} = \log(L) = \sum_{i=1}^{N} \sum_{j=1}^{p} [y_{ij} \log P_j(\theta_i) + (1 - y_{ij}) \log Q_j(\theta_i)] \tag{1.4}$$

Una vez realizado el proceso de derivación (el procedimiento detallado puede consultarse en Andrade et al. (2000)), las ecuaciones de estimación resultantes son

- Para a_j: $\sum_{i=1}^{N} [y_{ij} P_j(\theta_i)](\theta_i - b_j) w_{ij} = 0$
- Para b_j: $-a_j \sum_{i=1}^{N} [y_{ij} P_j(\theta_i)] w_{ij} = 0$

donde

$$w_{ij} = \frac{P_j^*(\theta_i) Q_j^*(\theta_i)}{P_j(\theta_i) Q_j(\theta_i)} \qquad P_j^*(\theta_i) = \frac{e^{-a_j(\theta_i - b_j)}}{1 + e^{-a_j(\theta_i - b_j)}}$$

Como alternativa, Harwell et al. (1988) propone agrupar el trazo latente por niveles: se consideran q grupos de n_{kj} individuos con nivel de trazo θ_k donde, además, r_{kj} representará el número de evaluados del grupo k que respondieron acertadamente al ítem j. Nuevamente, asumiendo independencia entre los individuos, r_{kj} sigue una distribución binomial de parámetros n_{kj} y $P_j(\theta_k)$, de manera que la función de verosimilitud ahora estará determinada por

$$L = \prod_{k=1}^{q} \prod_{j=1}^{p} \binom{n_{jk}}{r_{jk}} P_j(\theta_k)^{r_{kj}} Q_j(\theta_k)^{n_{kj} - r_{kj}} \tag{1.5}$$

y la función de log-verosimilitud será

$$\mathcal{L} = \log(L) = \sum_{k=1}^{q}\sum_{j=1}^{p}\left[r_{kj}\log P_j(\theta_k) + (n_{kj} - r_{kj})\log Q_j(\theta_k)\right] + \sum_{k=1}^{q}\sum_{j=1}^{p}\log\binom{n_{jk}}{r_{jk}} \quad (1.6)$$

Cabe notar que el segundo término del lado derecho de la ecuación (1.6) es una constante respecto a los parámetros de los ítems. De esta manera, las ecuaciones de la estimación máximo-verosímil son

- Para a_j: $\sum_{k=1}^{q}(\theta_k - b_j)\left[r_{kj} - n_{kj}P_j(\theta_k)\right]w_{kj} = 0$
- Para b_j: $-a_j\sum_{k=1}^{q}\left[r_{kj} - n_{kj}P_j(\theta_k)\right]w_{kj} = 0$

1.2.1. Estimación Máximo–Verosímil Conjunta (JMLE)

Birnbaum (1968) propone un método de estimación de máxima verosimilitud conjunta (JMLE) en la que, bajo el supuesto de independencia local, se estiman primero los parámetros de los ítems suponiendo conocido el trazo para cada individuo y posteriormente, asumiendo independencia entre individuos, se utilizan las estimaciones obtenidas para encontrar el trazo latente, esta metodología se repite hasta que el procedimiento converge.

Para realizar las estimaciones en cada paso, se implementa el algoritmo de Newton-Raphson, sin embargo, solo ha podido demostrarse la consistencia de los estimadores para el modelo de Rasch, (Haberman, 1977). Como limitación adicional, al maximizar la Ecuación (1.6) de manera simultánea para los parámetros de los ítems y el trazo latente, puede ocurrir que exista más de un valor máximo, es decir, el modelo no es identificable.

Generalmente, la metodología JMLE suele asociarse con un modelo de ANOVA de efectos fijos.

1.2.2. Estimación Máximo–Verosímil Marginal (MMLE)

Con la finalidad de superar las limitaciones de la metodología JMLE, Bock & Lieberman (1970) proponen una metodología de estimación máximo–verosímil marginal (MMLE) que produce estimadores consistentes, pero presenta inconvenientes a la hora de implementarse computacionalmente cuando se trabaja con test que contengan una gran cantidad de ítems.

Esta metodología supone que los individuos constituyen una muestra aleatoria en la que el trazo sigue una distribución determinada[1], $g(\theta|\epsilon)$, por lo que el método MMLE suele asociarse con un modelo de ANOVA de efectos mixtos en el que el trazo es considerado el efecto aleatorio y los ítems los efectos fijos.

Para la estimación MMLE, la probabilidad marginal de $\boldsymbol{Y_i}$ está determinada por

[1]Donde ϵ representa el vector de parámetros de la distribución

$$P(\boldsymbol{Y_i}|a_j, b_j) = \int P(\boldsymbol{Y_i}|\theta, a_j, b_j;)g(\theta_i|\epsilon)d\theta \tag{1.7}$$

De manera que, bajo la suposición de independencia entre individuos, la función de verosimilitud estará dada por

$$L = \prod_{i=1}^{N} P(\boldsymbol{Y_i}|a_j, b_j) \tag{1.8}$$

Las ecuaciones (1.7), (1.8) y el procedimiento para la obtención de las ecuaciones de máxima verosimilitud, darán como resultado expresiones en las que aparecen involucradas integrales que pueden generar inconvenientes a la hora de implementar los algoritmos computacionalmente. Como alternativa para aproximar las integrales resultantes, se emplean métodos numéricos de cuadratura.

En este caso en particular, se considera que la distribución del trazo latente, $g(\theta|\epsilon)$, es continua y tiene momentos finitos, de hecho, se asume usualmente que sigue una distribución normal estándar (*cuadratura gaussiana*). Luego de hallar una partición del soporte de la distribución, se busca aproximar el área bajo la curva normal mediante la suma de áreas de rectángulos cuyas bases son los intervalos de la partición; los puntos medios de los q intervalos, denotados por X_k, se denominan *nodos de cuadratura*, y los valores que relacionan cada nodo con la distribución normal se conocen como *pesos de cuadratura* y se denotan por $A(X_k)$.

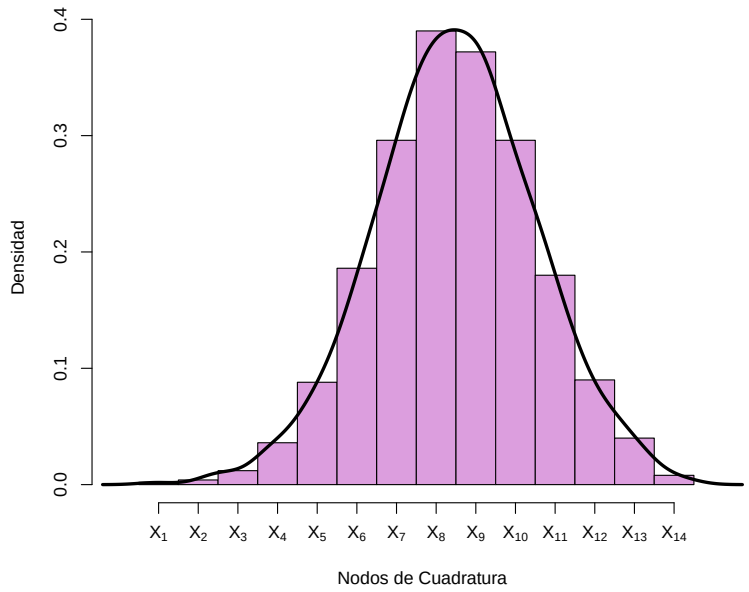


FIGURA 1.2. Cuadratura Gaussiana

Existen diversos métodos para obtener tanto los nodos como los pesos de la cuadratura gaussiana, Stroud & Secrest (1966), por ejemplo, construyeron tablas para distintas cantidades de puntos de cuadratura, sin embargo, como lo sugieren Bock & Lieberman

(1970), bajo la suposición de normalidad del trazo los valores de X_k de la tabla deben multiplicarse por $\sqrt{2}$ y los pesos deben dividirse por $\sqrt{\pi}$. En el software `R`, las funciones `gauss.quad` y `gauss.quad.prob` del paquete `statmod` también calculan los pesos y nodos de cuadratura empleando varios métodos y distribuciones (Smyth et al., 2017).

1.2.3. Estimación Máximo–Verosímil Marginal Implementando el Algoritmo EM (MMLE–EM)

Para resolver las limitaciones que computacionalmente presenta la metodología MMLE, Bock & Aitkin (1981) proponen un procedimiento de estimación de máxima verosimilitud marginal en el que se asume que el trazo sigue una distribución normal estándar con función de densidad $g(\theta_i|\epsilon)$, de manera que la probabilidad marginal de $\boldsymbol{Y_i}$ está dada por la ecuación (1.7), en tanto que la probabilidad posterior de cada individuo queda determinada por

$$P(\theta_i|\boldsymbol{Y_i}) = \frac{P(\boldsymbol{Y_i}|\theta_i)g(\theta_i|\epsilon)}{\int P(\boldsymbol{Y_i}|\theta_i)g(\theta_i|\epsilon)d\theta} \tag{1.9}$$

En Harwell et al. (1988) se definen

$$P(X_k|\boldsymbol{Y_i}) = \frac{\prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} A(X_k)}{\sum_{k=1}^{q}\prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} A(X_k)} \tag{1.10}$$

$$\overline{n}_{kj} = \sum_{i=1}^{N} P(X_k|\boldsymbol{Y_i}) = \sum_{i=1}^{n} \frac{\prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} A(X_k)}{\sum_{k=1}^{q}\prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} A(X_k)} \tag{1.11}$$

$$\overline{r}_{kj} = \sum_{i=1}^{N} y_{ij} P(X_k|\boldsymbol{Y_i}) = \sum_{i=1}^{n} y_{ij} \frac{\prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} A(X_k)}{\sum_{k=1}^{q}\prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} A(X_k)} \tag{1.12}$$

que representan respectivamente, el número esperado de individuos en el nivel trazo X_k y el número esperado de respuestas correctas al ítem j en el nivel del trazo X_k. Bock & Aitkin (1981) también definen

$$L(X_k) = \prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} \tag{1.13}$$

con lo que las ecuaciones (1.11) y (1.12) pueden reescribirse como

$$\overline{n}_{kj} = \sum_{i=1}^{N} \frac{L(X_k)A(X_k)}{\sum_{k=1}^{q} L(X_k)A(X_k)} \qquad \overline{r}_{kj} = \sum_{i=1}^{N} \frac{y_{ij}L(X_k)A(X_k)}{\sum_{k=1}^{q} L(X_k)A(X_k)} \tag{1.14}$$

Además, la probabilidad posterior dada en la ecuación (1.9) puede escribirse como

$$P(X_k|\boldsymbol{Y_i}) = \frac{L(X_k)A(X_k)}{\sum_{k=1}^{q} L(X_k)A(X_k)} \tag{1.15}$$

Así pues, el procedimiento de estimación de Bock & Aitkin (1981) puede resumirse en 4 pasos fundamentales como se indica en Harwell et al. (1988):

1. Después de obtener los nodos y pesos de cuadratura, se emplea $P_j(X_k)$ y una estimación inicial de los parámetros de los ítems para calcular $L(X_k)$ por medio de la ecuación (1.13).
2. Calcular la probabilidad posterior del trazo de cada individuo mediante la ecuación (1.15).
3. Calcular $\overline{n}_{kj}$ y $\overline{r}_{kj}$ empleando las expresiones dadas en (1.14).
4. Resolver las ecuaciones de maximización utilizando $\overline{n}_{kj}$ y $\overline{r}_{kj}$.

A diferencia del procedimiento propuesto por Bock & Lieberman (1970), los estimadores resultantes en la metodología de Bock & Aitkin (1981) son consistentes, no importa cuánto aumente el tamaño de la muestra, el número de parámetros a estimar es siempre el mismo.

En Dempster et al. (1977) se define el algoritmo EM (*Expectation-Maximization*) como un procedimiento iterativo que permite encontrar estimadores máximo–verosímiles cuando se tienen variables aleatorias latentes, que en este caso es el trazo latente de cada individuo.

En el algoritmo EM se utilizan las respuestas de los individuos a los ítems para hacer inferencias sobre el trazo latente, donde $(\boldsymbol{Y}, \theta)$ representan los datos no observables (completos), y $\boldsymbol{Y}$ representa los datos observados (incompletos) (Orchard & Woodbury, 1972). Sea $f(\boldsymbol{Y}, \boldsymbol{\theta}|\boldsymbol{\beta})$[2] la función de densidad conjunta de los datos completos, una vez obtenida una estimación inicial de los parámetros de los ítems, el procedimiento de estimación se lleva a cabo en dos etapas:

1. **Paso E (*Expectation*):** Se obtienen los valores esperados de la función de log-verosimilitud, es decir, se calcula $E\left[\log f(\boldsymbol{Y}, \boldsymbol{\theta}|\boldsymbol{\beta})|\boldsymbol{Y}; \boldsymbol{\beta}^{(k)}\right]$. Para llevar a cabo el paso E, se distribuyen las respuestas y los scores a lo largo de la escala del trazo latente, para luego obtener el número de individuos esperado n_{kj}, y el número de respuestas acertadas r_{kj}, en cada nivel del trazo que, previamente, ha sido muestreado utilizando el procedimiento de cuadratura mencionado en secciones anteriores.

[2] Aquí $\boldsymbol{\beta}$ representa el vector de parámetros a estimar

2. **Paso M (*Maximization*):** Considerando a la función obtenida en el paso E como una función de verosimilitud, se determina $\boldsymbol{\beta}^{(k+1)}$de manera que la maximice.

Una vez estimados los parámetros de los ítems, pueden emplearse los valores obtenidos para estimar el trazo de los individuos a través, por ejemplo, del método de Newton–Raphson.

1.2.4. Estimación Bayesiana

Existen varias metodologías propuestas desde la perspectiva bayesiana para estimar los parámetros en un modelo TRI logístico. Mislevy (1986) propone, por ejemplo, un procedimiento de estimación bayesiana marginal que generaliza los aportes de Bock & Aitkin (1981). En este tipo de procesos se incorporan distribuciones *a priori* para los parámetros que se desean estimar y se construye una distribución *a posteriori* que resulta proporcional al producto entre la función de verosimilitud y la distribución a priori; las estimaciones se obtienen con base en alguna característica de la distribución posteriori como la media o la moda.

Dentro de las propuestas bayesianas también se encuentran propuestas de estimación conjunta que, en principio, fueron formuladas por Swaminathan & Gifford (1986), y en las que las estimaciones se obtienen sin marginalizar sobre los parámetros.

CAPÍTULO 2

Modelo 2PL con Trazo Latente Asimétrico

En el capítulo anterior se presentaron los supuestos básicos de la TRI, así como algunos de los métodos de estimación que se implementan en el área. De manera particular, en los procedimientos de estimación propuestos por Bock & Lieberman (1970) y Bock & Aitkin (1981) se planteó la necesidad de suponer una distribución de probabilidad para el trazo latente, y aunque es posible considerar distintas distribuciones, es común asumir que $\theta_i \sim N(\mu = 0, \sigma = 1)$ para cada uno de los $i = 1, ..., N$ evaluados.

Sin embargo, como se establece en Woods (2008), en ámbitos como el análisis de rasgos de personalidad, el trazo podría presentar características que sugieran que una distribución diferente a la normal estándar es más adecuada. En Boulet (1996),Levent et al. (2001),Preston & Reise (2014), se hace referencia a que los estudios de simulación indican que el sesgo en las estimaciones de los modelos TRI, cuando se asume una distribución equivocada para el trazo, pueden llegar a ser considerables.

Si bien es cierto que la selección de una distribución diferente a la normal estándar podría derivar en dificultades computacionales, como la obtención de los nodos y pesos de cuadratura en las metodologías de estimación de máxima verosimilitud marginal, también resulta indispensable generar la mejor escala posible para el trazo latente, de manera que la caracterización de los rasgos de interés de lugar a interpretaciones apropiadas.

Dentro de las metodologías de estimación recientes, en modelos donde no se presume normalidad para el trazo, se encuentran aquellas que utilizan splines para obtener estimaciones adecuadas de los ítems y la esperanza posteriori del trazo (EAP). En Woods & Thissen (2006a) se propone un método particular denominado *Curva de Ramsay en TRI* para datos dicotómicos en el que se utilizan tanto splines como máxima verosimilitud marginal para llevar a cabo la estimación y, en publicaciones posteriores, el mismo autor generaliza los resultados para modelos logísticos de tres parámetros y datos politómicos (Woods (2006b),Woods (2008)).

2.1. Modelo Propuesto

Se considera un test dicotómico de p ítems aplicados a N individuos, donde la probabilidad de que el evaluado i acierte al ítem j estará dada por un modelo TRI logístico de

dos parámetros como en la ecuación (1.1). Además de los supuestos usuales de la TRI, se asumirá lo siguiente:

- El trazo latente subyacente al test es unidimensional.
- El trazo latente sigue una distribución Gamma[1], esto es, $\theta_i \sim Gamma(\alpha, \beta)$, para $i = 1, 2, ..., N$, de manera que

$$g(\theta|\alpha, \beta) = \frac{\theta^{\alpha-1} e^{-\frac{\theta}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)} \qquad (\theta > 0, \alpha, \beta > 0) \tag{2.1}$$

2.2. Estimación

Para llevar a cabo el proceso de estimación se empleará la metodología propuesta por Bock & Aitkin (1981) que fue descrita de manera general en el capítulo previo. En una primera etapa se realizará la estimación de los parámetros de los ítems implementando el algoritmo EM y, una vez obtenidas las estimaciones, se estimará el trazo latente utilizando máxima verosimilitud.

Cabe resaltar que en el proceso de estimación se eliminan aquellos registros de respuesta que tienen score (puntaje) perfecto, p, o score nulo, debido a los siguiente:

- Si el i-ésimo individuo responde equivocadamente los p ítems del test, la función de verosimilitud queda reducida a

$$L(\theta_i) = \prod_{j=1}^{p} Q_j(\theta_i)$$

que es una función decreciente respecto al trazo y, por lo tanto, el estimador máximo verosímil para θ_i será $\theta_i = -\infty$

- Si el individuo responde acertadamente todos los ítems, la función de verosimilitud será

$$L(\theta_i) = \prod_{j=1}^{p} P_j(\theta_i)$$

que en este caso resulta creciente respecto al trazo, de manera que el estimador de máxima verosimilitud para θ_i será $+\infty$

2.2.1. Estimación de los Parámetros de los Ítems

Con el fin de implementar los métodos de cuadratura, supóngase que el trazo latente toma un conjunto finito de q valores, θ_k, con probabilidades π_k,$(k = 1, 2, ..., q)$. Se define n_{kj} como el número de individuos con nivel del trazo θ_k que respondieron al ítem j, de manera que $\sum_{k=1}^{q} n_{kj} = N$; sea además r_{kj} el número de evaluados con nivel de trazo θ_k que acertaron el j-ésimo ítem. Entonces, al seleccionar aleatoriamente n individuos

[1]Para la distribución Gamma se utilizará la parametrización forma-escala

que hayan respondido el j-ésimo ítem, la probabilidad conjunta de que los n_{kj} evaluados tengan nivel de trazo θ_k para $k = 1, 2, ..., q$, quedará determinada por

$$P(n_{kj}|\pi_k) = \frac{n!}{\prod_{k=1}^{q}(n_{kj})!} \prod_{k=1}^{q} \pi_k^{n_{kj}}, \qquad (j = 1, 2, ..., p) \tag{2.2}$$

Además, para n_{kj} y θ_k dados, la probabilidad de obtener r_{kj} respuestas correctas al ítem j en los n_{kj} ensayos de los individuos con nivel de trazo θ_k será

$$P(r_{kj}|n_{kj}, \theta_k) = \prod_{j=1}^{p} \binom{n_{kj}}{r_{kj}} P_j(\theta_k)^{r_{kj}} Q_j(\theta_k)^{n_{kj}-r_{kj}} \tag{2.3}$$

De manera que, para $\boldsymbol{n} = (n_{1j}, n_{2j}, ..., n_{qj})'$ y $\boldsymbol{r} = (r_{1j}, r_{2j}, ..., r_{qj})'$,

$$P(\boldsymbol{n}, \boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{\pi}) = \prod_{k=1}^{q} \prod_{j=1}^{p} \binom{n_{kj}}{r_{kj}} P_j(\theta_k)^{r_{kj}} Q_j(\theta_k)^{n_{kj}-r_{kj}} \frac{n!}{\prod_{k=1}^{q}(n_{kj})!} \prod_{k=1}^{q} \pi_k^{n_{kj}} \tag{2.4}$$

Nótese que

$$\begin{aligned} P(\boldsymbol{n}, \boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{\pi}) &= P(\boldsymbol{n}|\boldsymbol{\theta}, \boldsymbol{\pi}) P(\boldsymbol{r}|\boldsymbol{n}, \boldsymbol{\theta}, \boldsymbol{\pi}) \\ &= P(\boldsymbol{n}|\boldsymbol{\pi}) P(\boldsymbol{r}|\boldsymbol{h}, \boldsymbol{\theta}) \end{aligned}$$

Por lo que, tal y como se afirma en Harwell et al. (1988), $(\boldsymbol{n}, \boldsymbol{r})$ es una estadística suficiente para los datos completos $(\boldsymbol{Y}, \boldsymbol{\theta})$. Así pues, la función de log-verosimilitud, sin tener en cuenta los términos constantes, para $(\boldsymbol{Y}, \boldsymbol{\theta})$ será

$$\mathcal{L} = \log(L) = \sum_{k=1}^{q} \sum_{j=1^p} [r_{kj} \log P_j(\theta_k) + (n_{kj} - r_{kj}) \log Q_j(\theta_k)] \tag{2.5}$$

Ahora bien, $(\boldsymbol{n}, \boldsymbol{r})$ no es observable, pero, puede calcularse la esperanza de la log-verosimilitud condicionándola en $\boldsymbol{\beta} = (a_j, b_j)$, de manera que

$$E[\mathcal{L}] = \sum_{k=1}^{q} \sum_{j=1}^{p} [E(r_{kj}|\boldsymbol{Y}, \boldsymbol{\beta}) \log P_j(\theta_k) + E[(n_{kj} - r_{kj})|\boldsymbol{Y}, \boldsymbol{\beta}] \log Q_j(\theta_k)] \tag{2.6}$$

Haciendo $\overline{n}_{kj} = E[n_{kj}|\boldsymbol{Y}, \boldsymbol{\beta}]$ y $\overline{r}_{kj} = E[r_{kj}|\boldsymbol{Y}, \boldsymbol{\beta}]$, la ecuación (2.6) puede escribirse como

$$E[\mathcal{L}] = \sum_{k=1}^{q} \sum_{j=1^p} [\overline{r}_{kj} \log P_j(\theta_k) + (\overline{n}_{kj} - \overline{r}_{kj}) \log Q_j(\theta_k)] \tag{2.7}$$

Cabe resaltar que maximizar la ecuación (2.7) es similar a maximizar la expresión (1.6) para completar el paso E del algoritmo EM.

Como se propone en Bock & Aitkin (1981), Harwell et al. (1988) y Andrade et al. (2000), para resolver las ecuaciones de maximización derivadas de la expresión (2.7), es necesario recurrir a métodos de cuadratura. Tal y como se planteó anteriormente, la distribución del trazo latente es Gamma de parámetros α y β, de manera que ésta será la distribución empleada para obtener los nodos y los pesos de cuadratura.

El objetivo entonces es encontrar un conjunto de $k = 1, 2, ..., q$ nodos, X_k, sobre el soporte $g(\theta|\alpha, \beta)$ y sus pesos respectivos $A(X_k)$, (Gráfica 2.1), de manera que

$$\frac{P(\boldsymbol{Y_i}|\theta_k)g(\theta_k|\epsilon)}{\int P(\boldsymbol{Y_i}|\theta_k)g(\theta_k|\epsilon)d\theta} \approx \frac{\prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} A(X_k)}{\sum_{k=1}^{q}\prod_{j=1}^{p} P_j(X_k)^{y_{ij}} Q_j(X_k)^{1-y_{ij}} A(X_k)} \tag{2.8}$$

para $i = 1, 2, ..., N$, $j = 1, 2, ..p$ y donde $\boldsymbol{\epsilon} = (\alpha, \beta)$, $A(X_k) = g(\theta_k|\boldsymbol{\epsilon})\Delta_k$, $P_j(X_k) = \left[1 + e^{-a_j(X_k - b_j)}\right]^{-1}$ y $Q_j(X_k) = 1 - P_j(X_k)$.

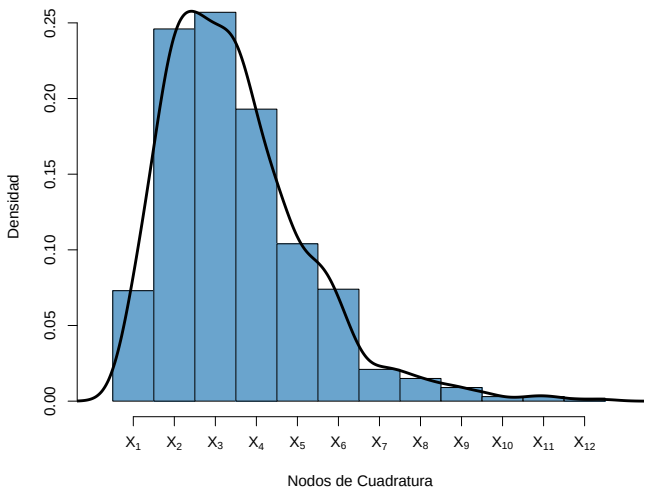


FIGURA 2.1. Nodos de Cuadratura con Distribución de Pesos Gamma

Para llevar a cabo este procedimiento se implementó la metodología propuesta en Fernandes & Atchley (2006), en la que se realiza el método de cuadratura utilizando distribuciones de pesos con soporte en $\mathbb{R}^+$ para aproximar

$$\int_a^b f(x)w(x)dx \approx \sum_{i=1}^{n-1} f(x_i)w_i \qquad (i = 1, 2, ..., n-1) \tag{2.9}$$

en este caso, x_i representa los nodos del proceso y w_i los pesos correspondientes. Fernandes & Atchley (2006) afirma que, una vez se aproxima a f por medio de una serie polinómica, la cuadratura dependerá de los momentos de la distribución de pesos $w(x)$, por lo que se requiere que éstos existan y sean finitos.

Siempre es posible encontrar un conjunto de polinomios ortogonales con relación al producto interno[2] $\langle f, g\rangle$, tales que

$$\begin{aligned} p_{-1} = 0, \qquad , p_0 = 1 \\ p_{i+1}(x) = (x - \gamma_i)p_i(x) - \eta_i p_{i-1}(x) \end{aligned} \tag{2.10}$$

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

$$\gamma_i = \frac{\langle xp_i, pi\rangle}{\langle p_i, p_i\rangle} \qquad \eta_i = \frac{\langle p_i, p_i\rangle}{\langle p_{i-1}, p_{i-1}\rangle} \tag{2.11}$$

teniendo en cuenta que $\eta_0 = \int w(x)dx$. Después de obtener γ_i y η_i, se construye la matriz

$$J = \begin{pmatrix} \gamma_0 & \sqrt{\eta_1} & & & & \\ \sqrt{\eta_1} & \gamma_1 & \sqrt{\eta_2} & & & \\ & \sqrt{\eta_2} & \ddots & \ddots & & \\ & & \ddots & \ddots & \sqrt{\eta_{n-2}} & \\ & & & \sqrt{\eta_{n-2}} & \gamma_{n-2} & \sqrt{\eta_{n-1}} \\ & & & & \sqrt{\eta_{n-1}} & \gamma_{n-1} \end{pmatrix} \tag{2.12}$$

De aquí, los nodos x_i se obtienen como los autovalores de la matriz J y los pesos se obtienen mediante

$$w_i = \eta_0 u_{i,0}^2 \tag{2.13}$$

teniendo en cuenta que $u_{i,0}^2$ es el primer elemento del autovector normalizado u_i de la matriz J.

Para la parametrización de la distribución Gamma dada en la ecuación (2.1), los coeficientes de las expresiones (2.11) y (2.12) están dados por $\gamma_i = \beta(\alpha + 2i)$, para $i = 0, 1, 2, ..., (n-1)$ y $\eta_i = \beta^2 i(\alpha + i - 1)$ para $i = 1, 2, ..., (n-1)$.

Para los nodos $x_i = X_k$ y los pesos $w_i = A(X_k)$, $(k = 1, 2, ..., q)$, se satisfacen las siguientes igualdades,

- $\sum_{k=1}^{q} A(X_k) = 1$
- $\sum_{k=1}^{q} X_k A(X_k) = 0$
- $\sum_{k=1}^{q} X_k^2 A(X_k) = 1$

Una vez obtenidos los nodos de cuadratura y sus pesos respectivos, se procede a organizar la matriz de datos compuesta por los valores de la variable

$$y_{ij} = \begin{cases} 1, & \text{el individuo } i \text{ responde correctamente el ítem } j \\ 0, & \text{en caso contrario} \end{cases}$$

[2] $\langle f, g\rangle = \int f(x)g(x)w(x)dx$

Puede ocurrir entonces que varios individuos tengan asignado el mismo vector $\boldsymbol{Y}_i$ de respuestas a los ítems del test, por lo que, en lugar de trabajar con la matriz completa se consideraran solamente los patrones de respuesta diferentes. En resumen:

- El test consta de p ítems, N evaluados y s patrones de respuesta con sus respectivas frecuencias denotadas por n_l.
- Se obtendrán q nodos de cuadratura y sus pesos correspondientes.
- $j = 1, 2, 3, ..., p$, indexa los ítems del test.
- $i = 1, 2, 3, ..., N$, es el índice para los individuos.
- $l = 1, 2, 3, ..., s$, corresponde a los patrones de respuesta del test.
- $k = 1, 2, 3, ..., q$, indexa los nodos y pesos de cuadratura.

Después de obtener los patrones de respuesta, Bock & Aitkin (1981) y Harwell et al. (1988) plantean las expresiones para $\overline{n}_{kj}$ y $\overline{r}_{kj}$ en términos de los mismos

$$\overline{n}_{kj} = \frac{\sum_{l=1}^{s} n_l L_l(X_k) A(X_k)}{\sum_{l=1}^{s}\sum_{k=1}^{q} L_l(X_k) A(X_k)} \qquad \overline{r}_{kj} = \frac{\sum_{l=1}^{s} y_{lj} n_l L_l(X_k) A(X_k)}{\sum_{l=1}^{s}\sum_{k=1}^{q} L_l(X_k) A(X_k)} \tag{2.14}$$

donde $L_l(X_k)$ está dada por la ecuación (1.13) haciendo $i = l$.

Cuando se haya obtenido la ecuación (2.7), el paso E del algoritmo EM se completa maximizándola por medio de la función `optim` del software estadístico `R`, que implementa la metodología propuesta en Nelder & Mead (1965). Como valores iniciales para los parámetros de los ítems se consideran los propuestos en Harwell et al. (1988) y Andrade et al. (2000), es decir $\boldsymbol{\beta}^{(0)} = (a_j = 1, b_j = 0)$ para $j = 1, 2, ..., p$.

Una vez obtenido $\boldsymbol{\beta}^{(k)} = (a_j, b_j)^{(k)}$, se efectúa el paso M del algoritmo actualizando en cada iteración la función de verosimilitud obtenida en el paso E hasta la convergencia. En particular, como criterio de parada se propone $|\boldsymbol{\beta}^{(k+1)} - \boldsymbol{\beta}^{(k)}| < 0.001$.

Para obtener el error estándar de los estimadores de los parámetros de los ítems se debe calcular la matriz de varianzas–covarianzas del modelo, que en este caso es la inversa de la matriz de información

$$I_j = \begin{pmatrix} I_{a_j a_j} & I_{a_j b_j} \\ I_{a_j b_j} & I_{b_j b_j} \end{pmatrix} \qquad (j = 1, 2, ..., p) \tag{2.15}$$

donde,

- $I_{a_j a_j} = \sum_{i=1}^{N} (\theta_i - b_j)^2 P_j(\theta_i) Q_j(\theta_i)$
- $I_{a_j b_j} = a_j \sum_{i=1}^{N} (\theta_i - b_j) P_j(\theta_i) Q_j(\theta_i)$

- $I_{b_j b_j} = a_j^2 \sum_{i=1}^{N} P_j(\theta_i) Q_j(\theta_i)$

2.2.2. Estimación del Trazo Latente

Después de haber obtenido los parámetros de los ítems se procede a la estimación del trazo latente individuo por individuo considerando como valores iniciales los puntajes del score estandarizado para cada evaluado.

Como los parámetros de discriminación y dificultad de modelo ya han sido estimados, para determinar el trazo debe maximizarse la función de verosimilitud dada por la expresión (1.4), para ello puede recurrirse al algoritmo de Newton–Raphson o Scoring-Fisher (Andrade et al., 2000). Computacionalmente, el proceso de optimización puede llevarse a cabo implementando la función `optimize` de `R` que se basa en la metodología propuesta por Brent (1973).

En Baker (2001) se propone un procedimiento iterativo que se lleva a cabo mediante la expresión:

$$\theta_{k+1} = \theta_k + \frac{\sum_{j=1}^{p} a_j \left[y_{ij} - P_j(\theta_k) \right]}{\sum_{i=1}^{N} a_j^2 P_j(\theta_k) Q_j(\theta_k)} \tag{2.16}$$

El procedimiento se detiene cuando $|\theta_{k+1} - \theta_k| < 0.001$ por ejemplo. Como medida adicional puede obtenerse el error estándar de la estimación del trazo a través de la expresión

$$SE(\widehat{\theta_i}) = \frac{1}{\sqrt{I(\theta)}} = \frac{1}{\sqrt{\sum_{j=1}^{p} a_j^2 P_j(\theta_i) Q_j(\theta_i)}} \tag{2.17}$$

donde $I(\theta)$ es la función de información del trazo (Hambleton et al. (1991)).

Otra alternativa para la estimación del trazo es la media (o esperanza) posteriori (EAP), en donde se emplean las estimaciones de los parámetros de los ítems, las respuestas observadas y el valor esperado de la distribución posteriori para estimar el trazo. Vale la pena recordar que la distribución posteriori es el resultado del producto entre una distribución a priori y la verosimilitud del modelo.

Partiendo del supuesto de que el trazo sigue una distribución gamma, y aproximando nuevamente las integrales con los métodos de cuadratura descritos anteriormente, el estimador EAP de θ puede obtenerse mediante,

$$\overline{\theta_i} = \frac{\sum_{k=1}^{q} X_k L(X_k) A(X_k)}{\sum_{k=1}^{q} L(X_k) A(X_k)} \tag{2.18}$$

Sin embargo, en casos en los que el trazo siga una distribución asimétrica, usar la moda o la mediana posterior podría arrojar estimaciones con mayor precisión.

2.3. Pruebas de Bondad de Ajuste

En el ámbito de la TRI, *"...se define un residual r_{ij} como la diferencia entre el comportamiento observado del ítem para un subgrupo de individuos y su comportamiento esperado en el mismo conjunto..."* (Hambleton et al., 1991). Es decir,

$$res_{ij} = P_j(\theta_i) - E[P_j(\theta_i)] \tag{2.19}$$

donde para el j-ésimo ítem

- i denota el nivel de habilidad del trazo para el subgrupo de individuos.
- $P_j(\theta_i)$ representa la proporción de respuestas acertadas al ítem j en el nivel de trazo i
- $E[P_j(\theta_i)]$ es la proporción esperada de aciertos al j-ésimo ítem, que se obtiene desde el modelo logístico 2PL.

Para calcular los residuales, suele dividirse la escala del trazo latente en intervalos de igual ancho, más o menos entre 10 y 15, garantizando que el número de individuos en cada nivel del trazo no sea demasiado pequeño o nulo.

Calcular la diferencia entre las frecuencias esperadas y observadas resulta bastante útil a la hora de examinar la bondad de ajuste del modelo. En la Gráfica 2.2 se presentan la comparación para un ítem arbitrario: la gráfica de la izquierda corresponde al caso en el que la diferencia entre las dos frecuencias es mínima (residuales cercanos a cero), mientras que la gráfica derecha presenta diferencias más altas, lo que indica un ajuste pobre del modelo.

Como se menciona en Hambleton et al. (1991), los residuales estandarizados están definidos por

$$z_{ij} = \frac{N_i \left(P_j(\theta_i) - E[P_j(\theta_i)]\right)}{\sqrt{E[P_j(\theta_i)]\left(1 - E[P_j(\theta_i)]\right)}} \tag{2.20}$$

donde N_i representa el número de individuos en el nivel i del trazo.

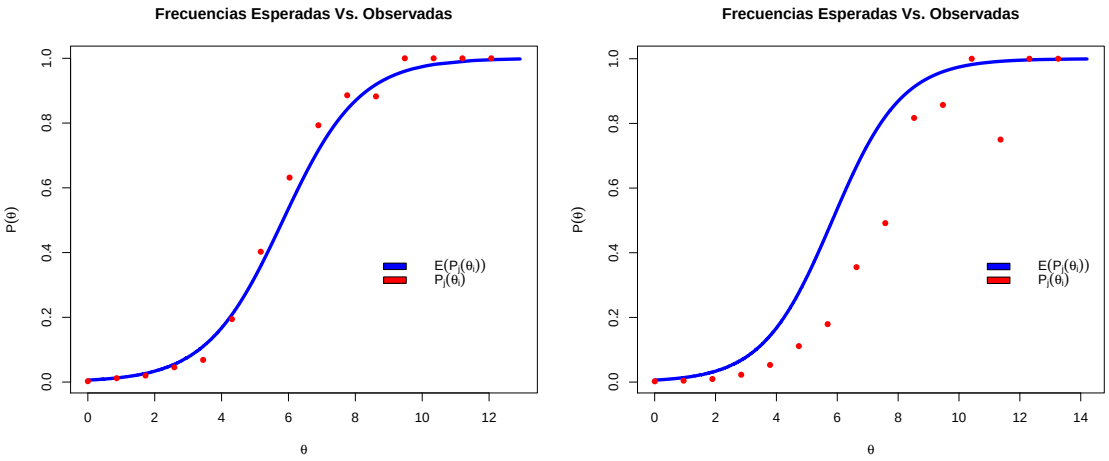


FIGURA 2.2. Frecuencias esperadas Vs. frecuencias observadas

A partir de los residuales estandarizados se construye una de las estadísticas de bondad de ajuste más utilizadas en TRI, la estadística Q_1 propuesta por Yen (1981), que está dada por

$$\begin{aligned} Q_1 &= \sum_{i=1}^{m} z_{ij}^2 \\ &= \sum_{i=1}^{m} \frac{N_i \left(P_j(\theta_i) - E[P_j(\theta_i)]\right)^2}{E[P_j(\theta_i)]\left(1 - E[P_j(\theta_i)]\right)} \end{aligned} \tag{2.21}$$

En este caso, m representa el número de intervalos en que fue dividido el trazo. La estadística Q_1 sigue una distribución Chi–cuadrada con $\nu = m - \rho$ grados de libertad, donde ρ es el número de parámetros de los ítems en el modelo.

Para el modelo propuesto, se considera $m = 10$ y $\rho = 2$, además, las categorías del trazo latente se construyen a partir de los deciles de la escala.

2.4. Implementación

Para llevar a cabo el procedimiento de estimación se empleará el software `R`.

2.4.1. Análisis Preliminar de los Datos

Con el objetivo de establecer la dimensión del test se realiza un Análisis de Componentes Principales (ACP) mediante la función `PCA` de la librería `FactoMineR`. Una vez obtenidos los autovalores, se determina el porcentaje de varianza explicada por cada componente y se calcula la proporción entre el primer y segundo autovalor. Para concluir que el test es *esencialmente unidimensional*, Reckase (1979) sugiere que el cociente entre el primer autovalor y el número de ítems sea cercano a 0.20. La afirmación anterior suele apo-

yarse además con la Curva de Cronbach-Mesbah que en este caso se construye utilizando la función `alpha.curve` del paquete `CMC` y que debe ser monótona creciente.

Posteriormente, se calculan los scores (puntajes), y se realiza el histograma correspondiente con la finalidad de obtener un primer indicio del comportamiento del trazo latente. Adicionalmente se eliminan del test los registros con score nulo y score perfecto para proceder con la estimación de los parámetros.

Para determinar la distribución apropiada del trazo latente se utilizan modelos aditivos generalizados para localización, escala y forma (GAMLSS), tomando como variable observada a la primera componente del ACP cuya correlación con el score suele ser alta (Rigby & Stasinopoulos (2005), Stasinopoulos & Rigby (2007)). Empleando la función `fitDist` del paquete `gamlss` se obtiene un listado de distribuciones que podrían describir el comportamiento del trazo con sus respectivos AIC; después de seleccionar la distribución para el trazo, se ajusta un modelo GAMLSS sin covariables para determinar el valor de los parámetros de localización, escala y/o forma necesarios.

2.4.2. Estimación

Una vez seleccionada la distribución del trazo latente, se sigue el procedimiento descrito en la Sección 2.2.1 para obtener los nodos y los pesos de cuadratura y se establecen los valores iniciales de los parámetros de dificultad y discriminación. A continuación se calculan los patrones de respuesta de la matriz de datos con sus frecuencias correspondientes.

Se calculan además las ecuaciones (2.6) y (2.14) para llevar a cabo el paso E del algoritmo EM y se completa el procedimiento maximizando la última expresión a través de la función `optim`. Para llevar a cabo el paso M, se actualizan las expresiones (2.6) y (2.14) con los valores obtenidos hasta que el algoritmo resulte convergente, en particular el criterio de parada aquí empleado es que la diferencia entre la estimación actual y la anterior sea inferior a 0.001.

Se utilizarán aquí las estimaciones de los parámetros de los ítems obtenidas en el apartado anterior. Se calcula en `R` la expresión (1.4) y se optimiza mediante la utilidad `optimize`.

Después de obtener los parámetros de los ítems y del trazo latente, se implementan las ecuaciones (2.15) y (2.17) para calcular los errores estándar de las estimaciones.

2.4.3. Análisis gráfico

Se calculan las frecuencias esperadas y las frecuencias observadas como se indica en la Sección 2.3 y se construyen las gráficas correspondientes. Adicionalmente se elaboran las gráficas de los residuales estandarizados (ecuación (2.20)) para cada uno de los ítems y se construyen las gráficas de las densidades de los parámetros estimados.

CAPÍTULO 3

Simulación

Con la finalidad de evaluar la calidad de recuperación de los parámetros del algoritmo de estimación se realiza la simulación de un conjunto de datos bajo un modelo TRI de dos parámetros en el que el trazo latente sigue una distribución asimétrica.

3.1. Generación del Test

Se genera un test dicotómico con el que se evalúan $p = 20$ ítems a $N = 5000$ individuos y en el que la probabilidad de que el individuo i acierte al ítem j está dada por

$$P(Y_{ij} = 1|\theta_i) = \frac{1}{1 + e^{-a_j(\theta_i - b_j)}}; \qquad (i = 1, 2, ..., N;\ j = 1, 2, ..., p) \tag{3.1}$$

Así pues, para $j = 1, 2, ..., p$, la distribución del parámetro de discriminación es $LN(0, 1/4)$, mientras que el parámetro de dificultad se genera a partir de una distribución $GA(2, 1/2)$. Adicionalmente, y para $i = 1, 2, ..., N$, se considera que la distribución del trazo es $GA(2, 1/2)$; algunas medidas descriptivas de los parámetros simulados se presentan en la Tabla 3.1.

Medida	a_j	b_j	θ_i
mín	0.641	0.0951	0.0079
Q_1	0.863	0.3639	0.4924
Q_2	1.074	0.7992	0.8488
$\overline{x}$	1.056	0.9666	1.0053
Q_3	1.252	1.2956	1.3614
máx	1.635	2.8078	6.5959
sd	0.271	0.7547	0.6955

Tabla 3.1. Medidas Descriptivas de los Parámetros Simulados

3.2. Recuperación de los Parámetros

El proceso de recuperación de los parámetros se llevará a cabo utilizando la metodología propuesta por Bock & Aitkin (1981) y Harwell et al. (1988), es decir, se implementará el algoritmo EM para recuperar los parámetros de los ítems y, una vez obtenidos, se empleará el procedimiento de Máxima Verosimilitud[1] para obtener el trazo.

Se eliminaron del test los registros de individuos que obtuvieron scores nulos o perfectos en la prueba, con lo que el test depurado consta de $N = 4983$ individuos y $p = 20$ ítems.

Se implementa el algoritmo para recuperar los parámetros de los ítems empleando $q = 20$ puntos de cuadratura cuyos pesos están determinados por una distribución gamma con $\alpha = 2$ y $\beta = 1/2$ siguiendo la propuesta de Fernandes & Atchley (2006). Se consideraron como valores iniciales[2] $a_j = 1$ y $d_j = 0$ tal y como se propone en Harwell et al. (1988) y Baker (1985), teniendo en cuenta además que $d_j = -a_j \cdot b_j$.

Una vez obtenidos los parámetros de los ítems se procede a recuperar el trazo latente empleando la función `optimize` de `R` considerando como valores iniciales a los scores estandarizados de cada individuo; con el objetivo de identificar el intervalo de optimización se construyeron las gráficas de la función de verosimilitud para cada uno de los respondientes y se observó que, aunque algunas de las curvas son multimodales, todas parecen tener su valor máximo dentro del intervalo [0,5](ver Gráfica 3.1).

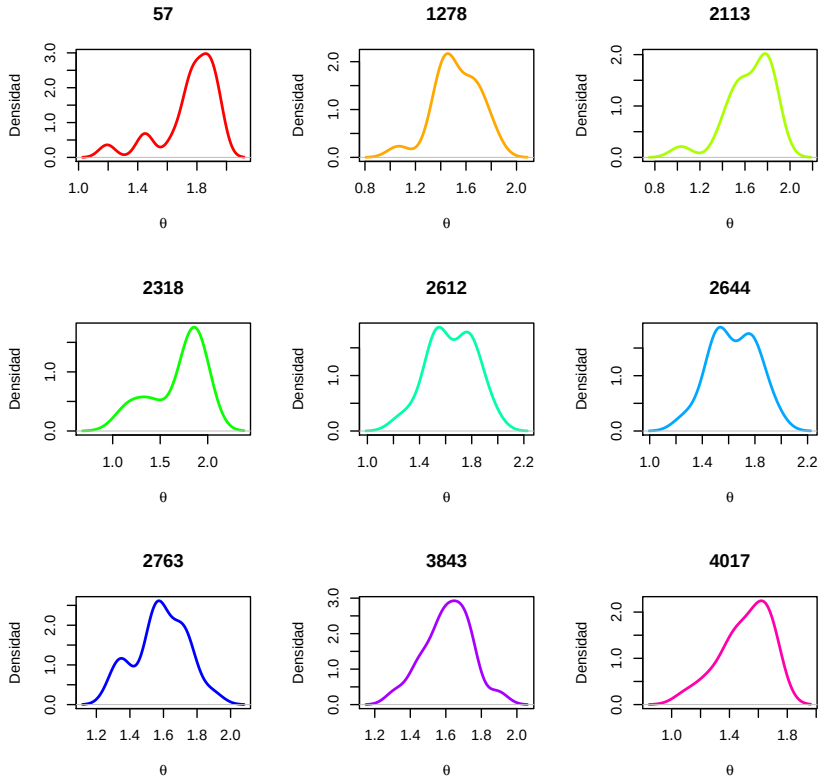


FIGURA 3.1. Función de Verosimilitud para algunos individuos seleccionados aleatoriamente

[1]También se presenta la recuperación del trazo por medio del método EAP en el apéndice B.

[2]En el Apéndice B se presenta otra alternativa para obtener los valores iniciales de los parámetros de los ítems.

3.3. Valores Simulados Vs. Recuperados

Para evaluar la calidad del ajuste del modelo se realizaron las comparaciones entre los valores de los parámetros en la simulación y los valores obtenidos en el proceso de recuperación.

Tal y como se observa en la Gráfica 3.2 el procedimiento implementado recupera de manera adecuada los parámetros de discriminación de los ítems: la correlación obtenida entre a_j y $\widehat{a}_j$ fue de 0.96947, los errores estándar se encuentran entre 0.0393 y 0.0597, en tanto que el error absoluto está entre 0.0001 y 0.1390.

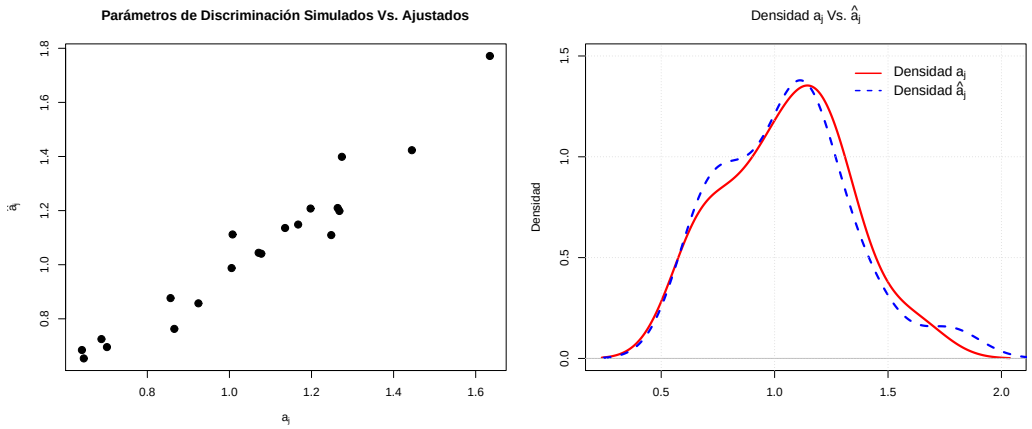


FIGURA 3.2. Valores simulados de los parámetros de discriminación Vs. los valores recuperados

Análogamente, para el parámetro de dificultad puede observarse en la Gráfica 3.3, una adecuada recuperación de los valores simulados; en este caso los errores estándar oscilan entre 0.0190 y 0.0721, mientras que el error absoluto se encuentra entre 0.0003 y 0.0591; la correlación entre los valores simulados y recuperados fue de 0.9991.

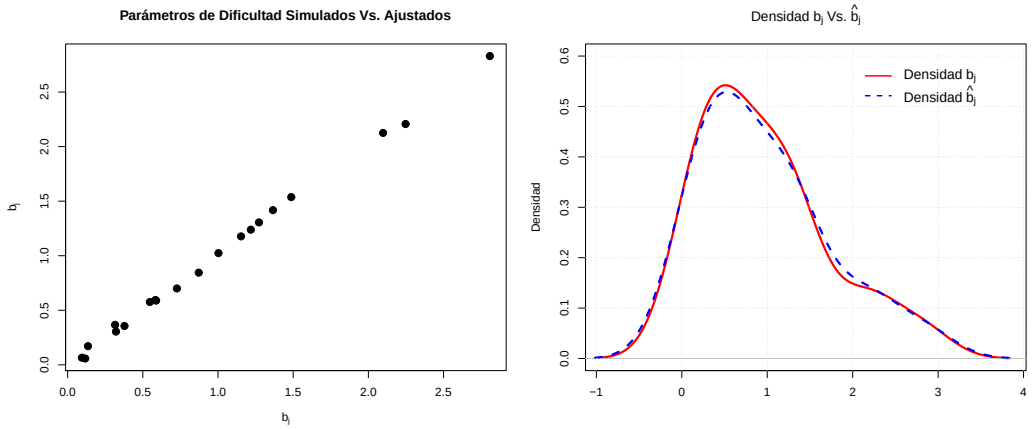


FIGURA 3.3. Valores simulados de los parámetros de dificultad Vs. los valores recuperados

En la Tabla 3.2 se detalla el valor simulado, el valor recuperado, el error estándar de la estimación y el error absoluto para los parámetros de los ítems, mientras que en la Tabla 3.3 se presentan las principales medidas descriptivas del trazo recuperado.

j	a_j	$\widehat{a}_j$	sd_{a_j}	$\lvert a_j - \widehat{a}_j \rvert$	b_j	$\widehat{b}_j$	sd_{b_j}	$\lvert b_j - \widehat{b}_j \rvert$
1	1.2740	1.3989	0.0565	0.1249	0.3155	0.3664	0.0280	0.0509
2	1.6345	1.7715	0.0597	0.1370	0.7267	0.6999	0.0190	0.0268
3	1.2480	1.1095	0.0462	0.1385	0.8717	0.8450	0.0277	0.0267
4	1.0077	1.1118	0.0503	0.1041	0.3780	0.3560	0.0350	0.0220
5	1.2679	1.1986	0.0521	0.0693	2.8078	2.8301	0.0683	0.0223
6	0.8657	0.7628	0.0451	0.1029	0.1170	0.0579	0.0624	0.0591
7	1.1977	1.2076	0.0549	0.0099	0.1354	0.1710	0.0378	0.0356
8	1.1674	1.1485	0.0454	0.0189	1.1531	1.1772	0.0279	0.0241
9	1.0053	0.9876	0.0427	0.0177	1.2723	1.3053	0.0333	0.0330
10	1.0711	1.0441	0.0468	0.0270	0.5838	0.5940	0.0318	0.0102
11	0.7016	0.6955	0.0409	0.0061	0.5883	0.5886	0.0471	0.0003
12	0.6406	0.6847	0.0420	0.0441	0.3215	0.3046	0.0573	0.0169
13	0.8565	0.8766	0.0478	0.0201	0.0951	0.0654	0.0548	0.0297
14	1.1356	1.1357	0.0450	0.0001	1.2183	1.2389	0.0288	0.0206
15	0.6452	0.6536	0.0402	0.0084	0.5468	0.5758	0.0504	0.0290
16	0.6882	0.7251	0.0393	0.0369	2.2472	2.2070	0.0721	0.0402
17	1.0777	1.0407	0.0441	0.0370	2.0977	2.1255	0.0502	0.0278
18	0.9244	0.8573	0.0407	0.0671	1.3655	1.4178	0.0395	0.0523
19	1.2637	1.2096	0.0458	0.0541	1.4871	1.5375	0.0315	0.0504
20	1.4445	1.4233	0.0505	0.0212	1.0031	1.0239	0.0225	0.0208

TABLA 3.2. Valores simulados y recuperados de los parámetros de discriminación y dificultad

mín	Q_1	Q_2	$\overline{x}$	Q_3	máx	sd
0.0335	0.4220	0.9020	1.0500	1.4900	4.9400	0.8184

TABLA 3.3. Resumen Descriptivo de $\widehat{\theta}_i$

En la Gráfica 3.4 se comparan las densidades de los valores simulados y recuperados del trazo. En este caso, la correlación obtenida fue de 0.829116 y el error absoluto se encuentra alrededor de $\widetilde{x} = 0.36$ con una desviación estándar de 0.2867 como se indica en la Tabla 3.4. Finalmente, en la Tabla 3.5 se presenta un resumen de las medidas descriptivas de los errores absolutos de los parámetros recuperados.

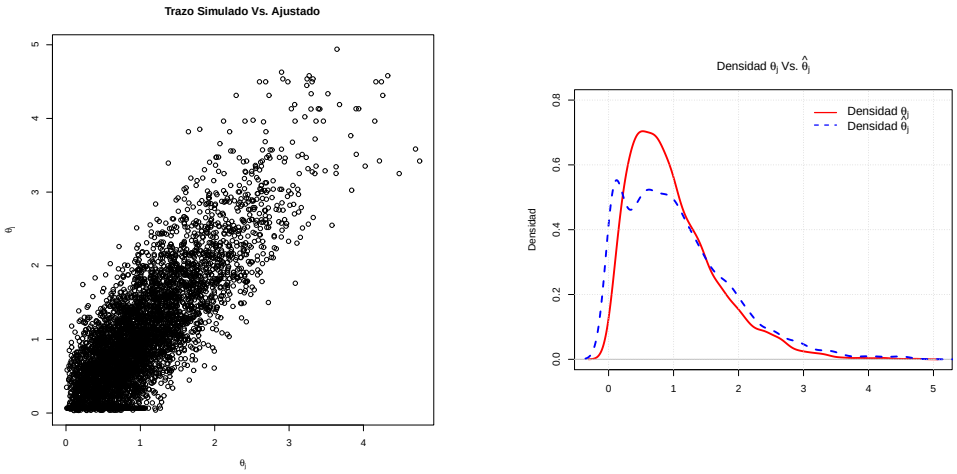


FIGURA 3.4. Valores simulados del trazo Vs. valores obtenidos en la estimación

mín	Q_1	Q_2	$\overline{x}$	Q_3	máx	sd
0.00025	0.14000	0.29500	0.36000	0.51000	2.17000	0.28668

TABLA 3.4. Medidas Descriptivas para $|\theta_i - \widehat{\theta}_i|$

	mín	Q_1	$\widetilde{x}$	$\overline{x}$	Q_3	máx	sd	Correlación
$\lvert a_j - \widehat{a}_j \rvert$	0.0001	0.0186	0.0370	0.0523	0.0777	0.1390	0.0457	0.969
$\lvert b_j - \widehat{b}_j \rvert$	0.0003	0.0217	0.0273	0.0299	0.0368	0.0591	0.0148	0.999
$\lvert \theta_i - \widehat{\theta}_i \rvert$	0.0002	0.1400	0.2950	0.3600	0.5100	2.1700	0.2867	0.829

TABLA 3.5. Medidas Descriptivas de los Errores Absolutos

3.4. Pruebas de Bondad de Ajuste

Utilizando los valores recuperados de los parámetros de los ítems y del trazo se construyeron las curvas características de cada ítem y se compararon con las frecuencias esperadas en distintos niveles de θ como se propone en Hambleton et al. (1991).

En la Gráfica 3.5 puede observarse para 4 de los ítems seleccionados aleatoriamente, que las frecuencias esperadas y observadas presentan comportamientos similares, lo que apoya la suposición inicial de que el trazo sigue una distribución asimétrica.

Se presentan las gráficas de los residuales estandarizados para los mismos 4 ítems en la Figura 3.6 y, adicionalmente, en la Tabla 3.6 se encuentra la correlación entre las frecuencias observadas y esperadas para cada uno de los 20 ítems.

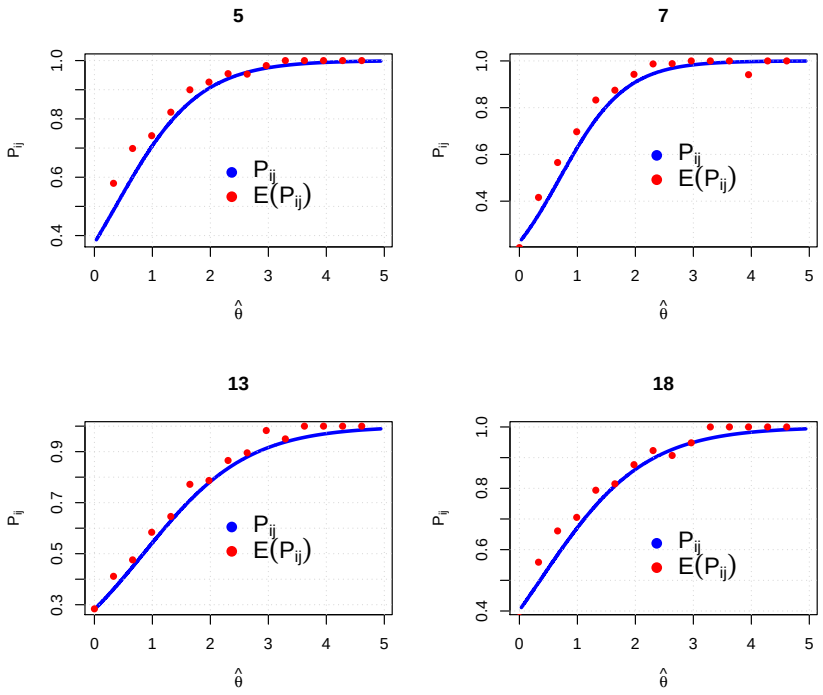


FIGURA 3.5. Frecuencias Esperadas Vs. Observadas para algunos ítems seleccionados aleatoriamente

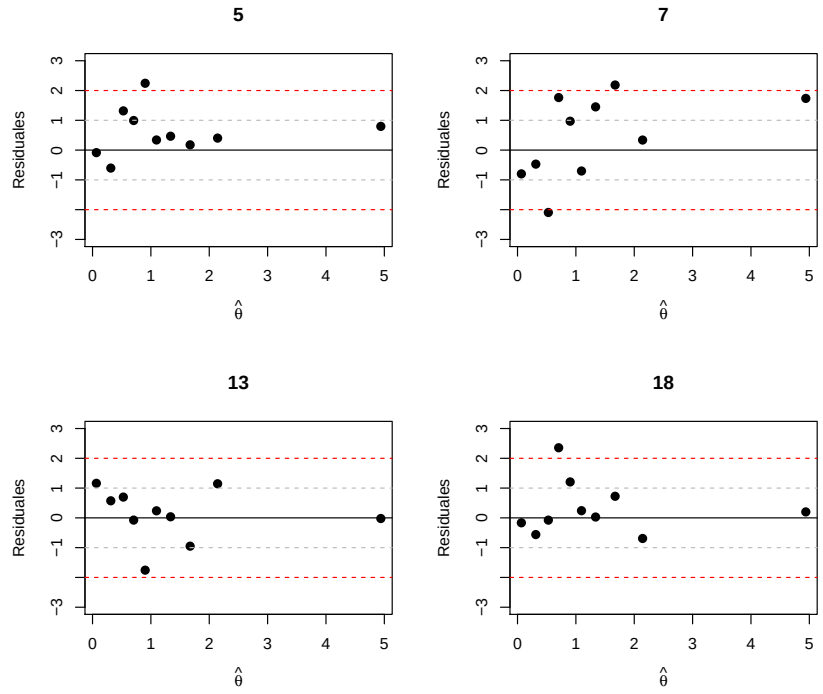


FIGURA 3.6. Residuales Estandarizados para algunos ítems

j	1	2	3	4	5	6	7	8	9	10
Cor	0.996	0.995	0.988	0.993	0.994	0.966	0.993	0.993	0.995	0.972

j	11	12	13	14	15	16	17	18	19	20
Cor	0.974	0.984	0.934	0.991	0.979	0.987	0.981	0.994	0.991	0.995

TABLA 3.6. Correlaciones entre frecuencias esperadas y observadas para cada uno de los ítems del test

Se construyó también la curva característica del test, y se compararon las frecuencias esperadas y observadas tal y como se indica en la Gráfica 3.7.

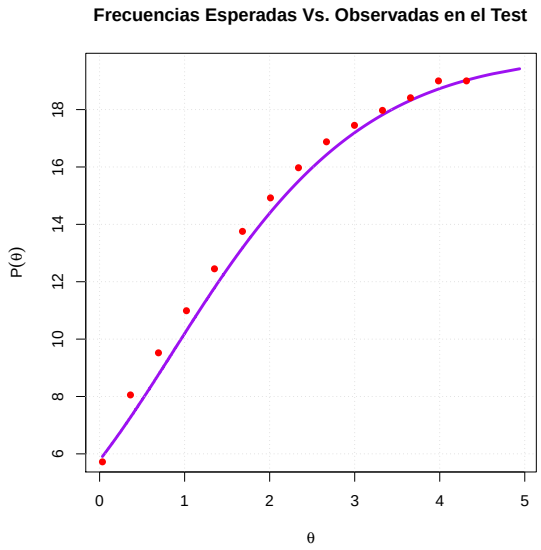


FIGURA 3.7. Frecuencias Esperadas Vs. Observadas en el Test

Para llevar a cabo la prueba de bondad de ajuste para los ítems se utilizó la metodología propuesta en Bock (1972) que constituye una variación de la estadística propuesta en Yen (1981) que es la utilizada con mayor frecuencia en la literatura y que se describió en la Sección 2.3.

Una vez efectuada la prueba con una significancia de $\alpha = 0.05$, se obtuvo que las curvas características de los ítems 2, 19 y 20 no son ajustados adecuadamente por el modelo propuesto.

CAPÍTULO 4

Aplicación a Datos Reales

4.1. Descripción de los Datos

Los datos que se analizarán en este capítulo corresponden al conjunto de Pruebas aplicadas durante el proceso de admisión de 2006 en la Facultad de Humanidades y Educación de la Universidad Central de Venezuela. La base de datos consta de 5565 registros en los que se evaluaron variables relacionadas con el éxito escolar y conductas que afectan el rendimiento académico [1], así como variables relacionadas con razonamiento verbal y conocimientos generales.

El conjunto de datos fue previamente analizado por Páyer (2013), con la finalidad de determinar la relación entre algunos rasgos de personalidad específicos y los puntajes obtenidos en pruebas de razonamiento.

Los constructos diseñados para medir rasgos de personalidad en los 5665 registros corresponden al Inventario PIHEMA R-2 (Ponce & Granell, 2006), que es un test psicométrico aplicado a estudiantes mayores de 15 años que aspiraban ingresar a la Facultad de Humanidades y Educación de la Universidad Central de Venezuela en el año 2006. El conjunto de pruebas indaga sobre rasgos y/o características que pueden asociarse tanto con el rendimiento escolar como con factores que interfieren en el mismo.

Como se describe en Páyer (2013), el test se aplica de manera escrita e individual con un tiempo límite de 30 minutos en los que el evaluado responde a 160 ítems con tres categorías de respuesta cada uno: verdadero, falso, no sabe. En Ponce & Granell (2006) se indica que el test aporta información en tres dimensiones:

- El primer constructo está relacionado con índices de éxito académico, motivación y estabilidad emocional.
- El segundo constructo aporta información sobre rasgos de personalidad que afectan el desarrollo académico: neuroticismo, indicios de psicopatía y paranoia[2].
- El tercer constructo observa la reacción del individuo frente a situaciones que resultan socialmente cuestionables. Textualmente, Ponce & Granell (2006) definen esta

[1] Los rasgos conductuales fueron evaluados a través del Inventario PIHEMA Ponce & Granell (2006)

[2] La descripción de cada uno de los rasgos de personalidad puede consultarse en Ponce & Granell (2006)

tercera dimensión del test como "*un conjunto de declaraciones que implican posiciones socialmente deseables pero imposibles de cumplir. Esta escala detecta, pues, un intento deliberado y bastante ingenuo de la persona de identificarse de una manera favorable y reflejar una imagen socialmente deseable*".

4.2. Selección del Constructo y Análisis Preliminar

Se construyeron los histogramas para los scores obtenidos por los individuos en cada uno de los constructos, los resultados se presentan en la gráfica 4.1, allí puede observarse que la mayoría de los constructos presentan puntajes con distribución asimétrica. Finalmente se seleccionó la variable *Sinceridad* que está relacionada con el tercer constructo de las pruebas, para implementar el procedimiento de estimación. Así pues, el test asociado consta de $p = 19$ ítems para los $N = 5665$ individuos.

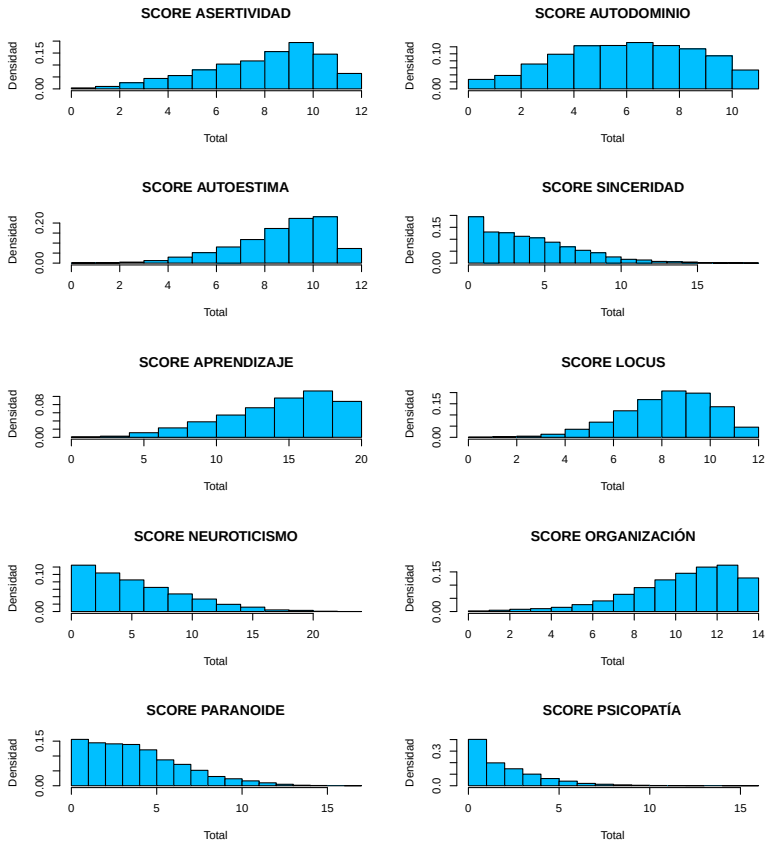


FIGURA 4.1. Scores por Constructo

Una vez seleccionado el conjunto de datos se realiza un análisis preliminar para determinar la dimensión del test, detectar ítems problemáticos y eliminar los registros con scores perfectos y nulos.

En la Gráfica 4.2 se representan los autovalores de los componentes obtenidos con el análisis de componente principales; puede observarse que el 19.5 % de la varianza total está representado por el autovalor del primer componente y que ésta es 2.2 veces la varianza

explicada por el segundo componente. El coeficiente de Cronbach fue $\alpha = 0.751$, además se construyó la curva de Cronbach–Mesbah en la Gráfica 4.3, donde se evidencia que el ítem 12 afecta la unidimensionalidad del test por lo que será retirado.

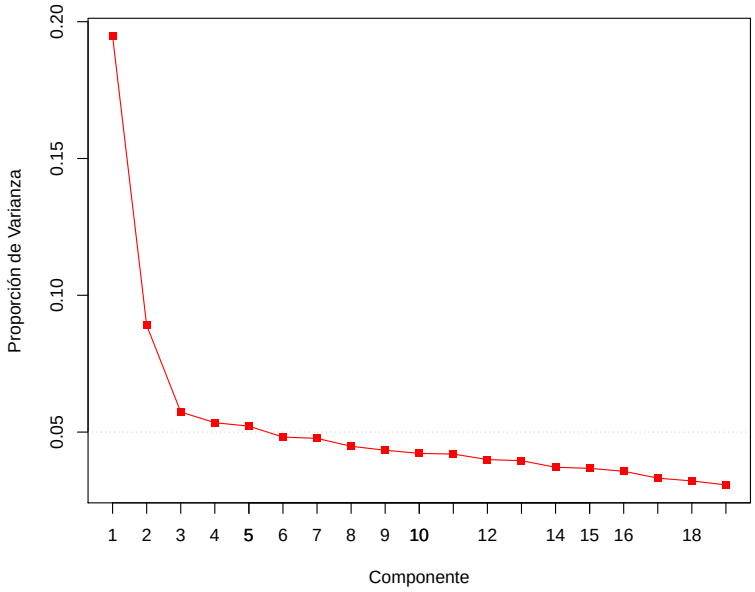


FIGURA 4.2. Proporción de varianza representada por autovalor en cada componente

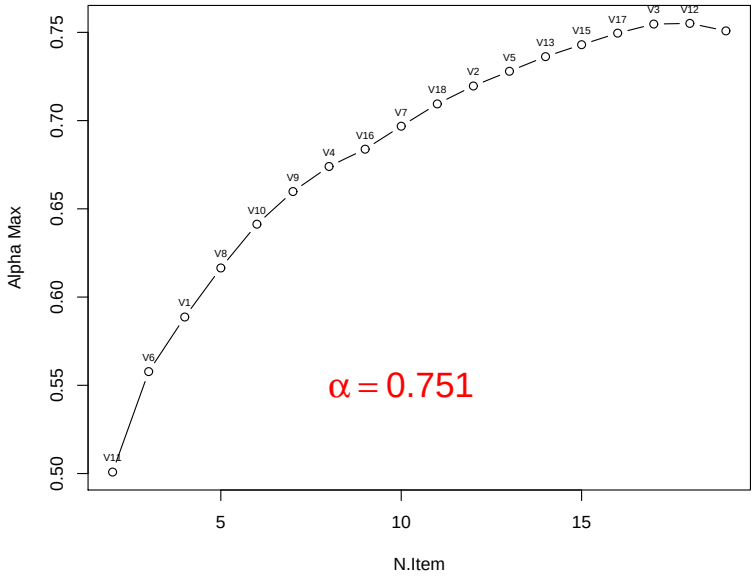


FIGURA 4.3. Curva de Cronbach-Mesbah

Después de eliminar los ítems problemáticos, se realizó nuevamente el ACP: esta vez, la varianza explicada por el primer componente fue de 20.5 %, y la razón entre los dos primeros autovalores se mantuvo en 2.2.

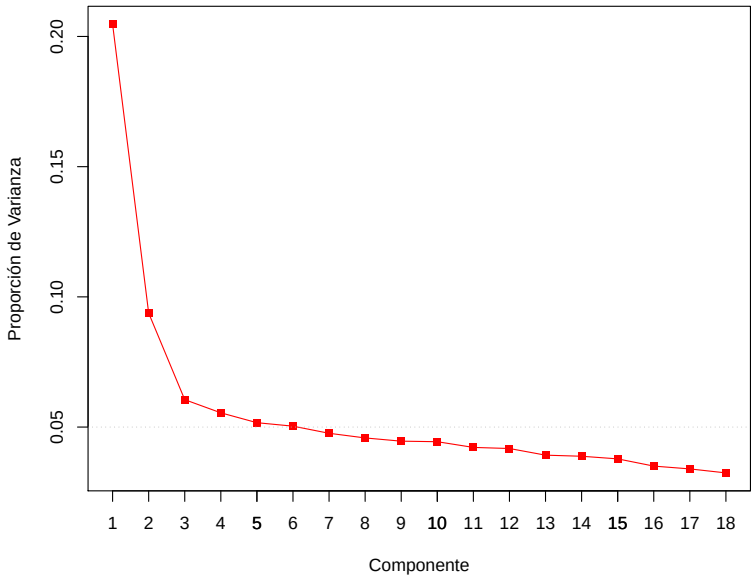


FIGURA 4.4. Proporción de varianza representada por autovalor en cada componente

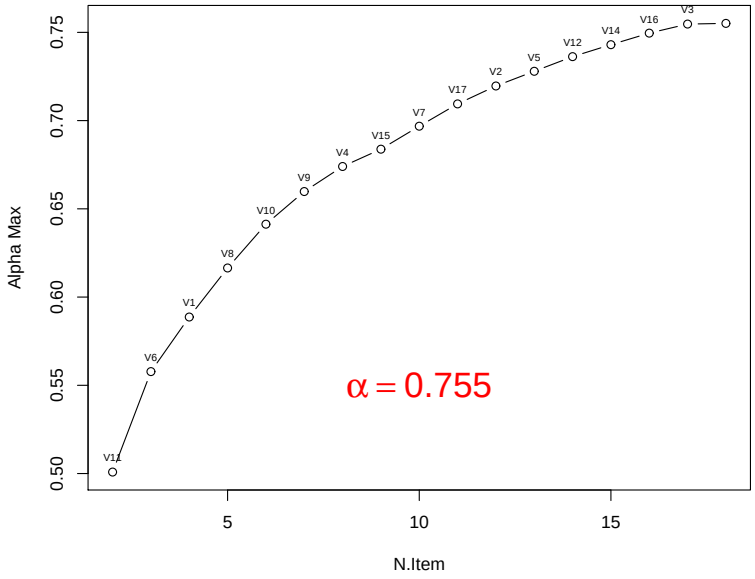


FIGURA 4.5. Curva de Cronbach-Mesbah

El coeficiente de Cronbach aumentó a $\alpha = 0.755$ y la curva de Cronbach–Mesbah resultó completamente creciente, como se detalla en la Gráfica 4.5.

En base a estos resultados se concluye que el test correspondiente a la variable *Sinceridad* es esencialmente unidimensional.

De acuerdo con el análisis preliminar de los datos, se supone que el test puede ser caracterizado a través de un modelo TRI logístico de dos parámetros en donde el trazo latente sigue una distribución asimétrica. Con el fin de determinar cuál es la distribución que mejor describe el comportamiento del trazo se ajustaron modelos aditivos generalizados para localización, escala y forma (GAMLSS), considerando como variable observada al primer componente obtenido del ACP inicial cuya correlación con el score es de 0.9714, (Rigby & Stasinopoulos (2005), Stasinopoulos & Rigby (2007)).

La función `fitDist` del paquete `gamlss` de `R` (Stasinopoulos & Rigby, 2007), permite determinar las posibles distribuciones que modelan un determinado conjunto de datos utilizando criterios de información (GAIC); en particular se consideraron como candidatas para describir al trazo latente distintas distribuciones asimétricas con soporte en $\mathbb{R}^+$. En la tabla 4.1 se presentan las distribuciones sugeridas y su correspondiente GAIC[3].

GIG	BCPEo	GG	GA	WEI3	BCCGo
18244	18255	18274	18275	18277	18331

BCTo	EXP	PARETO2	LOGNO	IG	IGAMMA
18333	18597	18599	18850	19535	20684

TABLA 4.1. Posibles Distribuciones para Modelar el Trazo Latente

Debido a la necesidad de utilizar procedimientos de cuadratura durante la estimación de los parámetros, se seleccionó la distribución Gamma para modelar el trazo latente. Una vez especificada la distribución se ajustó un modelo GAMLSS sin covariables para determinar los parámetros de la misma, considerando como variable observada al primer componente. En la Gráfica 4.6 y en la Tabla 4.2 se encuentra el diagnóstico del modelo ajustado.

De acuerdo a las estadísticas presentadas en la Tabla 4.2, los residuales siguen una distribución normal estándar, y el coeficiente de correlación de Filliben es cercano a 1, lo que permite concluir que la distribución gamma es un buen modelo para el conjunto de datos.

Ahora bien, la parametrización de la distribución gamma que se emplea en el paquete `gamlss` tiene media μ y varianza $\sigma^2\mu^2$, por lo que, en términos de la función (2.1), $\alpha = 1/\sigma^2$ y $\beta = \mu\sigma^2$. Así pues, después del ajuste se obtuvo que $\alpha = 1.3967$ y $\beta = 1.5691$ como parámetros de forma y escala respectivamente para la distribución gamma.

En resumen, para el proceso de estimación se asume que $\theta_i \sim GA(\alpha = 1.3967, \beta = 1.5691)$.

[3]Para la descripción de las distribuciones y sus parametrizaciones, revisar Rigby & Stasinopoulos (2005) y Stasinopoulos & Rigby (2007)

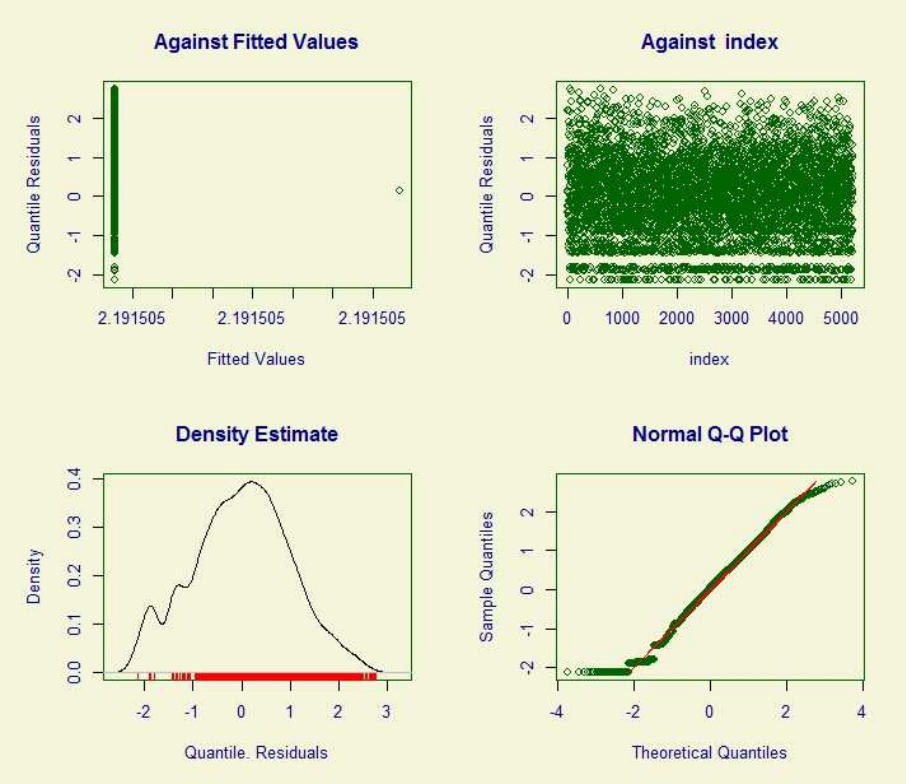


FIGURA 4.6. Gráficas de Diagnóstico para el modelo GAMLSS ajustado con la distribución Gamma

Media	-4.7869e-05
Varianza	0.99981
Coef. Asimetría	-0.021542
Coef. Curtosis	2.6114
Coef. Correlación Filliben	0.9961

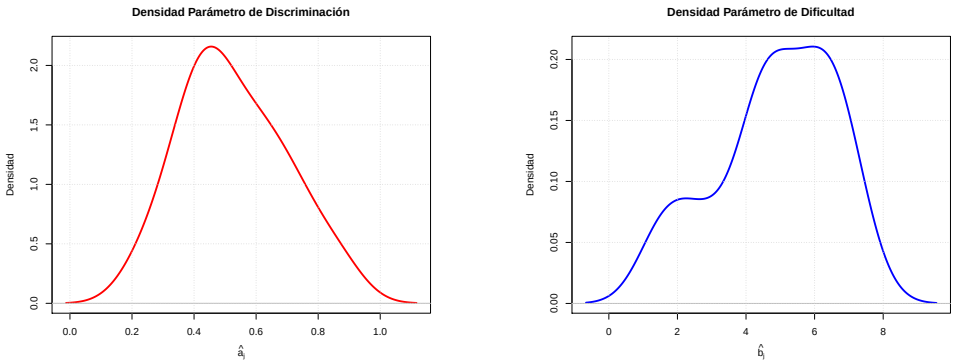
TABLA 4.2. Estadísticas de resumen para el modelo GAMLSS ajustado con la distribución Gamma

4.3. Estimación

4.3.1. Estimación de los Parámetros de los Ítems

El constructo de Sinceridad consta de $N = 5210$ respondientes y $p = 18$ ítems después de haber eliminado los ítems problemáticos y los scores nulos y perfectos. Para la estimación de los parámetros de los ítems se implementará el algoritmo EM descrito en el Capítulo 2, utilizando $q = 20$ puntos de cuadratura cuyos pesos están determinados por la distribución $GA(\alpha = 1.3967, \beta = 1.5691)$ obtenida en la Sección 4.2.

Las estimaciones y sus respectivos errores estándar se presentan en la Tabla 4.3, mientras que en la Gráfica 4.7 se encuentran las densidades de los parámetros estimados.



(a) (b)

FIGURA 4.7. (a) Densidad $\widehat{a}_j$ (b) Densidad $\widehat{b}_j$

j	$\widehat{a}_j$	$sd_{\widehat{a}}$	$\widehat{b}_j$	$sd_{\widehat{b}}$
1	0.452	0.018	4.656	0.107
2	0.450	0.022	7.285	0.217
3	0.231	0.016	5.415	0.240
4	0.712	0.030	6.550	0.131
5	0.411	0.020	6.798	0.204
6	0.639	0.022	4.386	0.075
7	0.587	0.025	6.613	0.151
8	0.500	0.020	5.361	0.119
9	0.874	0.034	5.837	0.091
10	0.495	0.019	4.406	0.093
11	0.558	0.019	3.094	0.062
12	0.380	0.017	1.601	0.079
13	0.786	0.024	1.625	0.042
14	0.353	0.017	4.348	0.123
15	0.704	0.030	6.500	0.130
16	0.316	0.016	4.524	0.143
17	0.436	0.019	5.956	0.156
18	0.612	0.020	2.554	0.053

TABLA 4.3. Estimaciones y errores estándar para los parámetros de los ítems

4.3.2. Estimación del Trazo

Para la estimación del trazo se emplearon dos metodologías descritas en la Sección 2.2.2: en primer lugar se llevó a cabo el procedimiento de optimización de la verosimilitud dada en (1.4) mediante la función `optimize` de `R`, y en segundo lugar se llevó a cabo la estimación a través del método EAP implementando la función `eap` de la librería `irtoys` en `R`, [4].

[4]La documentación del paquete puede consultarse en Partchev (2009)

Los principales estadísticos de resumen del trazo estimado con máxima verosimilitud se encuentran en la Tabla 4.4, en tanto que los resultados obtenidos con EAP se presentan en la Tabla 4.5, vale la pena mencionar que la correlación entre los trazos estimados por los dos métodos es de 0.9784.

mín	Q_1	Q_2	$\overline{x}$	Q_3	máx	sd
0.0335	0.3420	2.0000	2.2100	3.4700	9.9700	1.929

TABLA 4.4. Estadísticas de resumen para $\widehat{\theta}_i$ obtenido con máxima verosimilitud

mín	Q_1	Q_2	$\overline{x}$	Q_3	máx	sd
0.724	1.050	1.700	2.180	2.870	10.400	1.473

TABLA 4.5. Estadísticas de resumen para $\widehat{\theta}_i$ obtenido con EAP

Una vez estimado el trazo latente, se construyó la función de información pertinente para obtener el error estándar de la estimación. En las Tablas 4.6 y 4.7 se presentan los estadísticos descriptivos para $SE(\widehat{\theta})$ obtenido con máxima verosimilitud y EAP respectivamente, se debe destacar que el error estándar de la estimación máximo verosímil presenta valores más bajos que el error estándar de la estimación EAP.

mín	Q_1	Q_2	$\overline{x}$	Q_3	máx	sd
0.959	1.060	1.180	1.240	1.480	1.730	0.216

TABLA 4.6. Estadísticas de resumen para $SE(\widehat{\theta}_i)$ con MV

mín	Q_1	Q_2	$\overline{x}$	Q_3	máx	sd
0.959	1.100	1.220	1.200	1.320	1.920	0.129

TABLA 4.7. Estadísticas de resumen para $SE(\widehat{\theta}_i)$ con EAP

4.4. Comparación con otros Paquetes

Para el mismo conjunto de datos se realizó el proceso de estimación empleando la librería `ltm` de `R`, cuya documentación puede revisarse en Rizopoulos (2006) y en el que se asume que la distribución del trazo es normal estándar. Una vez realizada la estimación de los parámetros de los ítems se obtuvieron los errores estándar respectivos y se compararon con los resultados obtenidos en la Sección 4.3.1.

j	a.em	sda.em	a.ltm	sda.ltm	b.em	sdb.em	b.ltm	sdb.ltm
1	0.451	0.018	0.850	0.047	4.656	0.107	1.370	0.072
2	0.450	0.022	0.981	0.062	7.285	0.217	2.456	0.126
3	0.231	0.015	0.404	0.038	5.415	0.239	1.862	0.179
4	0.712	0.030	1.708	0.098	6.550	0.131	2.048	0.072
5	0.411	0.020	0.858	0.054	6.798	0.203	2.298	0.124
6	0.640	0.022	1.228	0.058	4.386	0.075	1.232	0.050
7	0.587	0.025	1.313	0.075	6.613	0.151	2.134	0.087
8	0.501	0.020	0.948	0.052	5.361	0.119	1.738	0.083
9	0.874	0.034	2.119	0.119	5.837	0.091	1.761	0.053
10	0.495	0.019	0.920	0.048	4.406	0.093	1.255	0.063
11	0.558	0.019	0.986	0.049	3.094	0.062	0.576	0.040
12	0.380	0.017	0.594	0.040	1.602	0.079	-0.333	0.054
13	0.786	0.023	1.192	0.057	1.625	0.042	-0.270	0.031
14	0.353	0.016	0.583	0.040	4.348	0.123	1.332	0.097
15	0.704	0.030	1.644	0.093	6.500	0.130	2.052	0.073
16	0.316	0.016	0.560	0.040	4.524	0.143	1.349	0.101
17	0.436	0.019	0.826	0.050	5.956	0.156	2.043	0.109
18	0.612	0.020	1.026	0.050	2.554	0.053	0.286	0.035

TABLA 4.8. Estimaciones de los parámetros de los ítems con la metodología propuesta y usando el paquete ltm

Para la comparación de la estimación del trazo latente, se emplearon los resultados obtenidos en la Sección 4.3.2 con el método de máxima verosimilitud y los resultados de la estimación con `ltm`. En la Tabla 4.9 se presentan las estadísticas de resumen para las estimaciones del trazo.

Medida	Método Propuesto	`ltm`
mín	0.0335	-1.1967
Q_1	0.3422	-0.5992
Q_2	2.0011	-0.0208
$\overline{x}$	2.2123	0.0538
Q_3	3.4725	0.5890
máx	9.9667	2.9822
sd	1.9292	0.8097

TABLA 4.9. Resumen estadístico para el trazo estimado con MV y ltm

En la Gráfica 4.8 puede observarse la comparación entre las frecuencias observadas y esperadas después del ajuste para algunos ítems con ambas metodologías: las tres gráficas superiores corresponden a los resultados obtenidos con `ltm` para los ítems 1,9 y 18 del test, mientras que las tres inferiores se obtuvieron con la metodología propuesta en este documento para los mismos tres ítems.

Cabe destacar que las frecuencias observadas parecen ajustarse mejor a las curvas características de los ítems del test con trazo asimétrico que con las estimaciones obtenidas con `ltm`.

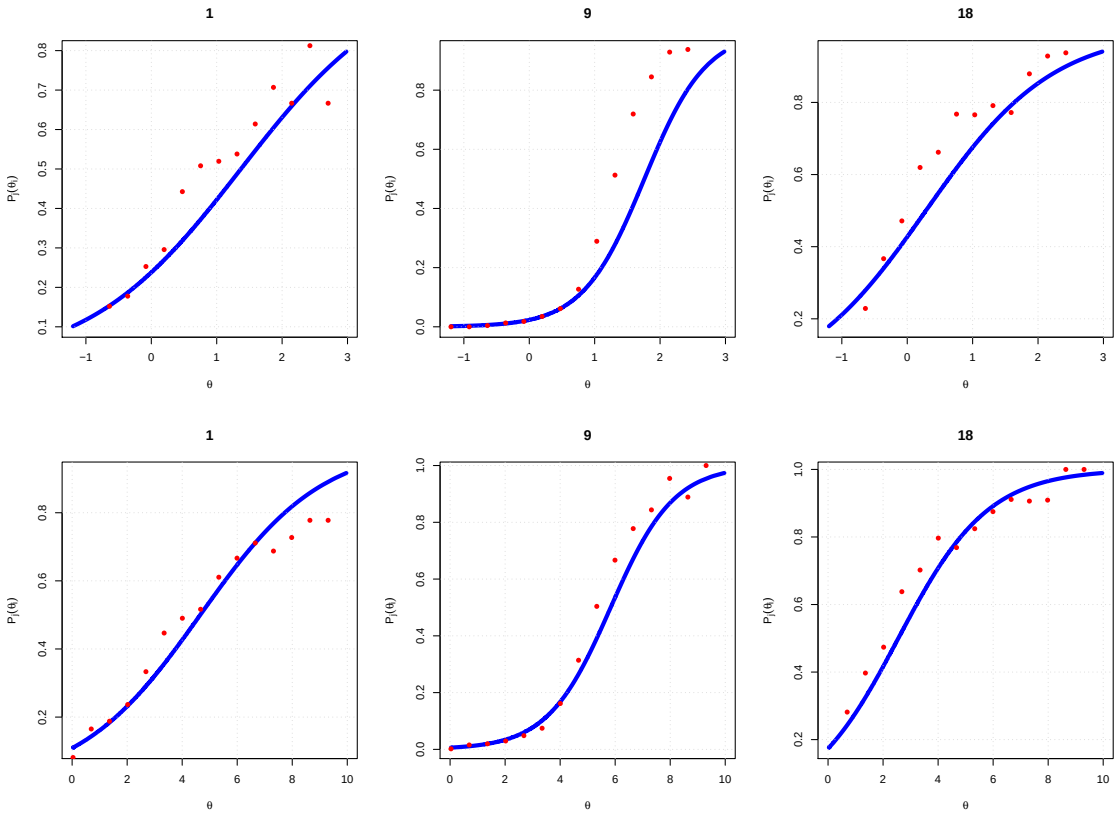


FIGURA 4.8. Comparación ltm vs. metodología propuesta

Adicionalmente, en la Gráfica 4.9 se presentan las frecuencias observadas y esperadas del test empleando ambas estimaciones: la gráfica del lado izquierdo corresponde a la propuesta de este documento, mientras que la gráfica del lado derecho se obtuvo con la estimación del paquete `ltm` de `R`; resulta evidente que asumir un trazo normal estándar en el constructo conlleva a un ajuste inadecuado.

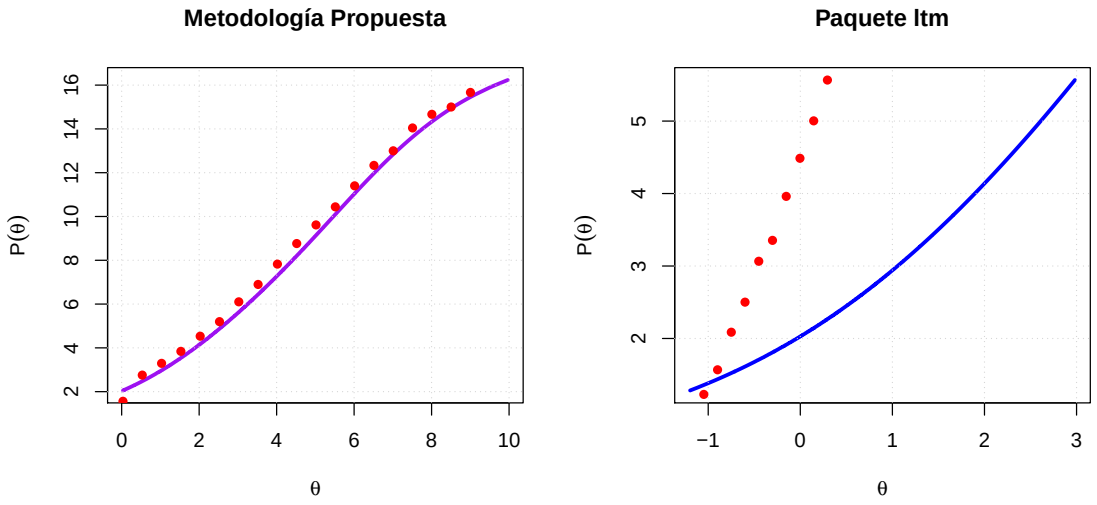


FIGURA 4.9. Frecuencias esperadas Vs. observadas del test

Usando el criterio de información de Akaike (AIC), se obtuvo que para la metodología propuesta $AIC = 80078$, mientras que, con el paquete `ltm` el $AIC = 80925$, lo que indica que la suposición de una distribución Gamma para el trazo proporciona un mejor ajuste.

APÉNDICE A

Código R

A continuación se presenta el código implementado en R para llevar a cabo el procedimiento descrito en la Sección 2.4.

A.1. Estimación Parámetros de los Ítems

```
# N: Número de individuos
# p: Número de ítems
# q: Número de nodos
# Parámetros de la función de pesos w(r)=Gamma(alfa,beta)
# x.k:Nodos del proceso de cuadratura
# Ax.k: Pesos del proceso de cuadratura
# cuadratura: arreglo de los nodos y pesos de cuadratura
# patrones: patrones de respuesta del test
# s: número de patrones de respuesta
# r: frecuencia de los patrones de respuesta
# a: parámetro de discriminación del modelo
# b: parámetro de dificultad del modelo


# Valores iniciales de los parámetros de los ítems

d.inicial<-c(rep(0,p))
a.inicial<-c(rep(1,p))

#Matriz de probabilidades de respuesta correcta para cada patrón con cada uno de los nodos
#de cuadratura.

eta<-c()
PQJ<-matrix(nrow = q,ncol = p)
mmm<-array(NA,dim=c(q,p,s))
for(k in 1:q){
  for(j in 1:p){
    eta[k]<--(a.inicial[j]*x.k[k]-d.inicial[j])
    for(l in 1:s){
      if(patrones[l,j]==1){
        PQJ[k,j]=1/(1+exp(eta[k]))
      }
```

```
      else{
        PQJ[k,j]=1-(1/(1+exp(eta[k])))
      }
      PQJ[k,j]=ifelse(PQJ[k,j]<1e-5,1e-5,PQJ[k,j])
      PQJ[k,j]=ifelse(PQJ[k,j]>0.99,0.99,PQJ[k,j])
      mmm[k,j,l]=array(PQJ[k,j])
    }
  }
}

# Almacena la probabilidad conjunta del patrón k para cada nodo.
prod.patron<-c()
L.x.k<-matrix(nrow = q,ncol = s)
for(k in 1:q){
  for(l in 1:s){
    prod.patron[k]<-prod(mmm[k,,l])
    L.x.k[k,l]=prod.patron[k]
  }
}

# Construcción de las expresiones 1.13
ver <- list()
for(l in 1:s){
  ver[[l]]<-cbind(L.x.k[,l],Ax.k,L.x.k[,l]*Ax.k)
}

P.l<-c()
for(i in 1:s){
  P.l[i]<-sum(ver[[i]][,3])
}

tablas<- list()
for(k in 1:q){
  tablas[[k]]<-cbind(r,L.x.k[k,],Ax.k[k],P.l,(r*L.x.k[k,]*Ax.k[k])/P.l)
}

n.j.k<-c()
for(k in 1:q){
  n.j.k[k]<-sum(tablas[[k]][,5])
}

nnn<-array(NA,dim=c(s,q,p))
tablas3<-matrix(nrow = s,ncol = q)
for(k in 1:q){
  for(j in 1:p){
    tablas3[,k]<-(patrones[,j]*r*L.x.k[k,]*Ax.k[k])/P.l

    nnn[,k,j]=array(tablas3[,k])
  }
}

r.j.k<-matrix(nrow=q,ncol=p)
for(j in 1:p){
  r.j.k[,j]<-apply(nnn[,,j], 2, sum)
}


# Tabla con los valores de n y r para los p ítems y los k nodos.

n.r<-cbind(r.j.k,n.j.k)
```

```
# Función a maximizar: Ecuación 2.7

Fun<-function(a,d){
  eta<--(a*x.k-d)
  P<-1/(1+exp(eta))
  P=ifelse(P<1e-5,1e-5,P)
  P=ifelse(P>0.99,0.99,P)
  return(sum(r.j.k[,j]*log(P)+(n.j.k-r.j.k[,j])*log(1-P)))

}

# Estimación con los valores iniciales
parametros<-matrix(nrow = p,ncol=2)
for(j in 1:p){
  parametros[j,]<-matrix(optim(par=c(a.inicial[j],d.inicial[j]),
                 (function(x){Fun(x[1],x[2])}),control=list(fnscale=-1))$par)
}
parametros

# Se realiza el procedimiento iterativo con el criterio de parada
# requerido para obtener las estimaciones.
```

A.2. Estimación del Trazo Latente

```
# y: matriz de datos eliminando registros con scores nulos y perfectos
# score: puntaje por individuo
# a: parámetro de discriminación estimado
# b: parámetro de dificultad estimado

score<-apply(y,1,sum)
summary(score)
score1<-(score-mean(score))/sd(score)
score.es<-score1-min(score1)+0.01

d<-a*b

t<-score.es
l.t<-length(t)

# Función de verosimilitud

Fun<-function(t,a,d,y){
  eta<--(a*t-d)
  P<-1/(1+exp(eta))
  P=ifelse(P<1e-5,1e-5,P)
  P=ifelse(P>0.99,0.99,P)
  z=ifelse(y==1,log(P),log(1-P))
  return(sum(z))
}
```

```
# Procedimiento de optimización

trazo<-numeric(length = dim(y)[1])
for(i in 1:dim(y)[1]){
  if(i%%10==0) print(i)
  trazo[i]<-optimize(Fun,c(0,20),a=a,d=d,y=y[i,],maximum =TRUE)$maximum
}
```

APÉNDICE B

Resultados Simulación

B.1. Análisis Preliminar

Se genera un test dicotómico con el que se evalúan $p = 20$ ítems a $N = 5000$ individuos y en el que la probabilidad de que el individuo i acierte al ítem j está dada por la Ecuación (3.1). La distribución del parámetro de discriminación es $LN(0, 1/4)$, mientras que el parámetro de dificultad se genera a partir de una distribución $GA(2, 1/2)$. Adicionalmente, se considera que la distribución del trazo es $GA(2, 1/2)$; algunas medidas descriptivas de los parámetros simulados se presentan en la Tabla 3.1.

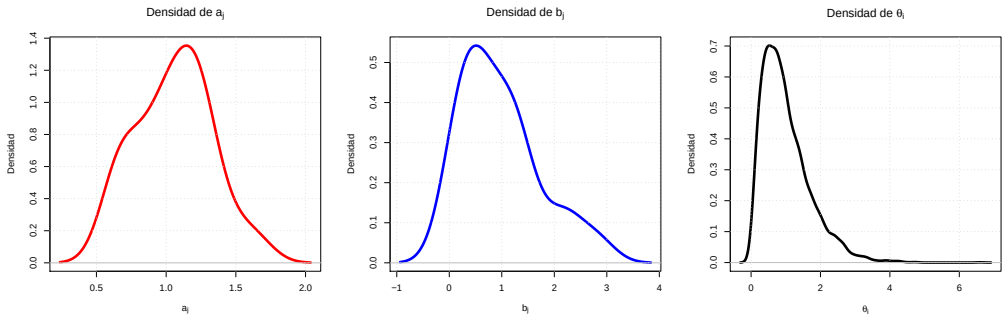
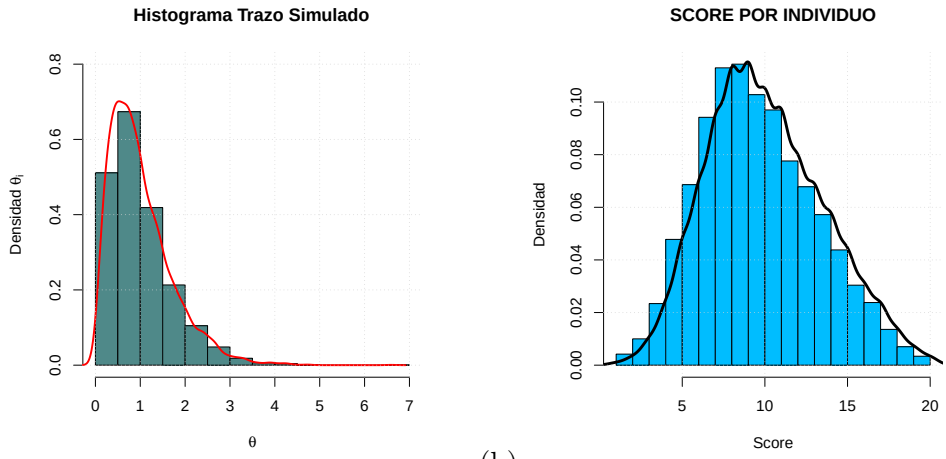


FIGURA B.1. Densidades de los parámetros simulados

En la Gráfica B.2 se presentan los histogramas del trazo latente simulado y del score del test, en ambos casos, la asimetría de sus distribuciones es característica.

Después de eliminar del test los registros con puntuaciones nulas y perfectas se procede con la recuperación de los parámetros de los ítems mediante el algoritmo EM; para tal fin, se calculan los nodos y pesos de cuadratura utilizando como función de pesos a la distribución Gamma de parámetros $\alpha = 2$ y $\beta = 1/2$, los resultados se encuentran en la tabla B.1.



(a) (b)

FIGURA B.2. (a) Histograma Trazo Simulado (b) Histograma Score del Test

k	1	2	3	4	5	6	7	8	9	10
Nodos	34.19	28.78	24.63	21.20	18.24	15.66	13.38	11.35	9.55	7.94
Pesos	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

k	11	12	13	14	15	16	17	18	19	20
Nodos	6.51	5.24	4.13	3.16	2.33	1.64	1.07	0.62	0.29	0.09
Pesos	0.00	0.00	0.00	0.02	0.07	0.16	0.26	0.28	0.17	0.04

TABLA B.1. Pesos y nodos del procedimiento de cuadratura

B.2. Recuperación

Se obtuvieron 4286 patrones de respuesta con sus frecuencias respectivas. Como valores iniciales para los parámetros de los ítems se emplearon los propuestos por Harwell et al. (1988) y Baker (1985), es decir, $a_j = 1$ y $d_j = 0$. Sin embargo, vale la pena resaltar, que también podrían utilizarse como valores iniciales a los resultados de la estimación de una regresión logística con el primer componente como variable observada. En la Tabla B.2 se encuentran los resultados del ajuste mencionado, obtenidos mediante la función `glm` de `R`.

j	1	2	3	4	5	6	7	8	9	10
d_j	-1.98	-3.05	-2.23	-1.66	-5.68	-0.96	-1.42	-2.80	-2.58	-1.82
a_j	0.67	0.83	0.55	0.55	0.74	0.39	0.57	0.59	0.52	0.52

j	11	12	13	14	15	16	17	18	19	20
d_j	-1.31	-1.00	-1.04	-2.92	-1.22	-2.70	-3.78	-2.38	-3.50	-3.15
a_j	0.37	0.34	0.44	0.60	0.35	0.41	0.58	0.46	0.64	0.71

TABLA B.2. Valores iniciales para los parámetros de los ítems

Los valores de los parámetros recuperados tras 22 iteraciones se presentan en la Tabla 3.2.

Para la obtención del trazo por máxima verosimilitud, se emplearon como valores iniciales los scores estandarizados de cada individuo, sin embargo, también podrían considerarse los resultados del primer componente del ACP, cuyos principales descriptivos se encuentran en la Tabla B.3 y cuya correlación con el escore es de 0.993. El resumen del trazo recuperado se encuentra en la Tabla 3.3.

mín	Q_1	Q_2	$\overline{x}$	Q_3	máx	sd
0.10	3.10	4.17	4.32	5.43	8.81	1.65

TABLA B.3. Resumen Descriptivo del primer componente del ACP

También se llevó a cabo la recuperación del trazo empleando el método EAP descrito en la Sección 2.2, los resultados obtenidos se resumen en la Tabla B.4.

mín	Q_1	Q_2	$\overline{x}$	Q_3	máx	sd
0.231	0.564	0.832	0.998	1.280	3.380	0.581

TABLA B.4. Resumen Descriptivo del Trazo Recuperado con EAP

La correlación entre el trazo recuperado con máxima verosimilitud y el obtenido con EAP es de 0.994, en la Gráfica B.3 se presentan su densidades correspondientes.

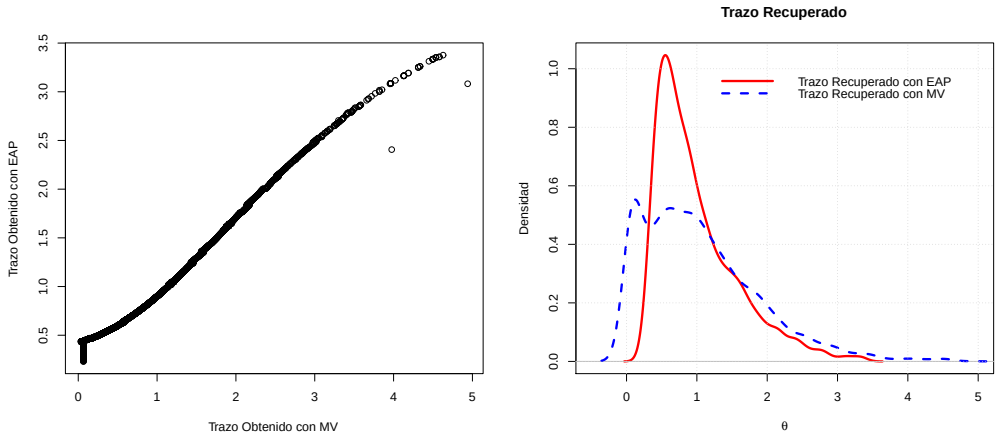


FIGURA B.3. Densidades de los parámetros simulados

Después de obtener los parámetros de los ítems y del trazo (empleando máxima verosimilitud), se construyeron las curvas características de los ítems y del test, como se observan en la Gráfica B.4.

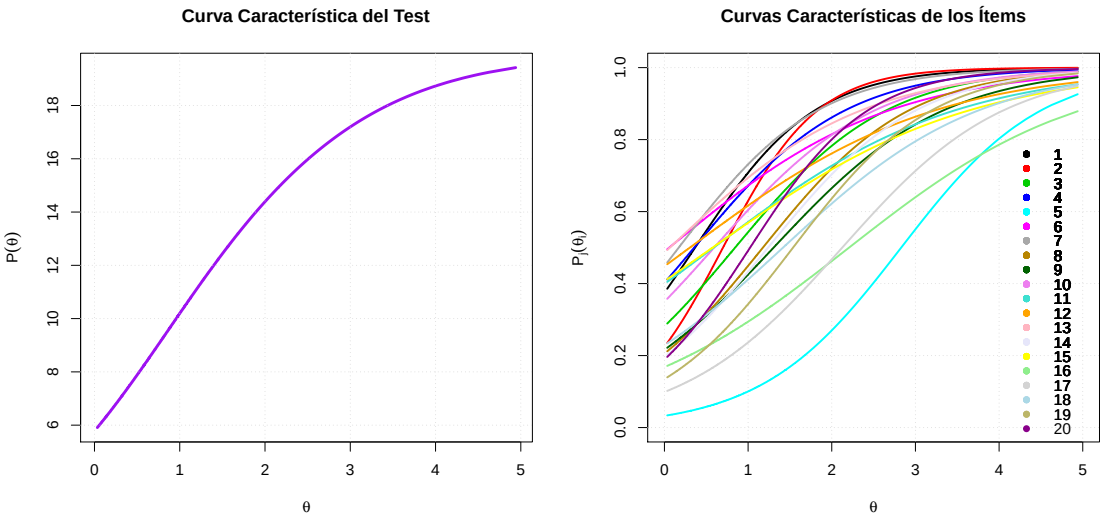


FIGURA B.4. Curvas Características del Test y de los Ítems respectivamente

B.3. Bondad de Ajuste

Una vez construidas las curvas características del test, se compararon la frecuencias esperadas y observadas utilizando tanto la recuperación máximo verosímil del trazo como la recuperación con EAP. Los resultados se encuentran en la Gráfica B.5.

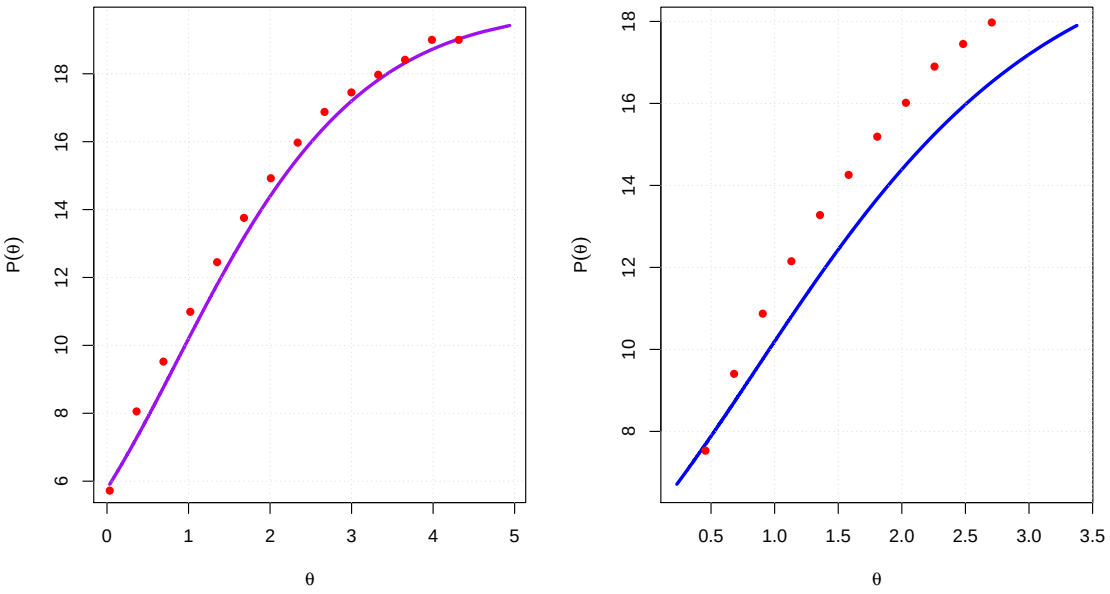


FIGURA B.5. Frecuencias esperadas vs. observadas: la gráfica del lado derecho corresponde al trazo obtenido con MV, la gráfica izquierda al trazo recuperado con EAP.

En las gráficas siguientes se presentan las comparaciones entre frecuencias esperadas y observadas para cada uno de los ítems del test, así como los residuales correspondientes.

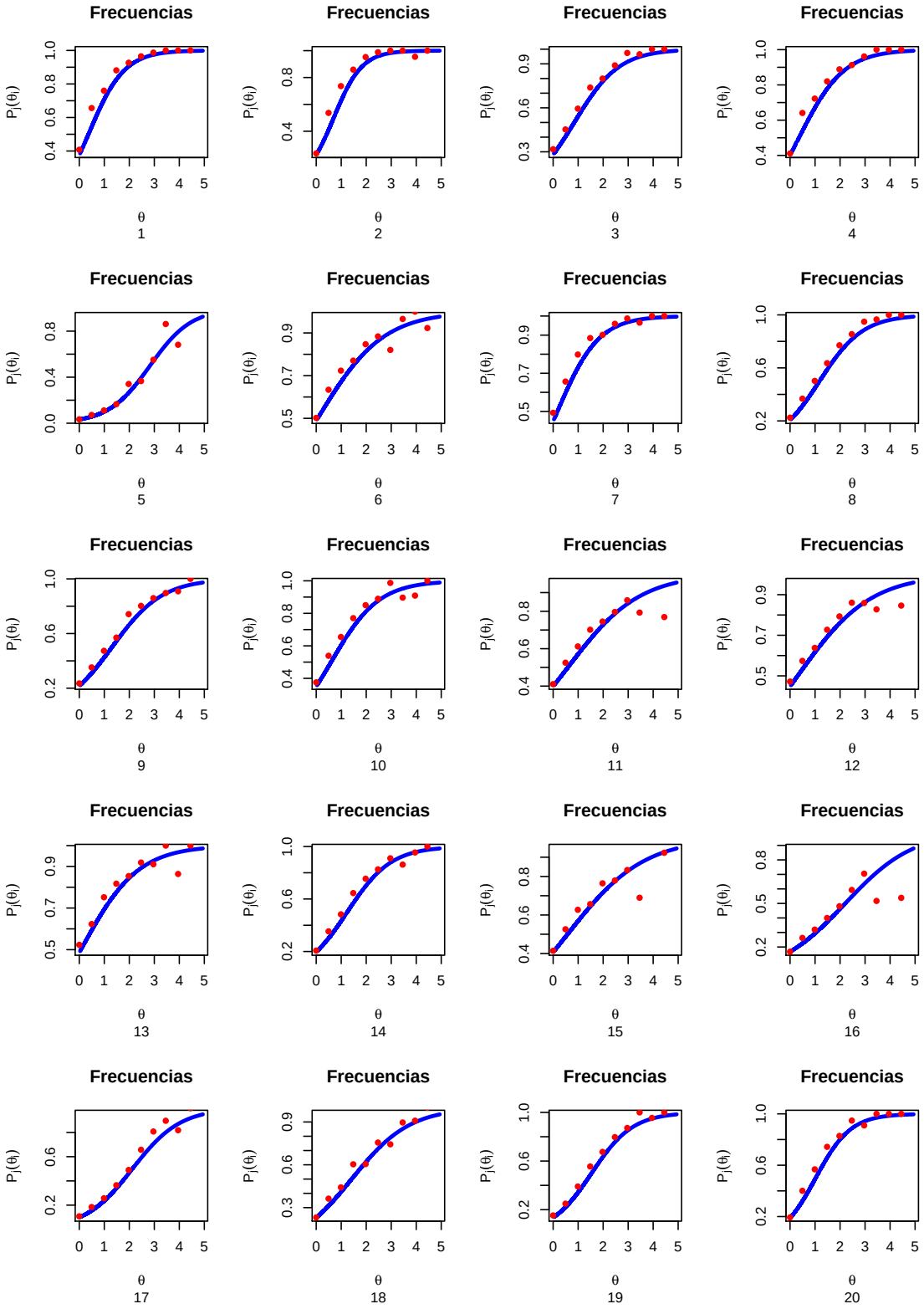


FIGURA B.6. Frecuencias esperadas vs. observadas

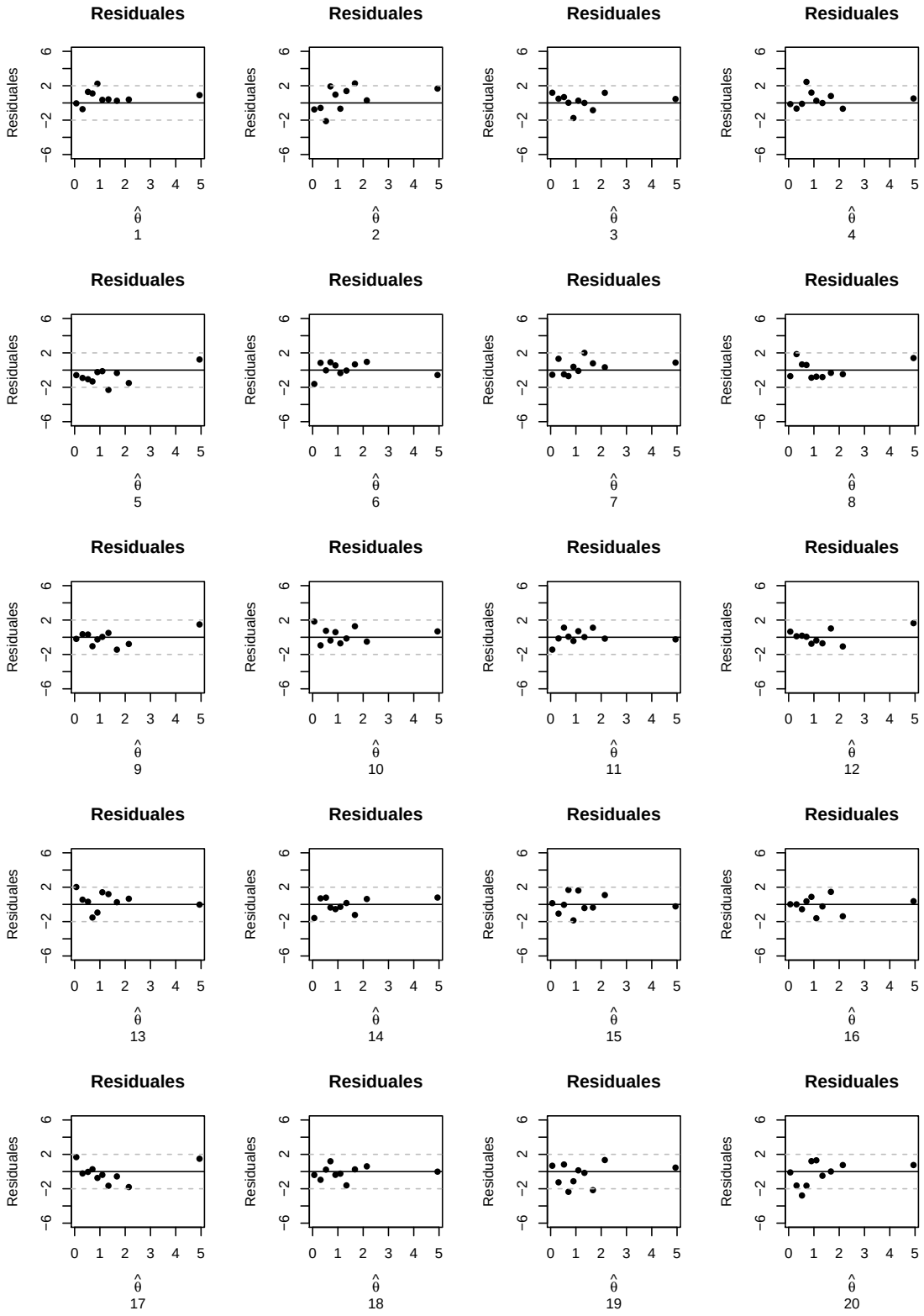


FIGURA B.7. Residuales estandarizados para cada uno de los ítems del test

Los valores del estadístico de prueba dado por la ecuación (2.21) y los P-valores correspondientes se presentan en la Tabla B.5.

j	1	2	3	4	5	6	7	8	9	10
Estadístico	9.73	20.39	7.59	9.32	13.33	6.25	8.41	9.04	6.58	8.17
P–Valor	0.28	0.01	0.47	0.32	0.10	0.62	0.39	0.34	0.58	0.42

j	11	12	13	14	15	16	17	18	19	20
Estadístico	5.34	6.56	11.66	6.70	11.61	8.02	12.16	5.78	16.24	17.67
P–Valor	0.72	0.58	0.17	0.57	0.17	0.43	0.14	0.67	0.04	0.02

Tabla B.5. Valores del estadístico calculado y p-valores para la prueba de Bondad de Ajuste

APÉNDICE C

Resultados Datos Reales

C.1. Selección de Datos

Dentro del conjunto de datos PIHEMA R-2 descrito en la Sección 4.1 se encontraron 10 constructos asociados con distintos rasgos de personalidad para los que se identificó el número de ítems como se indica en la Tabla C.1.

Constructos	N° de Ítems
Asertividad	12
Autodominio	11
Autoestima	12
Sinceridad	19
Aprendizaje	20
Locus	12
Neuroticismo	24
Organización	14
Paranoide	18
Psicopatía	18

Tabla C.1. Número de Ítems por Constructo

En la Gráfica C.1 se presentan los diagramas de cajas y bigotes para los scores de cada uno de los constructos de la prueba que se tuvieron en cuenta, junto con los histogramas de la Gráfica 4.1 para seleccionar el test objeto de análisis.

Debido a la distribución de los scores y al número de ítems, se seleccionó el constructo de Sinceridad para implementar la metodología de estimación propuesta en el documento.

Una vez seleccionado el conjunto de datos, se realiza un ACP para determinar la dimensión del test. Como se observa en la Gráfica C.2(a), el ítem 12 resulta problemático por lo que es retirado del test y se repite el ACP obteniendo la Gráfica C.2(b). Así pues, el test unidimensional consta de $N = 5665$ individuos y $p = 18$ ítems.

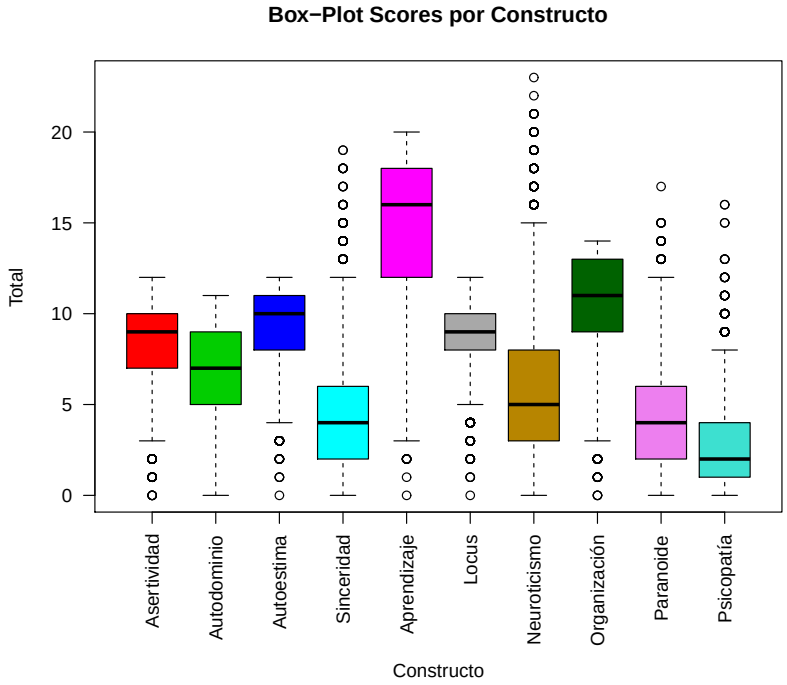
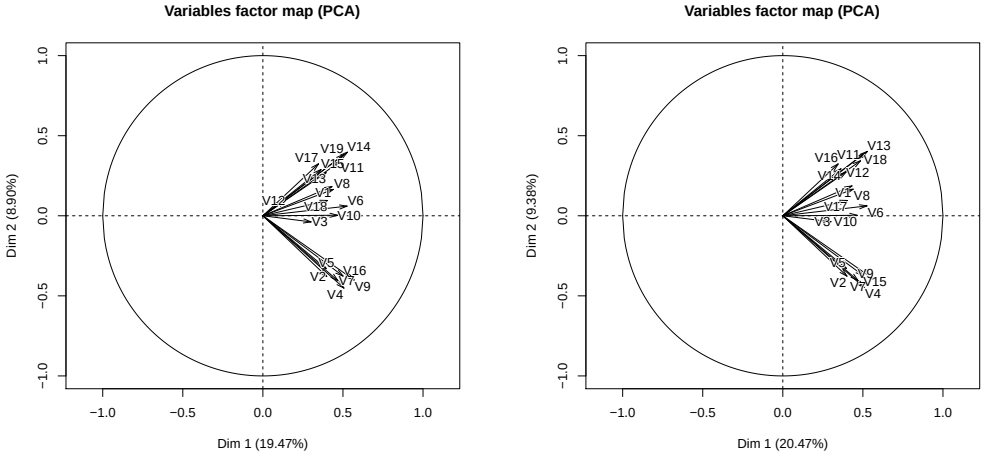


FIGURA C.1. Box-Plot para los scores de cada constructo.



(a) (b)

FIGURA C.2. (a) Resultados ACP Test Completo (b) Resultados ACP Test sin el Ítem 12

Además de eliminar los ítems que afectan la dimensionalidad del test, se retiraron los registros para los que el score es perfecto o nulo, de manera que el constructo de Sinceridad consta de $N = 5210$ evaluados y $p = 18$ ítems.

Empleando modelos GAMLSS con el primer componente del ACP como variable observada, se determinó que la distribución gamma es un buen modelo para los datos; las densidades del primer componente y de la distribución seleccionada se detallan en la Gráfica C.3.

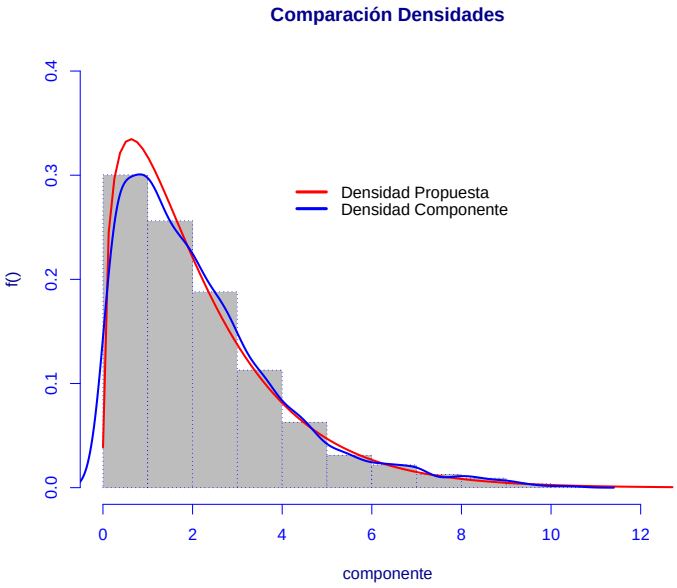


FIGURA C.3. Densidades Primer Componente y Distribución Gamma

Así pues, el trazo latente se modela con una distribución $Gamma(\alpha = 1.3967, \beta = 1.5691)$. Para estimar los parámetros de los ítems con el algoritmo EM, se emplearon $q = 20$ puntos de cuadratura determinando los nodos y los pesos con la distribución obtenida anteriormente. En la Tabla C.2 se encuentran los puntos y pesos de cuadratura.

k	1	2	3	4	5	6	7	8	9	10
Nodos	105.54	88.66	75.75	65.05	55.86	47.84	40.75	34.47	28.88	23.91
Pesos	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

k	11	12	13	14	15	16	17	18	19	20
Nodos	19.50	15.61	12.20	9.25	6.74	4.64	2.94	1.63	0.71	0.17
Pesos	0.00	0.00	0.00	0.01	0.04	0.11	0.21	0.29	0.24	0.09

TABLA C.2. Pesos y nodos del procedimiento de cuadratura

C.2. Estimación

Se identificaron 2671 patrones de respuesta con sus correspondientes frecuencias. Tal y como se ha indicado en secciones anteriores, pueden tomarse como valores iniciales $a_j = 1$ y $d_j = 0$ para $j = 1, 2, ..., p$, o bien los resultados del ajuste de una regresión logística con el primer componente del ACP como variable observada que se encuentran en la Tabla C.3.

j	1	2	3	4	5	6	7	8	9
d_j	-2.22	-3.66	-1.36	-5.45	-3.11	-3.02	-4.41	-2.80	-6.06
a_j	0.49	0.55	0.28	0.88	0.51	0.71	0.71	0.53	1.08

j	10	11	12	13	14	15	16	17	18
d_j	-2.39	-1.82	-0.67	-1.37	-1.56	-5.24	-1.47	-2.73	-1.67
a_j	0.57	0.60	0.42	0.85	0.36	0.84	0.33	0.47	0.67

TABLA C.3. Valores iniciales para la estimación de los parámetros de los ítems

Para la estimación del trazo se implementaron dos metodologías: máxima verosimilitud y EAP. Los valores de los parámetros obtenidos se presentan en las Tablas 4.3, 4.4 y 4.5.

En la Gráfica C.4 se detalla la curva característica del test, así como las curvas características de los $p = 18$ ítems.

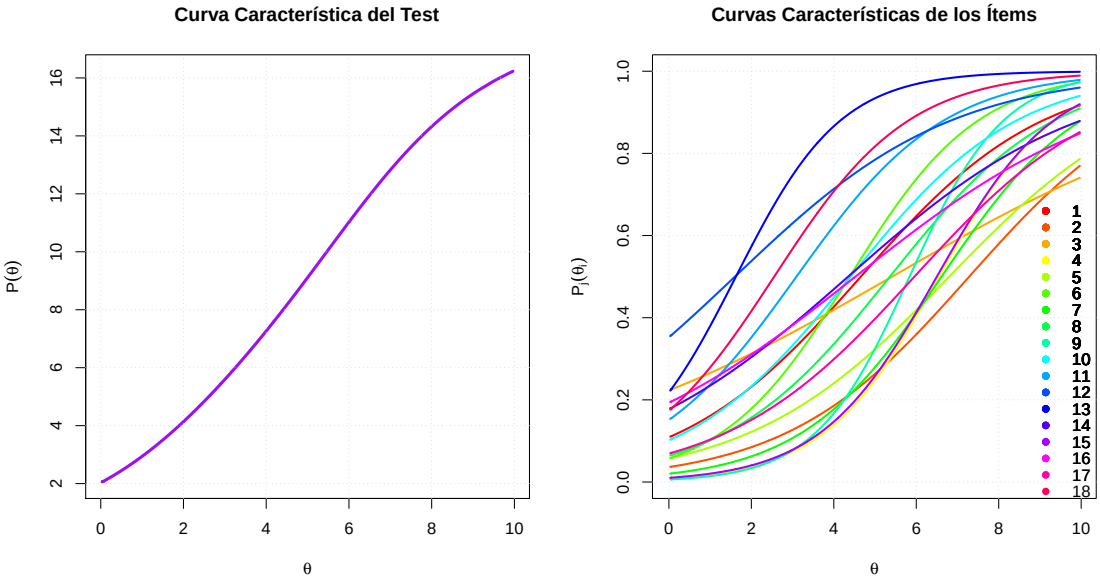


FIGURA C.4. Curvas características del test y de los ítems

Finalmente, se presentan las gráficas comparando las frecuencias esperadas y observadas de cada uno de los ítems.

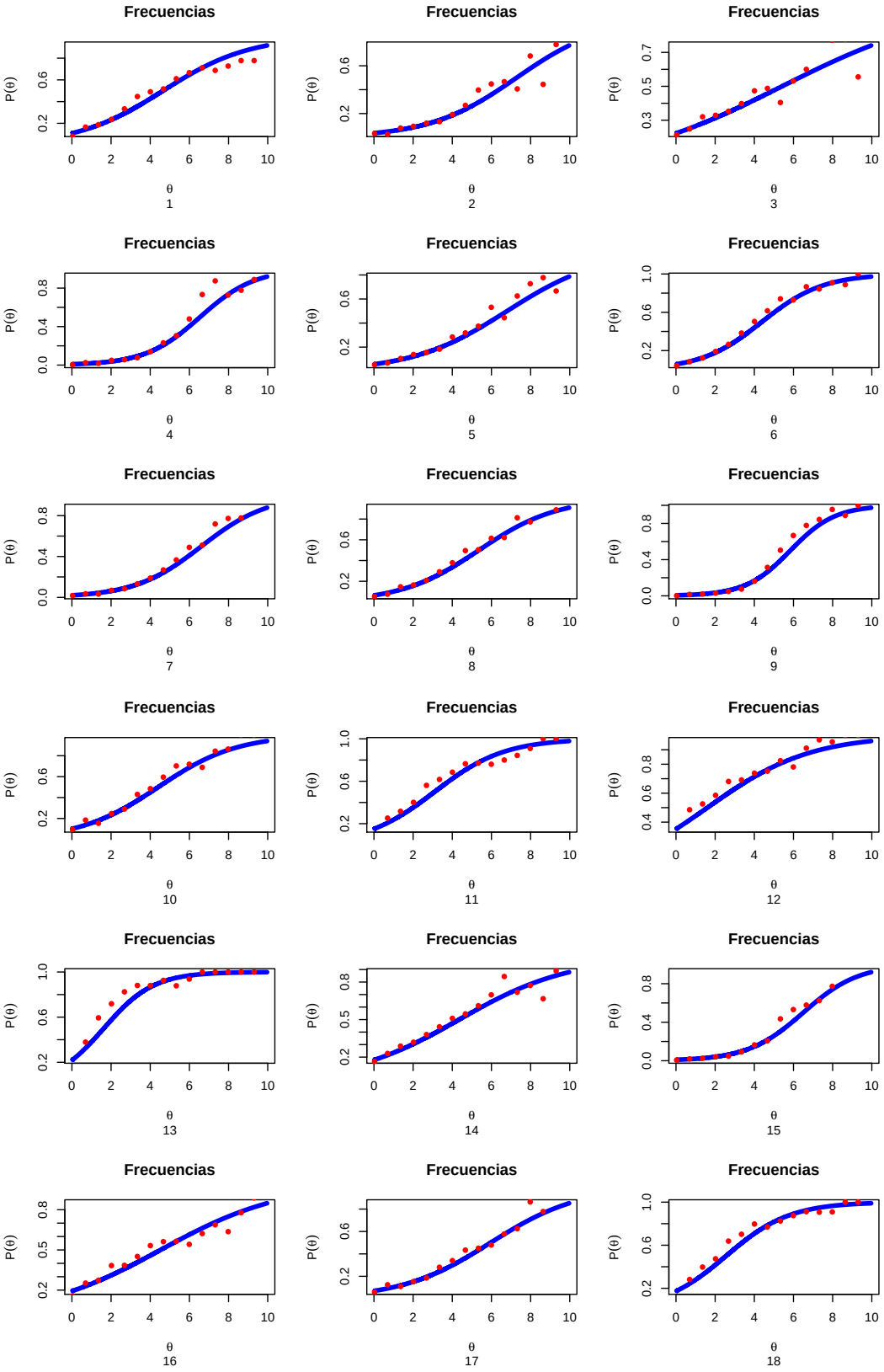


FIGURA C.5. Frecuencias Esperadas Vs. Observadas

Discusión y Conclusiones

En el contexto de la Teoría de Respuesta al Ítem, los trazos latentes (variables que no pueden medirse directamente), se han supuesto generalmente simétricos, de hecho, en la mayoría de la literatura relacionada, se asume que la distribución de los mismos es normal estándar. Sin embargo, algunos trazos presentan un comportamiento asimétrico, como aquellos relacionados con rasgos de personalidad y conducta, por lo que resulta indispensable estructurar metodologías de estimación que empleen diversas distribuciones.

El software estadístico desempeña un papel fundamental en los procesos de análisis y estimación de los modelos TRI, por lo que son cada vez más numerosos los aportes en este ámbito, lenguajes computacionales como LOGIST, Wingersky et al. (1982), BICAL, Wright & Mead (1976), BILOG, Mislevy & Bock (1982), MIRT, Chalmers (2012), ltm,Rizopoulos (2006), entre otros, permiten implementar distintos algoritmos de estimación pero siempre bajo el supuesto de normalidad del trazo latente.

A lo largo de este documento se propuso utilizar una distribución asimétrica para modelar el trazo latente en un modelo TRI logístico de dos parámetros, específicamente, la distribución Gamma y, bajo dicho supuesto, se implementó en `R` un procedimiento de estimación marginal de los parámetros de los ítems utilizando cuadratura para desarrollar el algoritmo EM, en tanto que, para la estimación del trazo se empleó máxima verosimilitud.

El procedimiento propuesto se aplicó en un test simulado donde el trazo y el parámetro de dificultad seguían una distribución Gamma, mientras que el parámetro de discriminación se modeló con una distribución lognormal. Los valores recuperados presentaron correlaciones altas con relación a los valores simulados: 0.969 para el parámetro de discriminación, 0.999 para el de dificultad y 0.829 para el trazo.

Posteriormente, se consideró un conjunto de datos real correspondiente al inventario PIHEMA R-2 que fue aplicado durante el proceso de admisión de 2006 en la Facultad de Humanidades y Educación de la Universidad Central de Venezuela.

Se emplearon modelos aditivos generalizados para localización, escala y forma (GAMLSS), para determinar la distribución del trazo en el test, encontrando que la distribución Gamma describía adecuadamente el comportamiento de la variable latente.

Una vez implementado el algoritmo de estimación propuesto utilizando `R`, se comparó el AIC con el obtenido mediante la implementación de la librería `ltm`, éste último resultó mayor, por lo que se concluye que el modelo logístico de dos parámetros con trazo latente distribuido Gamma es mejor que aquel que considera la variable latente con distribución normal.

Adicionalmente se compararon las gráficas de frecuencias esperadas vs. observadas con los resultados de las estimaciones obtenidos tanto con la metodología propuesta como con `ltm`, nuevamente, se observó un mejor ajuste con el primer método.

En las gráficas en que se comparan las frecuencias esperadas y observadas del test global resultó aún más evidente que la suposición de un trazo normal estándar conlleva a estimaciones inadecuadas, lo que puede derivar en interpretaciones equivocadas sobre el comportamiento de los individuos frente a los distintos ítems del test. Resulta por tanto indispensable considerar un mayor número de distribuciones para modelar el trazo latente e implementar metodologías de estimación en las que, características como la asimetría, sean tenidas en cuenta.

Trabajo Futuro

- Desarrollar la metodología propuesta para un mayor número de distribuciones.
- Implementar una metodología que permita seleccionar el mejor modelo posible para ajustar un test determinado.
- Proponer una metodología de Bondad de Ajuste que no se vea afectada por el tamaño del test o de la población.
- Generalizar el procedimiento a modelos en los que el trazo subyacente sea multidimensional.

Bibliografía

Andrade, D. F., Tavares, H. R. & Valle, R. C. (2000). *Teoria Da Resposta Ao Item: Conceitos E Aplicações*, São Paulo: Associação Brasileira De EstatíStica.

Attorresi, H., Lozzia, G., Abal, F., Galibert, M. & Aguerri, M. (2009). Teoría de respuesta al ítem. conceptos básicos y aplicaciones para la medición de constructos psicológicos, *Revista Argentina De Clínica Psicológica* **Xviii**: 179–188.

Baker, F. B. (1985). *The Basics Of Item Response Theory*, Heinemann.

Baker, F. B. (2001). *The Basics Of Item Response Theory*, 2nd ed. edn, Eric Clearinghouse On Assessment And Evaluation.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. in statistical theories of mental test scores., *Reading, Mass.: Addison-Wesley* pp. Chapters 17–20.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories, *Psychometrika* **37**(1): 29–51.

Bock, R. D. & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an em algorithm, *Psychometrika* **46**(4): 443–459.

Bock, R. & Lieberman, M. (1970). Fitting a response model forn dichotomously scored items, *Psychometrika* **35**(2): 179–197.

Boulet, J. R. (1996). The effect of nonnormal ability distributions on irt parameter estimation using full-information and limited-information methods, *University Of Ottawa, Canada* .

Brent, R. (1973). *Algorithms For Minimization Without Derivatives*, Prentice-Hall.

Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the R environment, *Journal Of Statistical Software* **48**(6): 1–29.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm, *Journal Of The Royal Statistical Society, Series B* **39**(1): 1–38.

Fernandes, A. D. & Atchley, W. R. (2006). Gaussian quadrature formulae for arbitrary positive measures, *Evolutionary Bioinformatics Online* **2**: 251–259.

Haberman, S. J. (1977). Maximum likelihood estimates in exponential response models, *Ann. Statist.* **5**(5): 815–841.

Hambleton, R., Swaminathan, H. & Rogers, H. (1991). *Fundamentals Of Item Response Theory*, Measurement Methods For The So, Sage Publications.

Harwell, M. R., Baker, F. B. & Zwarts, M. (1988). Item parameter estimation via marginal maximum likelihood and an em algorithm: A didactic, *Journal Of Educational Statistics* **13**: 243–271.

Levent, K., Tse-Chi, H. & Lifa, Y. (2001). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality, *Applied Psychological Measurement* **25**(2): 146–162.

Lord, F. (1952). *A Theory Of Test Scores*, number 7 in *A Theory Of Test Scores*, Psychometric Society.

Mislevy, R. J. (1986). Bayes modal estimation in item response models, *Psychometrika* **51**(2): 177–195.

Mislevy, R. J. & Bock, R. D. (1982). *Bilog: Maximum Likelihood Item Analysis And Test Scoring With Logistic Models For Binary Items*.

Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization, *Computer Journal* **7**: 308–313.

Neyman, J. & Scott, E. (1948). Consistent estimates based on partially consistent observations, *Econometrica* **16**: 1–32.

Orchard, T. & Woodbury, M. A. (1972). A Missing Information Principle: Theory And Applications, *University Of California Press* **1**: 697–715.

Partchev, I. (2009). *Irtoys: Simple Interface To The Estimation And Plotting Of Irt Models*. R Package Version 0.1.2.
URL: *Http://Cran.R-Project.Org/Package=Irtoys*

Páyer, M. (2013). Características de personalidad asociadas a estudiantes de permanencia prolongada, *Revista De Psicología* **29**(1).

Ponce, M. & Granell, P. (2006). Manual del inventario pihema r-2, *Universidad Central De Venezuela. Facultad De Humanidades Y Educación. Comisión Técnica De Admisión.* **Segunda Edición**.

Preston, K. S. J. & Reise, S. P. (2014). Estimating the nominal response model under nonnormal conditions, *Educational And Psychological Measurement* **74**(3): 377–399.

Rasch, G. (1960). *Probabilistic Models For Some Intelligence And Attainment Tests*, Studies In Mathematical Psychology, Danmarks Paedagogiske Institut.

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications, *Journal Of Educational Statistics* **4**(3): 207–230.

Rigby, R. A. & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape, *Applied Statistics* **54**: 507–554.

Rizopoulos, D. (2006). Ltm: An r package for latent variable modelling and item response theory analyses, *Journal Of Statistical Software* **17**(5): 1–25.

Smyth, G., Hu, Y., Dunn, P., Phipson, B. & Chen, Y. (2017). Statmod: Statistical modeling.

Stasinopoulos, D. M. & Rigby, R. A. (2007). Generalized additive models for location scale and shape (Gamlss) in r, *Journal Of Statistical Software* **23**(7).

Stroud, A. H. & Secrest, D. (1966). *Gaussian Quadrature Formulas*, Prentice-Hall Series In Automatic Computation, Prentice-Hall.

Swaminathan, H. & Gifford, J. A. (1986). Bayesian estimation in the three-parameter logistic model, *Psychometrika* **51**(4): 589–601.

Wingersky, M. S., Barton, M. A. & Lord, F. (1982). *Logist User's Manual*.

Woods, C. M. (2006b). Ramsay-curve item response theory (rc-irt) to detect and correct for nonnormal latent variables, *Psychometrika* **11**(3): 253–270.

Woods, C. M. (2008). Ramsay-curve item response theory for the three-parameter logistic item response model, *Applied Psychological Measurement* **32**(6): 447–465.

Woods, C. M. & Thissen, D. (2006a). Item response theory with estimation of the latent population distribution using spline-based densities, *Psychometrika* **71**(2): 281.

Wright, B. D. & Mead, R. J. (1976). *BICAL Calibrating Rating Scales With The Rasch Model*.

Yen, W. M. (1981). Using simulation results to choose a latent trait model, *Applied Psychological Measurement* **5**(2): 245–262.