



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

FACULTAD DE CIENCIAS MATEMÁTICAS

EAP. DE ESTADÍSTICA

**Modelo Rasch para estimar el nivel de las funciones
cognoscitivas en la escala "mini mental state
examination"**

TESINA

Para optar el Título de Licenciado en Estadística

AUTOR

César Augusto Acevedo Minaya

LIMA – PERÚ
2010

Modelo Rasch para estimar el nivel de las funciones cognitivas en la escala
“Mini Mental State Examination”

César Augusto Acevedo Minaya

Tesina presentada a consideración del cuerpo docente de la facultad de ciencias matemáticas de la Universidad Nacional Mayor de San Marcos, como parte de los requisitos para obtener el título profesional de licenciado en estadística.

Aprobada por:

Mg. Violeta Alicia Nolberto Sifuentes

Presidente

Mg. Emma Norma Cambillo Moyano

Miembro - Asesor

LIMA – PERÚ

ENERO - 2010

FICHA CATALOGRÁFICA

ACEVEDO MINAYA, CÉSAR AUGUSTO

Modelo Rasch para estimar el nivel de las funciones cognitivas en la escala “Mini Mental State Examination”, (Lima) 2010.

xii, 54 p., 29.7 cm. (UNMSM, Licenciado en Estadística, 2010).

Tesina, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas.

i. UNMSM / F. de C.M. ii. Título (serie)

A mis padres, hermano, amigos y profesores...

AGRADECIMIENTOS

Para la elaboración de esta tesina fue necesaria la ayuda de muchas personas e instituciones, sin las cuales no hubiera sido posible su realización.

A mi asesora por el tiempo dedicado a la lectura y corrección del presente trabajo, al Instituto Nacional de Salud Mental "HD- HN" por las facilidades brindadas para la realización de la tesina.

A todas aquellas personas que me ayudaron durante todo el tiempo de preparación de la tesina compartiendo conmigo con sus aportes y críticas.

MUCHAS GRACIAS

Resumen

Modelo Rasch para estimar el nivel de las funciones cognitivas en la escala
“Mini Mental State Examination”

César Augusto Acevedo Minaya

Enero 2010

Asesora: Mg. Emma Cambillo Moyano

Título obtenido: Licenciado en Estadística.

La evaluación de un instrumento psicométrico, bajo la teoría clásica de los test, brinda información sobre la confiabilidad y la validez del instrumento, también provee información de su capacidad predictiva si se compara con un criterio de referencia. El análisis aplicado a un instrumento bajo el modelo Rasch brinda información adicional como el nivel de dificultad del instrumento en función de la dificultad de los ítems que lo componen, además se obtienen estimaciones invariantes y comparables sobre el atributo evaluado por el instrumento para los individuos, a pesar que los sujetos provengan de diferentes grupos poblacionales.

Los resultados obtenidos al aplicar el modelo Rasch en el Mini Mental State Examination, determinaron que el instrumento posee una baja dificultad teniendo una gran precisión para el diagnóstico del nivel cognoscitivo considerado como deterioro severo, identificando como ítems de mayor dificultad aquellos que comprenden el cálculo, la evocación y la praxis, no ajustándose al modelo Rasch.

Palabras Clave: Modelo Rasch, Modelo de variable latente, Mini Mental State Examination, Teoría de respuesta a los ítems.

ABSTRACT

Rasch model to estimate the level of cognitive functions on the "Mini Mental State Examination"

Cesar Augusto Acevedo Minaya

January 2010

Adviser: Mg. Emma Cambillo Moyano

Degree: Bachelor of Statistics.

The evaluation of a psychometric instrument, under the classical theory of the test, provides information on the reliability and validity of the instrument, also provides information on its predictive power compared to a gold standard or benchmark. The analysis applied to an instrument under the Rasch model provides additional information such as the level of difficulty of the instrument in terms of the difficulty of the items that compose it also obtained comparable estimates invariants and the attribute assessed by the instrument for individuals, despite the subjects come from different populations groups.

The results obtained by applying the Rasch model in the Mini Mental State Examination determined that the instrument has a low difficulty is highly accurate for the diagnosis of cognitive level considered severe impairment, identified as those most difficult items involving computing, the evocation and practice, not to conform to the Rasch model.

Keywords: Rasch model, Latent variable model, MMSE, Theory of Item Response.

ÍNDICE

INDICE	iiiv
INTRODUCCIÓN	x
CAPÍTULO I. MODELOS PARA ANÁLISIS DE ÍTEMS	1
1.1 TEORÍA DE LOS TEST	2
1.1.1 TEORÍA CLÁSICA DE LOS TEST.	3
1.1.2 TEORÍA DE RESPUESTA A LOS ÍTEMS.	5
1.2 MODELOS DE VARIABLES LATENTES.	7
1.2.1 MODELO FACTORIAL.	7
1.2.2 MODELO FACTORIAL PARA DATOS BINARIOS.	9
1.3. MODELO RASCH	10
1.3.1 DEFINICIÓN.	10
1.3.2 ESTIMACIÓN DE PARÁMETROS.	11
1.3.3 AJUSTE DEL MODELO.	13
1.3.3.1 ESTADÍSTICO Q_2	13
1.3.3.2 BONDAD DE AJUSTE GLOBAL X^2	14
1.3.4 FUNCIÓN DE INFORMACIÓN.	15
1.3.5 CURVA CARACTERÍSTICA DEL TEST.	17
1.3.6 PUNTUACIÓN VERDADERA EN EL TEST.	17
CAPÍTULO II. APLICACIÓN DEL MODELO RASCH EN EL MINI MENTAL STATE EXAMINATION (MMSE) DE FOLSTEIN	18

DISCUSIÓN	29
CONCLUSIONES	30
BIBLIOGRAFÍA	32
APÉNDICE	34

INTRODUCCIÓN

El modelo Rasch es uno de los modelos más utilizados en la evaluación de instrumentos de medición en especial en el área de la psicometría. Este modelo se desarrolló basado en la teoría de respuesta a los ítems (TRI), cuyo soporte matemático está fundamentado en los modelos de variables latentes.

Los resultados que provee el modelo Rasch brindan información a nivel de sujeto y a nivel de ítems, siendo esto último de importancia para la evaluación del instrumento ya que permite identificar el nivel de dificultad o exigencia del instrumento además de obtener una estimación sobre la habilidad o característica latente del individuo bajo estudio. Dentro de la literatura sobre el tema se tiene algunos trabajos donde se aplicó el modelo Rasch cuyos resultados comentamos a continuación.

El objetivo del estudio “Propriedades psicométricas da versão brasileira da escala de qualidade de vida específica para acidente vascular encefálico: aplicação do modelo Rasch. Lima RCM y col. 2007” fue adaptar transculturalmente el instrumento Stroke Specific Quality of Live Scale y evaluar sus propiedades psicométricas. De los 49 ítems que conformó el instrumento 4 fueron retirados del estudio ya que comprometían la validez de constructo, los investigadores concluyeron que el instrumento adaptado es clínicamente útil para la población bajo estudio.

El estudio “Adaptive cognitive testing in dementia, por Hans Wouters y col., 2008” tuvo como objetivo investigar si el Examen Cognoscitivo de Cambridge (CAMCOG), una prueba de detección utilizada para la demencia, puede ser adaptada a cada paciente en una versión digital adaptada (CAT). Es un método de dos pasos con los datos recogidos previamente CAMCOG estimó las capacidades del paciente y la dificultad de los ítems bajo un modelo Rasch. Se concluye que la prueba adaptada CAT permite mayor eficiencia para las proyecciones de la demencia que las pruebas con un instrumento amplio. Actualmente la aplicación de los modelos Rasch se va extendiendo a nuevas áreas de las ciencias como la Neuropsicología, disciplina que estudia las enfermedades mentales bajo un enfoque psíquico y neurológico, como es el caso del estudio del deterioro cognitivo, que es abordado bajo un punto de vista multidisciplinario. El análisis de las propiedades psicométricas en pruebas de evaluación de funciones cognitivas como en el caso de *Mini Mental State Examination* propuesto por Folstein bajo el enfoque de la Teoría de Respuestas a los Ítems permitirá determinar de mejor manera las características de la prueba a través de la unidad elemental de ésta, los ítems; además de obtener una estimación del nivel de funcionamiento cognoscitivo que puede ser estándares y comparables, es decir que las estimaciones sean invariantes a la versión del test aplicado.

A pesar de que el modelo Rasch es una herramienta de gran utilidad para la evaluación de instrumentos, dicho modelo no es muy estudiado y aplicado dentro del ambiente académico. El objetivo del presente trabajo es estimar el nivel de funcionamiento cognoscitivo a través de la prueba Mini Mental State Examination (MMSE) bajo el modelo Rasch.

En este apartado hacemos una revisión de los antecedentes sobre el modelo bajo estudio, se formula el planteamiento del problema, el objetivo de la investigación y la finalidad del trabajo. En Capítulo I se hace una introducción a la teoría de los test comprendiendo la teoría clásica de los test y la teoría de respuesta a los ítems, además se presenta la teoría de los modelos de variables latentes en general y para datos binarios, posteriormente se explica la teoría sobre el modelo Rasch desde la definición, estimación, pruebas de bondad de ajuste. Los resultados del análisis sobre la estimación del modelo, la

dificultad de cada ítem, así como la clasificación de los ítems en niveles de deterioro cognoscitivo son presentados en Capítulo II, finalmente la discusión y conclusiones son presentadas. En los anexos se presentan información complementaria sobre el proceso de envejecimiento y la repercusión en las funciones cognoscitivas, modelos logísticos incorporados en los modelos de variables latentes, el algoritmo de programación de los análisis, y por último la bibliografía sobre el tema desarrollado en el presente trabajo.

CAPITULO I

MODELOS PARA EL ANÁLISIS DE ÍTEMS.

El modelo Rasch¹, forma parte de los modelos de variables latentes adoptado por la teoría de respuesta a los ítems para estudiar aquellos fenómenos que no se pueden medir de forma directa, sino a través de herramientas diseñadas particularmente para cuestionarios que permiten recoger información para tener una aproximación del fenómeno bajo estudio.

Los modelos Rasch y los demás modelos de variables latentes aparecieron a partir de la segunda mitad del siglo XX, sin embargo, su fundamento teórico empezó a surgir a inicios del siglo XX, cuando Thurstone en 1925 encontró la relación entre la edad de entrevistados y la proporción de aciertos de cada ítem de una prueba, cuyos datos fueron tomados del test de inteligencia de Binet. Esta curva descrita por Thurstone es la primera formulación de la curva Característica de los Ítems.

En 1936 Richarson ajustó por primera vez la ojiva normal a las respuestas de los ítems, debido a la necesidad de controlar la dificultad de los ítems según los

¹ Rasch George (1960), Probabilistic models for some intelligence and attainment ítems.

objetivos del test. Posteriormente Tucker y Lazardfield establecen los términos de Curva Característica del ítem y rasgo latente.

En 1954, Lord publicó su tesis doctoral bajo el seguimiento y asesoría de Gulliksen y Tucker. En dicho trabajo Lord presenta los principios básicos de la Teoría de Respuesta a los Ítems (TRI). Muñiz (1997) destaca la importancia de esta publicación y otras posteriores como "... la formulación más sistemática de los principales conceptos de la teoría de respuesta a los ítems, a partir de los cuales surgirán desarrollos posteriores". A partir de los trabajos de Lord, se empiezan a desarrollar modelos teóricos matemáticos basados en los supuestos de la TRI. Muñiz refiere que en 1957 Birnbaum utiliza por primera vez la función logística en vez de la ojiva normal, luego en 1960 G. Rasch publicó un libro en el que expone con detalle el modelo logístico de un parámetro, dándole el nombre de este al primer modelo formulado con detalle en su obra. Posteriormente, Lord y Novick con el aporte de Birnbaum publicaron en 1968 un libro que supone un fuerte impulso sobre la TRI. En 1969 Wrigth y Panchapakesan publican su trabajo sobre la estimación de los parámetros del modelo Rasch, presentando simultáneamente el primer programa para la estimación *BICAL*. Ese mismo año, Samejima propone modelos para ítems de respuesta politémicas y de respuesta ordinal.

El desarrollo de la teoría de respuesta a los ítems continua por varios años, a partir de la década de los 80, el avance de la teoría de los ítems se orientó mas por el desarrollo de software, definir las aplicaciones de la TRI, entre otros. Muñiz (1997) presenta en su publicación sobre TRI una breve pero concisa explicación sobre el avance de la Teoría de la TRI a través del tiempo, donde detalla los eventos mencionados y otros sucesos adicionales.

Una breve revisión sobre las teorías de los test es descrita a continuación.

1.1 TEORÍA DE LOS TEST

En algunas ciencias como la psicología, educación y ciencias del comportamiento, es de interés el estudio de fenómenos que en suma

son abstracciones de la realidad, como son los procesos conductuales, cognitivos, aptitudes o habilidades entre otras. Este tipo de fenómenos no son susceptibles a la medición de manera directa, sino que se requiere elaborar instrumentos que sean capaces de observar tales fenómenos indirectamente, es decir, que sean capaces de observar el fenómeno a través de la observación de otras variables que se asume están relacionadas con el fenómeno de interés.

Los instrumentos utilizados son cuestionarios o test, en donde las preguntas o ítems recogen información parcial sobre el rasgo o atributo de interés, en base a la teoría que sustenta su conceptualización, dichos instrumentos son operacionalizados para elaborar una escala que represente la magnitud del rasgo o atributo estudiado. Por tanto, los test han de generar puntuaciones que son una aproximación del rasgo o atributo estudiado.

Esta situación originó el desarrollo de procedimientos y técnicas para la elaboración de instrumentos que sean capaces de medir este tipo de fenómenos de manera precisa y fiable, además de tener la certeza de que el instrumento es válido para el estudio del fenómeno. Según menciona Abad y col. (2006), la Teoría de los Test comprende todos estos procedimientos, pues su objetivo es elaborar instrumentos (test) que sean lo más apropiado para describir el fenómeno.

La Teoría de los Test comprende dos enfoques que son la Teoría Clásica de los Test (TCT) y la Teoría de Respuestas a los Ítems (TRI).

1.1.1. TEORÍA CLÁSICA DE LOS TEST. (TCT)

Según Abad, esta teoría surge bajo la propuesta de Spearman en 1907, asume que existe una relación directa entre la puntuación obtenida por el instrumento y la magnitud del rasgo o atributo, esta relación se puede expresar bajo un modelo lineal simple

$$Y=V+E$$

donde

Y: Puntuación empírica.

V: Puntuación verdadera.

E: Error de medida.

El modelo clásico se centra en la puntuación total del test, ya que asume que la puntuación empírica es una aproximación de la puntuación verdadera y que la diferencia entre ambas es mínima,

Debido a que los test son una construcción arbitraria por parte del investigador basados en definiciones teóricas sobre el atributo, se debe comprobar que este instrumento está recogiendo en realidad la información relacionada sobre el atributo, ya que como se mencionó anteriormente la medición del atributo no se hace de forma directa sino sobre indicadores relacionados a este, propiedad conocida como validez del instrumento, y que comprende la validez: de contenido, de constructo y de criterio.

Otra característica fundamental que deben cumplir los test, es la fiabilidad, es decir que los test realicen mediciones estables, precisas o consistentes sobre un rasgo determinado. Observemos que un test puede ser susceptible a variación al momento de la medición o que el atributo en un individuo no permanezca constante. Podemos mencionar tres formas distintas de fiabilidad: fiabilidad como estabilidad temporal, como consistencia interna y fiabilidad como correlación entre formas paralelas.

Este enfoque presenta algunas dificultades, debido a que el número de ítems afectan la estimación del atributo latente, además existe el problema de comparar dos test distintos que miden el mismo atributo. Los test tienen diferente desempeño dependiendo a los sujetos a los que les aplica la prueba.

Según Muñiz (1997) la medición depende del instrumento utilizado y las propiedades de los instrumentos están en función de los objetos medidos, los sujetos.

Estos inconvenientes fueron tratados por varios investigadores hasta que finales de la década de los 60 del milenio pasado aparece un nuevo enfoque, la teoría de respuesta a los ítems.

1.1.2 LA TEORÍA DE RESPUESTA A LOS ÍTEMS (TRI)

Para Muñiz, la teoría de respuesta a los ítems es un nuevo enfoque de evaluación y análisis de la construcción de cuestionarios que busca superar algunas dificultades que se presenta en la teoría clásica de los test.

Su objetivo es encontrar estimaciones de la variable latente para cada sujeto, que sean invariante a los test aplicados. También se puede obtener estimaciones sobre las características de los test y de sus ítems independientemente de la población a la que se aplique el instrumento.

La teoría de respuesta a los ítems (TRI) y la teoría clásica de los test comparten el mismo principio, es decir, la variable latente es obtenida a partir de la medición de las variables no observadas, bajo un modelo establecido. Pero existen varias diferencias entre estos dos enfoques, que va desde la unidad elemental del análisis, hasta la teoría estadística en el cual se sustentan los modelos.

Para la TRI la unidad elemental de análisis es el ítem y no el test. Su análisis no se basa en las puntuaciones globales, sino que busca un modelo que represente la relación entre la probabilidad de escoger una alternativa de respuesta al ítem y el rasgo latente que es estudiado.

A consecuencia de lo antes mencionado, el modelo formulado en la TRI no es el mismo que en el modelo clásico, ya que cada ítem está representado bajo un modelo particular y debido a que los ítems no están en una escala métrica, el modelo lineal no es el

más adecuado; por lo tanto se recurre al análisis factorial apoyado por algunas nociones básicas de los modelos lineales generalizados, Bartholomew (2002) hace una revisión sobre este tema.

Por último, la TRI es un enfoque que va siendo aplicado en pruebas donde se miden las capacidades, destrezas o aptitudes, llamados test de rendimiento óptimo. Por lo tanto, las alternativas de los ítems son clasificados en dicotómicas (acierto / error), politómicas (muchas alternativa pero solo una es correcta) y ordinales (se mide un grado de acierto o intensidad del atributo).

Los modelos en la TRI surgieron bajo supuestos muy restrictivos, los supuestos de la TRI son:

1. La relación no lineal entre probabilidad de acierto de un ítem y la variable latente puede ser expresado a través de una función curvilínea, llamada Curva característica del ítem.
2. Los modelos de TRI asume que todos los ítems que componen al test están relacionados a una sola variable, es decir un solo rasgo es observado en el instrumento, a esta condición se le denomina Unidimensionalidad. Aunque más adelante se mencione los avances para el caso de TRI de rasgos múltiples.
3. La independencia local consiste en que para un determinado valor de la variable latente, la respuesta de un ítem no está sujeta a las repuestas de los otros ítems. Esto se puede expresar, bajo el concepto de probabilidad de la siguiente forma, la probabilidad de acertar un número n de ítems es igual al producto de las probabilidades de acertar n ítems.

Los modelos de variables latentes aplicados en la teoría de respuesta a los ítems, surgen con Rasch (1960) el cual postula un modelo simple donde asumen que las variables observadas de

escala nominal (dicotómicas) son manifestaciones de una sola variable latente en común “ θ ” en cuyo caso pertenece a una escala continua. En este modelo se asume que cada ítem tiene un parámetro que lo caracteriza, es el parámetro de dificultad “ b ”.

Posteriormente Birnbaum propone una generalización del modelo Rasch, considerando en el modelo dos parámetros adicionales; el índice de discriminación (a) y el acierto al azar (c).

También se extendieron estos modelos para el caso que los ítems sean politómicos o de respuesta ordinal (Samejima), también cuando se cuente con más de una atributo o variable latente en el modelo.

1.2 MODELOS DE VARIABLES LATENTES

Los modelos de variables latentes son un conjunto de métodos estadísticos que permiten encontrar variables no observadas de manera directa a partir de un conjunto de variables observadas que están relacionadas con las variables de interés.

Los modelos de variables latentes establecen que las variables observadas son explicadas por un conjunto de variables no observadas (latentes) a través de un modelo.

1.2.1 MODELO FACTORIAL

Definición 1. Sea (Y_1, Y_2, \dots, Y_q) un conjunto de variables aleatorias continuas observadas; y sea (X_1, X_2, \dots, X_p) un conjunto de variables no observadas (latentes). Entonces:

$$Y_j = \lambda_{j0} + \lambda_{j1}X_1 + \lambda_{j2}X_2 + \dots + \lambda_{jp}X_p + e_j \quad (j = 1, \dots, q) \quad (1)$$

Otra forma de expresar este modelo es bajo notación matricial propuesto por Peña:

$$Y = \mu + \Lambda X + u$$

donde:

X : es el vector de variables latentes o factores no observados que sigue una distribución $N_p(0, I)$.

Λ : Es una matriz $q \times p$ de constantes desconocidas, donde $q > p$.

u : Es el vector de errores no observados que tiene una distribución $N_q(0, \Psi)$, donde Ψ es diagonal, y los errores están incorrelacionados con los factores.

μ : Es el vector de medias de la matriz Y .

Y : Vector de variable con distribución normal multivariada $N_q(\mu, V)$.

Este es el modelo clásico de variables latentes, conocido como modelo lineal factorial, donde las variables observadas y las variables latentes son continuas, con distribución normal conjunta. Sin embargo, en la realidad se presentan otras situaciones donde las variables observadas son categóricas y en particular de respuestas binarias, siendo no viable aplicar el modelo factorial, debido a que el conjunto de datos no posee distribución normal conjunta. Ante esta situación se requiere de un modelo que pueda relacionar las variables latentes y las variables observadas cuando son categóricas binarias, para tal efecto se dispone del modelo factorial para datos binarios.

El modelo factorial para datos binarios incorpora dentro del modelo factorial lineal el esquema de los modelos lineales generalizados. Los modelos lineales generalizados que son una extensión de la familia de los modelos lineales generales, permite superar los problemas que surgen por la violación de algunos supuestos: normalidad de la variable respuesta,

homocedasticidad de los errores, pero además permite encontrar un enlace entre variables con distribución de probabilidad conocida y un conjunto de variables latentes a través de una función denominada función de enlace. Dentro de esta familia de modelos lineales se encuentran los modelos lineales para datos categóricos binarios conocidos como modelos de regresión logística que se verá en el apéndice. Bartolomew (2002) hace una presentación sistemática sobre los modelos de variable latente para datos binarios y que a continuación se presentan.

1.2.2 MODELO FACTORIAL PARA DATOS BINARIOS

Definición 2. Sea (Y_1, Y_2, \dots, Y_q) un conjunto de variables aleatorias discretas observadas, donde $Y_i \sim B(1, \pi_i)$; y sea (X_1, X_2, \dots, X_p) un conjunto de variables no observadas (latentes). Entonces:

$$g(E(Y_i)) = \lambda_{i0} + \lambda_{i1}X_1 + \lambda_{i2}X_2 + \dots + \lambda_{ip}X_p + e_i \quad (i = 1, \dots, q) \quad (2)$$

Bajo la notación matricial se tiene:

$$GY = \mu + \Lambda X + u$$

Donde G es una matriz de transformaciones que permiten enlazar la variable dicotómica con el predictor lineal del modelo factorial, es decir es una matriz de funciones de enlaces para cada variable Y, en este caso es una matriz de transformaciones logit. Las demás componentes del modelo son explicadas en (1).

Un caso particular de estos modelos es cuando nos encontramos con una sola variable latente que es capaz de resumir al conjunto de variables observadas.

Definición 3. Sea (Y_1, Y_2, \dots, Y_q) un conjunto de variables aleatorias discretas independientes e idénticamente distribuidas, donde $Y_i \sim B(\pi_i, 1)$; y sea X una variable no observadas (latente). Entonces:

$$g(E(Y_i)) = \lambda_{i0} + \lambda_{i1}X_1 + e_i \quad (i = 1, \dots, q) \quad (3)$$

Este modelo tiene bastante aplicación dentro de la teoría de respuesta a los ítems, tomando una notación particular en cada uno de los modelos de la TRI. Para el modelo Rasch la notación que tiene es:

$$E(Y_i) = \frac{e^{\lambda_{i0} + \lambda_{i1}X_1}}{1 + e^{\lambda_{i0} + \lambda_{i1}X_1}} = \frac{e^{b_i + \theta}}{1 + e^{b_i + \theta}}$$

donde:

λ_{i0} : Es el intercepto, también conocido como el índice de dificultad del ítem y su notación es $\lambda_{i0} = b_i$.

λ_{i1} : Es conocido como índice de discriminación, un caso particular es el modelo Rasch, donde $\lambda_{i1} = 1$ ó $\lambda_{i1} = 1.7$ como constante para todos los ítems, cuando λ_{i1} varía para cada ítem, el modelo es conocido como modelo de dos parámetros.

1.3 MODELOS RASH

1.3.1 DEFINICIÓN

Los modelos Rasch forman parte de una familia de modelos dentro de la TRI cuya finalidad es estimar un rasgo latente del cual no se puede medir de manera directa pero puede ser estimado en base a un instrumento elaborado. Este tipo modelo fue formulado por Rasch (1960), dicho modelo latente está compuesto por la variable de interés “ θ ” y adicionalmente un

parámetro que caracteriza al ítem “b” que representa la dificultad de acertar el ítem.

$$\pi_{(x_{ij}=1|\theta)} = \frac{e^{D(\theta_j - b_i)}}{1 + e^{D(\theta_j - b_i)}} \quad (4)$$

donde:

$\pi(\theta)$: Probabilidad de acertar el ítem i a un determinado nivel θ .

θ : Valores de la variable medida.

b_i : índice de dificultad (parámetro de posición).

e : Base de los logaritmos neperianos.

La constante D forma parte dentro del modelo pero no como parámetro, permite acondicionar el modelo logístico de tal forma que se asemeje a un modelo de ojiva normal, usualmente es 1 lo que indica que es un modelo logístico ó 1.7 que hace que la curva de la función logística se asemeje a la curva normal.

Los supuestos de la TRI son aplicados a los modelos Rasch:

1. La probabilidad de responder correctamente a un ítem está relacionada con la variable latente a través de la Curva Característica del ítem.
2. La unidimensionalidad atribuye que la respuesta a los ítems es generado con un solo atributo.
3. La independencia local determina que la respuesta a un ítem no está condicionada a la respuesta de los demás ítems, de forma análoga la probabilidad de acertar un ítem es independiente a acertar los demás ítems.

1.3.2 ESTIMACIÓN DE PARÁMETROS

Las estimaciones de los parámetros de dificultad “b”, así como la estimación del valor de la variable latente “ θ ”, para cada persona

son realizadas según el enfoque del análisis factorial para datos binarios a través del método de máxima verosimilitud, cuyos valores estimados serán aquellos que maximicen la probabilidad de ocurrencia de los datos obtenidos al aplicar los ítems.

La probabilidad de acertar un ítem “Y_i” para un determinado valor de θ está dado por

$$P(Y_i | \theta) = [\pi(\theta)]^{Y_i} [1 - \pi(\theta)]^{(1-Y_i)} \quad (5)$$

La probabilidad de que se produzca un determinado patrón de respuesta vendrá dado por el producto de probabilidades, basado en el supuesto de independencia local. Entonces la función de verosimilitud para un conjunto de n ítems está dado por:

$$L(y_1, y_2, y_3, \dots, y_n | \theta) = \prod_{i=1}^n [\pi_i(\theta)]^{y_i} [1 - \pi_i(\theta)]^{(1-y_i)} \quad (6)$$

Luego, se aplica el logaritmo neperiano;

$$\ln[L(y_1, y_2, y_3, \dots, y_n | \theta)] = \sum_{i=1}^n [y_i \ln \pi_i(\theta) + (1 - y_i) \ln(1 - \pi_i(\theta))] \quad (7)$$

Sea una muestra aleatoria de N sujetos, los cuales son sometidos a responder n ítems; la función de máxima verosimilitud es:

$$L(y|\theta) = \prod_{j=1}^N \prod_{i=1}^n [\pi_{ij}(\theta)]^{y_{ij}} [1 - \pi_{ij}(\theta)]^{(1-y_{ij})} \quad (8)$$

La función máxima verosímil $L(y|\theta)$ pasa por una transformación logaritmo neperiano².

$$\ln(L(y|\theta)) = \sum_{j=1}^N \sum_{i=1}^n [y_{ij} \ln(\pi_{ij}(\theta)) + (1 - y_{ij}) \ln(1 - \pi_{ij}(\theta))] \quad (9)$$

Para estimar los valores de θ se debe considerar si los parámetros de los ítems son conocidos o hace falta estimarlos, por tanto existen dos procedimientos que permitirán la estimación de los parámetros.

Si se conoce previamente la dificultad de los ítems, se procede a una estimación condicional y solo se estima θ :

$$\partial \ln L / \partial \theta_j = 0 \quad (10)$$

En cambio, si desconocemos la dificultad de los ítems se procede a una estimación conjunta para encontrar los estimadores de θ y b :

$$\partial \ln L / \partial (\theta_j, b) = 0 \quad (11)$$

Para encontrar solución a las ecuaciones 10 y 12, se requiere de métodos matemático. Bolfarine (2003) y Muñiz (1997) sugieren utilizar el método de Newton Raphson.

1.3.3 AJUSTE DEL MODELO.

Existen procedimientos estadísticos para evaluar el ajuste del modelo a través de la comparación entre los valores estimados y los observados conocidos también como la bondad de ajuste del modelo, análisis de los residuos y la comparación de las distribuciones de las puntuaciones.

1.3.3.1 Estadístico Q_2

Muñiz refiere que Wright y Panchapakesan propusieron en 1969 una medida de ajuste denominado Q_2 , que se distribuye a una X^2 con $k-1$ grados de libertad⁽⁶⁾,

$$Q_2 = \sum_{k=1}^K \frac{n_k [\pi(\theta_k) - \pi_e(\theta_k)]^2}{[\pi(\theta_k)][1 - \pi(\theta_k)]} \quad (12)$$

Donde

k : numero de categorías en las que se divide θ

n_k : numero de sujetos dentro de k -ésima categoría.

$\pi(\theta_k)$: Valor de la CCI dado por la fórmula del modelo con los parámetros estimados, para la k-ésima categoría.

$\pi_e(\theta_k)$: Proporción de sujetos que superan (empíricamente) el ítem para una categoría determinada.

El estadístico Q_2 se calcula agrupando por intervalos los valores θ , luego compara las probabilidades estimadas por el modelo y las probabilidades empíricas, obteniendo así el estadístico Q_2 . Esta prueba evalúa la bondad de ajuste del modelo Rasch para un ítem.

1.3.3.2 Bondad de ajuste global X^2

Una prueba para evaluar la bondad de ajuste global, basado en la X^2 con $(2^p - p(q+1) - 1)$ grados de libertad, siendo q el número de parámetros que se estima para los ítems que para los modelos Rasch es igual a 1.

$$X^2 = \sum_{r=1}^{2^p} \frac{(O_{(r)} - E_{(r)})^2}{E_{(r)}} \quad (13)$$

donde:

r : Es el número de posibles soluciones o respuestas al test, en máximo número de soluciones es 2^p .

$O_{(r)}$: El número de respuestas observadas para la r ésima combinación de respuesta.

$E_{(r)}$: El número de respuestas esperadas para la r ésima combinación de respuesta.

Rizopoulos (2006) y Bartolomew (2002) exponen con más detalle su desarrollo. Si la extensión del test (número de ítems) es largo, existe la posibilidad de encontrar muchas categorías (patrones de respuesta) con frecuencias

esperadas por debajo de 5; Rizopoulos sugiere aplicar pruebas de hipótesis Bootstrap, aquí la significancia estadística (p-valor) está dada por:

$$p^* = \frac{\left[1 + \sum_{i=1}^B I(T_i > T_{obs}) \right]}{B + 1}.$$

donde

I : Función indicadora que permite calcular la frecuencia de casos donde se cumple la condición establecida.

T_{obs} : El valor del test X^2 en muestra principal.

T : El valor del test X^2 en las muestras replicadas.

B : El número de réplicas.

1.3.4 FUNCIÓN DE INFORMACIÓN

El estimador $\hat{\theta}$ es un estimador máximo verosímil de θ , por propiedad de los estimadores de máxima verosimilitud estos estimadores se distribuyen asintóticamente normales con media θ y varianza dada por²

$$\text{var}(\hat{\theta}) = \sum_{i=1}^n \left\{ \frac{[\pi'_i(\theta)]^2}{\pi_i(\theta)(1-\pi_i(\theta))} \right\}^{-1} \quad (14)$$

donde

π_i : Probabilidad de acertar el ítem

π'_i : la derivada de la probabilidad de acertar el ítem

En tanto el error típico de medida o desviación típica está dado por la raíz cuadrada de la expresión (14), que permitirá establecer

intervalos de confianza para θ , Muñiz refiere que esta definición la propuso Birnbaum en 1968, basado en los aportes de Fisher.

Otra expresión del error típico está basada en la función de información.

$$I(\theta) = \sum_{i=1}^n \frac{[\pi'_i(\theta)]^2}{\pi_i(\theta)(1-\pi_i(\theta))} \quad (15)$$

Por tanto podemos ver la relación entre error típico de medida y la función de información en la siguiente expresión

$$Se = \frac{1}{\sqrt{I(\theta)}} \quad (16)$$

La precisión de la medición de los test, no es uniforme sino que ésta varía en función del nivel de los sujetos en la variable medida, por tanto para cada nivel (θ_j) se puede obtener un error típico de medida.

La función de información del ítem esta dado por:

$$I(\theta) = \frac{[\pi'_i(\theta)]^2}{\pi_i(\theta)(1-\pi_i(\theta))} \quad (17)$$

Su utilidad radica en que se puede identificar la cantidad de información que el ítem aporta al test, y además se puede identificar a qué nivel de θ aporta dicha información.

La eficiencia relativa (ER) de dos test para un determinado valor de θ es definida como la razón entre las funciones de información de cada test para dicho valor θ .

$$ER = \frac{I(\theta_x)}{I(\theta_y)} \quad (18)$$

1.3.5 CURVA CARACTERÍSTICA DEL TEST

La curva característica del test es la suma de las curvas características de los ítems que componen el test; es decir si a cada nivel de θ se suman los valores de $\pi_i(\theta)$ de cada ítem para ese nivel se obtiene la curva característica del Test.

$$\text{CCT} = \sum_{i=1}^n \pi_i(\theta) \quad (19)$$

1.3.6 LA PUNTUACIÓN VERDADERA EN EL TEST

La puntuación verdadera de cada individuo en el test viene estimada por la suma de probabilidades $\pi_i(\theta_j)$ dadas por las curvas características de los ítems que componen el test, para el valor θ_j :

$$V_j = \sum_{i=1}^n \pi_i(\theta_j) \quad (20)$$

Este es el valor calculado por la curva característica del test para θ_j .

CAPÍTULO II.

APLICACIÓN DEL MODELO RASCH EN EL MINI MENTAL STATE EXAMINATION (MMSE) DE FOLSTEIN.

Este instrumento fue creado por Folstein y McHugh (Folstein 1975), con la intención de examinar el estado mental de forma rápida y cuantificable para la evaluación de pacientes psiquiátricos de manera que permita separar pacientes con disturbio cognoscitivo de aquellos que no lo tienen.

El Minimal State Examination denominado también MMSE o Mini se ha convertido en un instrumento breve de exploración sobre trastorno de deterioro cognoscitivo agudo (demencia) más usado. Está incluido en el Diagnostic Interview Schedule (DIS) desde 1979, para estudios epidemiológicos. Ha sido recomendado por el National Institute of Neurological and Communicative Disorders Association (NINCDS-ADRDA) en 1984 para documentar el diagnóstico clínico de probable enfermedad de Alzheimer's.

El MMSE se compone de 11 temas o ejes que son calificados y dan un puntaje total de 30, según sus autores tiene dos secciones, la primera sección solo requiere respuestas orales (orientación en el tiempo y lugar, registro y evocación, atención) que dan un puntaje total para esta sección de 21; la segunda sección demanda habilidades para nombrar, seguir órdenes orales y escritas, escribir una frase y copiar un polígono compuesto, que reúne 9 puntos (Folstein y col., 1975)

Originalmente los ítems fueron agrupados en cinco bloques:

- Orientación en el tiempo y espacio.
- Registro

- Atención y cálculo
- Evocación
- Lenguaje

Otros investigadores consideran que los ítems se pueden clasificar en 7 bloques en vez de 5; donde separan la orientación del tiempo de la orientación del espacio, y en el lenguaje lo separan entre lenguaje y construcción civil.

El instrumento ha pasado por varios procesos de adaptación, Robles (2003) realizó una adaptación del MMSE aplicado a 368 sujetos mayores de 55 años, 45 casos con diagnóstico de depresión y 55 casos con diagnóstico DTA en la ciudad de Lima. Los resultados reportados en el trabajo determinaron que pacientes con mayor grado de instrucción tendrían mayor puntaje, pero el rendimiento disminuye conforme aumenta la edad. Además se encontró que un puntaje de corte igual a 23 permite tener una sensibilidad de 86% y una especificidad de 94% para el diagnóstico de deterioro cognoscitivo comparando con la evaluación clínica de un solo criterio. Se obtuvo una confiabilidad de 0.62 y la validez de constructo determinó 5 factores.

El Mini Mental State Examination aplicado en el estudio epidemiológico fue adaptado del original y ampliado con ítems alternativo (*ver anexo*). Para fines del presente estudio solo se consideró los ítems originales traducidos al español.

El Instituto Nacional de Salud Mental “Honorio Delgado – Hideyo Noguchi” ejecutó una investigación epidemiológica en Lima metropolitana y Callao en el 2002, que permite estudiar las características de salud mental de la población, así como la prevalencia de discapacidad psiquiátrica, acceso a servicios de salud y factores asociados.

El presente estudio investigó en Lima Metropolitana y el Callao los principales aspectos de la salud mental, como son los trastornos psiquiátricos más prevalentes, la violencia familiar y su vínculo con la salud mental, la salud mental del adulto mayor, así como problemas del adolescente evidenciados por el hallazgo de indicadores suicidas en investigaciones pasadas. Además de los temas propuestos, se ha considerado estudiar otros aspectos asociados a la salud mental, como: la calidad de vida, funcionamiento y discapacidad, estilos

parentales y su relación con los problemas psiquiátricos, aspectos económicos, acceso a servicios y salud mental en general, entre otros. Una particularidad de este estudio es que todos estos aspectos serán estudiados en un mismo hogar, por lo tanto, la potencialidad de la obtención de datos importantes y estudios a profundidad a través del cruce de información entre las unidades de análisis será de mucho valor para aspectos clínicos, epidemiológicos y de intervención.

El estudio epidemiológico fue de tipo descriptivo y de corte transversal. La población fueron los ocupantes que residieron permanente o habitualmente en la ciudad de Lima Metropolitana y Callao, con edad de 12 años a más en el año 2002. La muestra fue probabilística y bietápica. Las unidades de análisis fueron: adulto (persona con edad a partir de 18 años), la mujer unida, el adolescente y el adulto mayor (personas con edad igual o mayor a 60 años).

La muestra del adulto mayor recogida por el estudio epidemiológico fue 632 sujetos. Para el presente estudio solo se consideraron dentro de la muestra a los adultos mayores que no presentaron ninguna discapacidad física o ajena al deterioro cognoscitivo que impida responder a cualquiera de los ítems que componen el instrumento, que en total fueron 544, 285 varones y 259 mujeres; con edad media de 68 años, el 48.9% tiene nivel primaria y un 28.9% nivel secundaria.

Los resultados presentados a continuación comprenden: comprobación del supuesto de unidimensionalidad, evaluación de los ítems bajo la teoría clásica (dificultad del ítem bajo el modelo clásico, correlación ítem – test, coeficiente alfa de cronbach), evaluación de los ítems bajo el modelo Rasch (estimación de la dificultad del ítem y la función cognoscitiva, bondad de ajuste del modelo en el test y los ítems, función de información).

Según el enfoque clásico, los índices de dificultad de la mayoría de los ítems están próximos a 1, considerándose como ítems de baja dificultad o “fáciles”. Tomando en cuenta la clasificación expuesta por Abad y col³, se tiene que 22 ítems (73.3%) son considerados ítems de baja dificultad o “fáciles”; 6 ítems (20%) son considerados como de dificultad media y los dos ítems restantes son de gran dificultad.

Según la clasificación mencionada por Abad (Abad y col.) los ítems con menor dificultad varían entre 0.71 y 1; los de nivel medio entre 0.3 y 0.7, y los ítems con mayor dificultad entre 0 y 0.29.

Se identificaron 4 ítems con una baja correlación respecto al puntaje del test, los mismos ítems son considerados como ítems difíciles. Como referencia de la confiabilidad del test, se calculó en índice alfa de cronbach, dando una confiabilidad de 0.848 considerado dentro de la categoría muy confiable.

Como se mencionó en el capítulo II, los modelos Rasch generalmente usados presentan un coeficiente constante denominada “D”, las dos variantes usadas con mayor frecuencia son aquellos modelos cuyo valor D es igual a 1 y 1.7. Se plantearon dos modelos: modelo I (D=1) y modelo II (D=1.7) con la finalidad de encontrar aquel modelo que represente de mejor manera a los datos.

La prueba de bondad de ajuste del modelo no encontró evidencia de que existe discrepancia significativa entre los modelos y los datos reales (Tabla 2.1).

Tabla 2.1

Bondad de ajuste usando el estadístico X^2 bootstrap de Pearson

Modelo	T	n	p
Modelo I (D=1)	140977.4	200	0.115
Modelo II (D=1.7)	139132.9	200	0.225

Los resultados de test de razón de verosimilitudes (ver Tabla 2.2) ha permitido concluir que uno de los modelos es mejor que el otro (LRT: 131.97; $p < 0.05$), por el criterio de Akaike se concluyó que modelo II explica mejor los datos que el modelo I. A través la eficiencia relativa (Tabla 2.3) entre el modelo II vs. el modelo I nos muestra que el modelo II es más eficiente debido a que se obtienen 1.7 veces mayor información comparado con el modelo I.

Tabla 2.2

Tabla de comparación de modelos

Modelo	AIC	BIC	Logaritmo de verosimilitud	Test de razón de verosimilitud	p
Modelo I (D=1)	10210.53	10339.5	-5075.26	131.97	0.000*
Modelo II (D=1.7)	10078.56	10207.52	-5009.28		

Tabla 2.3

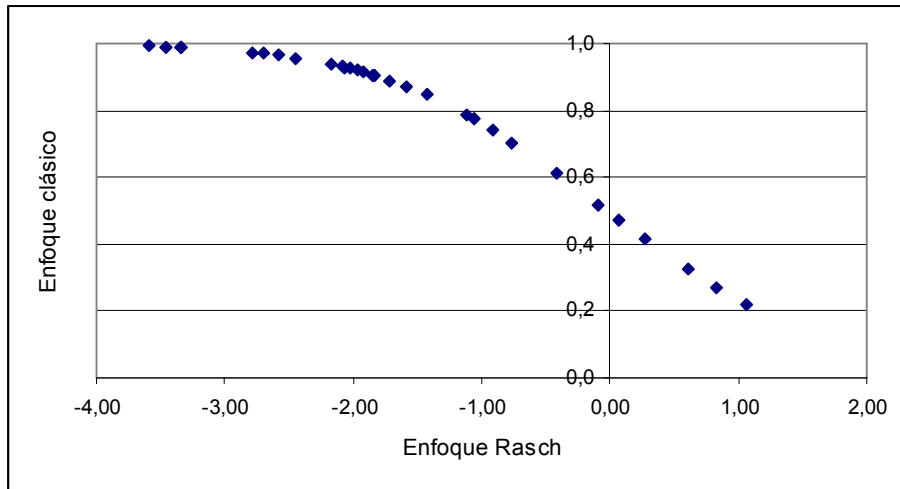
Eficiencia relativa entre el Modelo II vs. el modelo I

Modelo	Total	Eficiencia
Modelo I (D=1)	29.94	
Modelo II (D=1.7)	51	1.70

Los índices de dificultad estimados por el modelo Rasch II (D=1.7) mantienen una relación no lineal con los índices de dificultad bajo el modelo clásico, esto es coherente ya que a mayor proporción de respuesta el ítem es fácil, mientras crece la dificultad del ítem se espera que la proporción de respuesta disminuya, a pesar que las estimaciones de la dificultad de los ítems se encuentra en escalas distintas según cada enfoque. (Figura 2.1).

Figura 2.1

Relación entre los índice de dificultad por el modelo clásico y el modelo Rasch.



Los ítems más difíciles fueron aquellos que corresponden al cálculo, praxis (copiar una figura), y aprendizaje verbal (memoria de palabras).

Se realizó una clasificación de los ítems a través del criterio de un juez experto y se comparó con la clasificación usando como referencia los puntos de corte planteados por Folstein. Los resultados encontrados muestran que el 50% (15) de los ítems concuerdan con la clasificación asignada en ambos criterios.

Las pruebas de bondad de ajuste para cada ítem indicaron que 20 de los 30 ítems se ajustan bien al modelo Rasch. Los ítems que no se ajustaron al modelo Rasch pertenecen a las áreas de Atención y cálculo (5), aprendizaje verbal (2), lenguaje (2) y praxis (1) y están enlistados en la Tabla G en Anexos. Los ítems para los cuales el modelo Rasch no se ajusta son:

- Los ítems PP15 y PP18 que fueron los ítems con menor dificultad, siendo clasificados bajo el criterio de Folstein como ítems fáciles de resolver para individuos con deterioro cognoscitivo severo, leve o normal.
- El ítem PP12a fue calificado bajo el criterio de Folstein como ítem fácil de resolver para individuos con deterioro leve o normal.

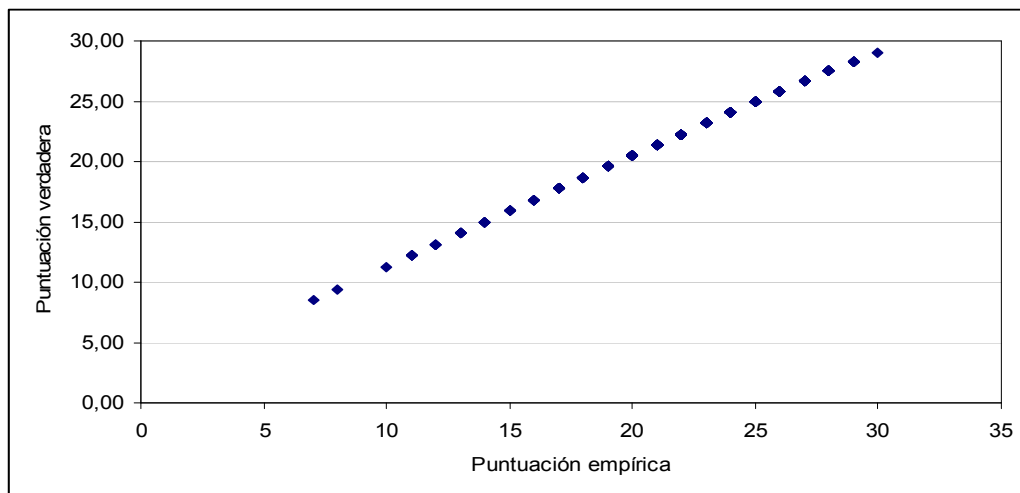
- Los ítems PP12b, PP12c, PP12d, PP12e, PP22b, PP14b, PP14c fueron los ítems de mayor dificultad y calificados bajo el criterio de Folstein como ítems fácil de resolver para individuos con nivel cognoscitivo normal.

A manera de exploración se compararon los porcentajes de acierto y las probabilidades de responder correctamente en cada ítem mencionado anteriormente, según el nivel cognoscitivo estimado $\hat{\theta}$.

La puntuación verdadera del test definida como la suma de probabilidades de acertar cada ítem por un sujeto, esta directamente relacionada con la puntuación empírica tal como se observa en el Figura 2.2.

Figura 2.2

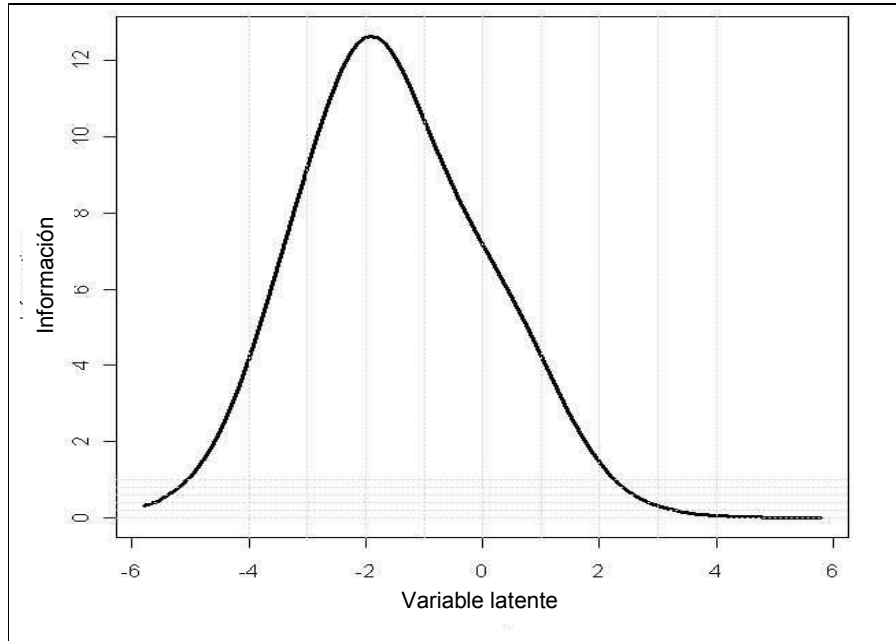
Puntuación verdadera del test y puntuación empírica



La función de información cuantifica la cantidad de información que es medida por el test. A través de la grafica de la función de información encontramos solo un punto donde la función se maximiza; que está alrededor de -1.9 y -2 (Figura 2.3)

Figura 2.3

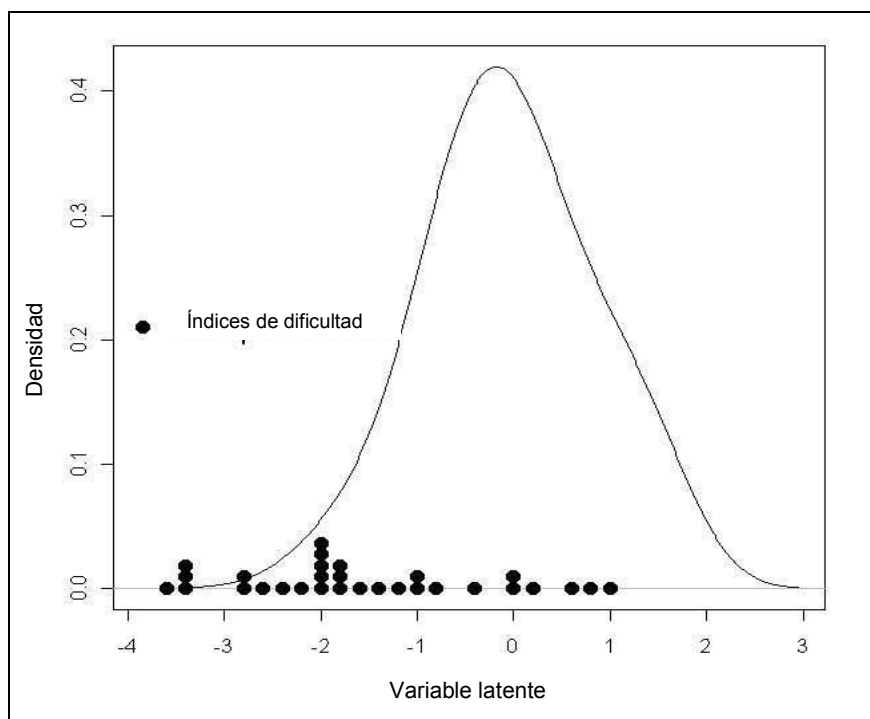
Función de Información del Test



En la figura 2.4 se muestran el estimador Kernel de la función de densidad para la función cognoscitiva en la muestra bajo estudio y un gráfico de frecuencias de la dificultad de los ítems, donde se aprecia que los valores de la variable latente $\hat{\theta}$ se encuentra por encima de la dificultad o exigencia de la mayoría de los ítems que componen al instrumento. Para describir mejor la función cognoscitiva y la dificultad de los ítems se estimaron estadísticas descriptivas que se presentan a continuación en la tabla 2.4. El nivel cognoscitivo medio de los entrevistados se centra alrededor de 0.16; en tanto que nivel de dificultad de los ítems se centra alrededor de -1.55.

Figura 2.4

Función de densidad estimada de Kernel para la función cognoscitiva de los adultos mayores



La función cognoscitiva media entre los adultos mayores es superior al nivel de dificultad media de los ítems que componen al instrumento, el 75% de los ítems no superen en nivel de dificultad -0.68, y este valor de dificultad puede ser superado por lo menos por el 75% de los entrevistados que tienen un nivel cognoscitivo mayor a -0.46. Además el 75% de los entrevistados se encuentran por encima del nivel medio de dificultad de los ítems (Tabla 2.4).

Tabla 2.4

Nivel de las funciones cognitivas de los adultos mayores e Índices de dificultad de los ítems

Función cognoscitiva estimada	b	θ
Media	-1.55	0,16
Intervalo de confianza para la media al 95%	Límite inferior	0,06
	Límite superior	0,26
Desviación estándar	1.29	1,20
Mínimo	-3.59	-5,68
Máximo	1.07	1,72
Mediana	-1.85	0,48
Percentil 25	-2.48	-0.46
Percentil 75	-0.68	1.02
Asimetría	0.49	-1,54
Curtosis	-0.58	3,19

Aplicando los puntos de corte según el criterio de juez experto y el criterio de punto de corte por orden de Folstein, comparándolos con el diagnóstico según puntaje total se encontró aplicando el criterio combinado de Folstein una eficiencia relativa del 52.21%, al momento de clasificar comparando con el procedimiento clásico de Folstein (Tabla 2.5), mientras que bajo el criterio del juez experto se tienen 47.43% de clasificación correcta (Tabla 2.6).

Tabla 2.5

Diagnóstico del nivel cognoscitivo según criterio punto de corte por orden Folstein vs. criterio por el puntaje total.

Punto de corte por orden de Folstein	Puntaje total		
	Normal	Leve	Severo
Normal	256 47,06%	155 28,49%	35 6,43%
Leve	23 4,23%	24 4,41%	9 1,65%
Severo	28 5,15%	10 1,84%	4 0,74%

Tabla 2.6

Diagnóstico del nivel cognoscitivo según criterio de juez experto vs. criterio por el puntaje total.

Juez experto	Puntaje total		
	Normal	Leve	Severo
Normal	180 33,09%	105 19,30%	20 3,68%
Leve	99 18,20%	74 13,60%	24 4,41%
Severo	28 5,15%	10 1,84%	4 0,74%

DISCUSIÓN

El uso del modelo Rasch como alternativa al modelo clásico, brinda información adicional sobre el test y los ítems que lo componen. Los resultados muestran que aplicar un modelo Rasch con una constante de curvatura $D=1.7$ permite explicar mejor la información registrada en los datos en lugar del modelo Rasch clásico ($D=1$).

Las puntuaciones estimadas por el modelo Rasch mantienen una relación no lineal con las puntuaciones observadas. La función de información del test alcanza su máximo cuando la habilidad del sujeto está alrededor de -1.97 , siendo este valor inferior al punto de corte para clasificación entre el deterioro cognoscitivo severo y leve, viendo gráficamente que la función de información es unimodal, el instrumento está diseñado para detectar con mayor eficiencia la estimación y clasificación del deterioro cognoscitivo severo alrededor del -1.97 . El análisis de bondad de ajuste por ítem reveló que el modelo Rasch con constante $D = 1.7$ no es el más apropiado para el ajuste de ciertos ítems. Los ítems que no ajustan al modelo Rasch pertenecen a las áreas de atención y cálculo, aprendizaje verbal, lenguaje y praxis. Haciendo referencia sobre otros trabajos, Robles comenta que los ítems que presentan mayores problemas en responder son cálculo, aprendizaje verbal, praxis y orientación en el tiempo. Los resultados encontrados son consistentes con los comentarios de Robles, específicamente en las áreas de cálculo, aprendizaje verbal y praxis, además se observó que la relación entre el porcentaje de acierto y la variable latente es inversa, es decir, que para valores muy bajos en variable latente existe bajo porcentaje de aciertos en algunos ítems pero alto porcentaje de aciertos en otros, mientras que para altos valores en la variable latente sucede lo contrario, ante esta situación Baker (2001) comentó que esto puede ser modelado considerando índices de discriminación diferentes para cada ítem, pero además menciona que la aparición de ítems cuyo comportamiento es inverso al modelo o con discriminación negativa deben ser eliminados del test o considerar modelos más elaborados. Estos resultados revelan que la variable latente ha polarizado a los ítems en dos bloques: el primer bloque comprendido por los ítems de cálculo, aprendizaje verbal y praxis y el segundo bloque

comprendido los ítems de lenguaje, orientación en lugar, orientación en el tiempo; y que posiblemente exista una estructura compleja que explique mejor los datos. Este comportamiento no esperado podría tener una explicación basado en un mecanismo compensatorio. En el proceso de deterioro las funciones cognoscitivas van disminuyendo, generalmente este proceso no afecta todas las áreas simultáneamente, sino que afectan primero alguna de las áreas específicas si es que esto es causado por algún problema de fondo, ante esta situación el individuo en un proceso de adaptación va desarrollando las otras áreas del cerebro no afectadas como mecanismo de compensación.

Se identificaron dos ítems donde la probabilidad de responder correctamente es casi constante para cualquier valor de la variable latente, lo cual indica que no tiene relación alguna con la variable latente, y en general el porcentaje de acierto está muy cerca del 100% por lo que estos ítems no son muy eficientes por su baja dificultad.

Existen otros modelos dentro de la teoría de respuesta a los ítems que pueden explicar mejor la respuesta de los entrevistados. La revisión bibliográfica ha mostrado que la mayoría de investigadores prefieren utilizar el modelo Rasch por su parsimonia y que es capaz de representar satisfactoriamente las respuestas de los encuestados. En el presente estudio la prueba de ajuste del instrumento no encontró diferencias significativas entre las respuestas registradas en el instrumento y el modelo propuesto. No obstante, será interesante realizar futuras investigaciones bajo otros modelos y poder compararlos con el modelo Rasch a fin de identificar el modelo que mejor se ajuste a los datos y con menor costo de información.

CONCLUSIONES

Los resultados de la TRI a través del modelo Rasch detectaron que los ítems se agrupan en dos bloques que expresan áreas de las funciones cognitivas que son contrarias entre sí.

La mayoría de ítems presentan niveles de dificultad por debajo del punto de corte para clasificar el deterioro severo y el punto máximo de información del test localizado se encuentra alrededor de -1.97 indicando que el instrumento es adecuado para identificar a individuos con deterioro cognoscitivo severo, pero poco plausible para persona con deterioro cognoscitivo leve o normal.

La mayoría de los ítems tienen índices de dificultad por debajo del rendimiento del primer cuartil de la población.

El modelo Rasch mostró ser un modelo adecuado para explicar las respuestas de los sujetos a la mayoría de los ítems, sin embargo existen otros modelos de variables latentes que podrían ser evaluados para su aplicación en los datos, particularmente los modelos que consideren más de una variable latente, y que incluyen además otros índices como capacidad de discriminación del ítem.

BIBLIOGRAFÍA.

- ABAD F., PONSODA V., OLEA J. Introducción a la psicometría. Teoría de los test y Teoría de respuesta al ítems, Univeridad Autonoma de Madrid, 2006. Internet 20 Setiembre de 2009, disponible en:
http://www.uam.es/personal_pdi/psicologia/fjabad/PSICOMETRIA/MATERIALES/IntPs_5.pdf
- BAKER F. The Basics of item response theory. ERIC Clearinghouse on Assessment and evaluation, 2001. Internet 29 diciembre de 2009, disponible en:
<http://info.worldbank.org/etools/docs/library/117765/Item%20Response%20Theory%20-%20F%20Baker.pdf>
- BARTOLOMEW D. J. The analysis and interpretation of multivariate data for social Scientists. CHAPMAN & HALL/CRC, 2002.
- BOLFARINE H. Inferencia Estadística para modelos de evaluación educativa. VII Coloquio Nacional de Estadística. UNMSM 2003.
- DOBSON, Annette J. An introduction to Generalized Linear Models. CHAPMAN & HALL, 1990.
- HOSMER David W., LAMESHOW Stanley. Applied Logistic Regression. Jhon Wiley & Sons, Inc. 2002.

- MAIR P., S. REISE, P. M. BENTLER. IRT Goodness-of-Fit Using Approaches from Logistic Regression. Internet 18 Junio 2009.
<http://preprints.stat.ucla.edu/540/logregIRTucla.pdf>
- MUÑIZ J. F. Introducción a la teoría de respuesta a los ítems. Ediciones Pirámide, S.A. 1997.
- PEÑA, Daniel. Análisis de Datos multivariantes. McGRAW-HILL /INTERAMERICANA DE ESPAÑA, 2002.
- RIZOPOULOS D. Irm: An R Package for latent variable modeling and Item Response Theory analyses *Journal of Statistical Software*, November 2006, Volumen 1, Issue 5.
<http://www.jstatsoft.org/v17/i05/paper>
- ROBLES A. Y. Adaptación del Mini Mental State Examination [Tesis de maestría] 2003, UPG de Psicología, UNMSM.
- INSTITUTO NACIONAL DE SALUD MENTAL. Estudio epidemiológico de salud mental en Lima metropolitana 2002. *Anales de Salud Mental* 2002; 18 Número 1-2.

Apéndice I

MODELO LINEAL GENERAL

El modelo lineal general engloba a una familia de modelos utilizado para describir una relación entre una variable observada y un conjunto de variables a las que se denomina predictoras. La estructura del modelo comprende tres componentes: la variable respuesta, el predictor o componente sistemática y la componente aleatoria o error aleatorio. La característica principal de este conjunto de modelos es que la variable respuesta es una variable aleatoria cuya distribución de probabilidad es $N(\mu, \sigma^2)$. Además se asume que los coeficientes o parámetros dentro del modelo son lineales y la aditividad del predictor lineal.

Definición 1. Sea Y una variable aleatoria con distribución $N(\mu, \sigma^2)$, y sea X un conjunto de variables observadas, entonces

$$\underbrace{Y}_{\text{variable aleatoria}} = \underbrace{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k}_{\text{predictor lineal } (\eta)} + \underbrace{e}_{\text{error aleatorio}}$$

Donde, e tienen distribución $N(0, \sigma^2)$, X_i son las variables predictivas, β_i son los coeficientes de los parámetros.

Este modelo fue formulado bajo algunos supuestos:

1. $E(Y | X_1 = x_1, \dots, X_k = x_k) = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$
 $E(Y | X_1 = x_1, \dots, X_k = x_k) = \mu = \eta$
2. $\text{Var}(Y | X_1 = x_1, \dots, X_k = x_k) = \sigma^2 < \infty$ (Constante)

3. $Cov(Y_i, Y_j) = 0$ (independencia)

4. $Y | X_1 = x_1, \dots, X_k = x_k \sim N(\mu, \sigma^2)$

La violación de cualquiera de los supuestos mencionados invalida el análisis de dichos modelos.

Estos modelos surgieron a partir del siglo XVII con Gauss, posteriormente los avances en la teoría estadística y el desarrollo computacional han permitido usar métodos análogos que nos ayuden a desarrollar modelos lineales en situaciones donde la variable respuesta tiene una distribución diferente a la Normal, y la relación entre las variables respuestas y las variables predictoras no necesariamente son de la forma lineal simple.

Uno de los principales hallazgos que han permitido extender el uso de los modelos lineales fue la identificación de una clase de distribuciones de probabilidad que poseen estadísticas suficientes para la estimación de sus parámetros (Fisher, 1934). Esta clase de distribuciones se les denominaron familia de distribución exponencial, debida a que dichas distribuciones pueden ser expresadas bajo una forma canónica.

$$f(Y, \theta) = \exp\{a(Y)b(\theta) + c(\theta) + d(Y)\}$$

A partir de esta definición, se llega a relacionar la variable aleatoria con el predictor lineal a través de una función, denominada función de enlace.

Modelos Lineales Generalizados

Definición 2. Sea Y una variable aleatoria independiente cuya distribución pertenece a la familia de distribución exponencial, tal que la función de distribución para cada Y se puede expresar de la siguiente forma

$$f(Y, \theta) = \exp\{a(Y)b(\theta) + c(\theta) + d(Y)\}$$

$a(Y)$: Estadística suficiente.

$b(\theta)$: Función del parámetro de localización.

$c(\theta)$: Función únicamente del parámetro.

$d(Y)$: Función únicamente de la estadística.

El modelo lineal generalizado está dado de la forma:

$$E(Y | X_1 = x_1, \dots, X_k = x_k) = g^{-1}(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)$$

Cuyas componentes son:

- La variable respuesta o componente aleatoria Y , la cual posee una distribución de probabilidades que pertenece a la familia exponencial y cuya esperanza es dada por:

$$E(Y | X_1 = x_1, \dots, X_k = x_k) = \mu$$

- El predictor lineal (η). Corresponde a la combinación lineal de las variables explicativas X_1, \dots, X_k . (regresión lineal)

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

- La función de enlace (g). Función matemática monótona y derivable que enlaza $E(Y)$ con η .

$$E(Y | X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = g^{-1}(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)$$

Los modelos lineales generalizados han permitido estimar los parámetros en modelos donde la variable respuesta no necesariamente cumple el supuesto de normalidad, así pues en las diversas disciplinas de investigación este aporte permite observar fenómenos como la frecuencia en que ocurre un evento en determinado tiempo o espacio, o simplemente la ocurrencia de un evento. Tales fenómenos están asociados a una distribución de probabilidad conocida, como en el primer caso el fenómeno está asociado a la distribución Poisson y el modelo recibe en nombre de modelo Poisson, mientras el segundo caso está asociado a la distribución Bernoulli y el modelo para esta variable recibe el nombre de modelo logístico.

El modelo Logístico.

El modelo logístico es aplicado en situaciones donde el fenómeno a ser modelado es simplemente la ocurrencia de un evento, es decir se observa o no la presencia de un atributo o característica de interés. La definición formal de este fenómeno está dado de la siguiente manera:

Definición 3. Sea $Y_i \square B(\pi_i, 1)$, donde la esperanza de Y_i es la probabilidad de éxito π_i , y sea $X = (X_1, X_2, \dots, X_p)$ variables predictivas (continuas o discretas), donde. La función de enlace para este caso es la función logit, definida como:

$$\text{logit}(\pi_i(X)) = \ln\left(\frac{\pi_i(X)}{1 - \pi_i(X)}\right) = \beta_{i0} + \sum_{j=1}^p \beta_{ij} X_j$$

Supuestos:

1. Las respuestas (y_i) : $i=1, 2, \dots, n$ son independientes
2. Las variables explicativas (X_1, X_2, \dots, X_k) son independientes entre si.
3. $Y_i \square B(\pi_i)$; $i=1, 2, \dots, n$ (Distribución Bernoulli)
4. $V(Y_i) = \pi_i(1 - \pi_i)$; $i=1, 2, \dots, n$ (Heterocedasticidad)

Mediante la transformación logit podemos obtener un modelo lineal. La

probabilidad denota la probabilidad de "éxito" y la razón $\frac{\pi_i(y)}{1 - \pi_i(y)}$ se conoce

como la chance (odd) de "éxito".

El procedimiento más utilizado para realizar la estimación es el método score de Fisher el cual se basa en la función score obtenida a partir de la función de verosimilitud (Nelder y Wedderburn, 1972).

Para una muestra de tamaño n , se tiene el vector $\mathbf{y} = (y_1, \dots, y_n)$

$$L(\beta) = L(\beta, y, X) = \prod_{i=1}^n \pi_i^{y_i} (x_i) (1 - \pi_i(x_i))^{1 - y_i}$$

Calculando el logaritmo de la función de verosimilitud

$$l(\beta) = Ln[L(\beta, y, X)] = \sum_{i=1}^n y_i \ln \left(\frac{\pi_i(x_i)}{1 - \pi_i(x_i)} \right) + \sum_{i=1}^n \ln(1 - \pi_i(x_i))$$

Apéndice II

DIAGNÓSTICO TEMPRANO DEL NIVEL DE DETERIORO COGNOSCITIVO DEL ADULTO MAYOR

Durante mucho tiempo las políticas de todos los países elaboraron programas que ayuden a mejorar la calidad de vida de los individuos desde sus inicios hasta la etapa adulta, y solo proveían las atenciones básicas para la vejez. Pero un fenómeno está cambiando el rostro del mundo. La población mundial está envejeciendo. Según OMS, la población mundial de adulto mayor de 60 años se triplicará entre el 2000 al 2050, especialmente en los países en desarrollo. Este proceso de envejecimiento es consecuencia de los cambios de los niveles de calidad de vida, el aumento de la esperanza de vida y la disminución de los índices de fecundidad y natalidad. Esta situación trae consigo nuevas demandas de la población en temas de salud pública, pensión y empleo, integración social, etc.

Proyecciones del INEI advierte que para el 2010 habrán 1728 759 habitantes mayores a 64 años⁸, un número considerable de habitantes que a esa edad ya están formando parte del grupo de jubilados y retirados. Publicaciones de la Naciones Unidas informan que entre el 2010 y el 2050 el incremento de la tasa de crecimiento de personas mayores de 60 años se acercará al 4% medio por año. Por tanto la población se verá afectada por un incremento de carga de dependencia que trae consigo el mantenimiento del adulto mayor de 60 años.

En el Perú, el INSM (2002) ejecutó un estudio epidemiológico en la población adulto mayor de Lima metropolitana, encontrando que en ese momento el 10.5% de los adultos mayores presentaban un funcionamiento cognoscitivo

anormal, siendo este problema mayor en los adultos mayores de 75 años (30.2%).

El reto del mundo actual frente al envejecimiento poblacional es sin duda brindar una buena calidad de vida a las personas para que al momento de llegar a la tercera edad, ellos puedan seguir viviendo de forma independiente y puedan seguir contribuyendo en la economía familiar.

Para esto es importante contar con estrategias que procuren brindar una buena salud y bienestar social, así como de indicadores confiables que permitan una evaluación del estado actual de esta creciente subpoblación que va tomando mayor importancia dentro de la estructura poblacional.

Considerando la importancia de la salud física y mental del adulto mayor, la vigilancia de su salud permitirá al adulto mayor permanecer como agente activo y productivo dentro de la sociedad.

El deterioro cognoscitivo es la pérdida o alteración de las funciones mentales, tales como memoria, orientación, lenguaje, reconocimiento visual, conducta, que interfiere con la actividad e interacción social de la persona afectada. Este deterioro ocurre principalmente en la vejez, aunque también existen situaciones donde el deterioro es causado por problemas ajenos a la edad de la persona. El deterioro cognoscitivo y la demencia son trastornos conceptualmente similares, pero la diferencia entre ambos se encuentra en el grado de la alteración en las funciones del individuo. En consideración al proceso de envejecimiento, se prevé normal que aparezca el deterioro cognoscitivo en el adulto mayor como consecuencia del envejecimiento cerebral, pero el trastorno de demencia es una patología que minimiza las funciones mentales del individuo a un punto de dependencia absoluta, sin posibilidad de disfrute de la vida. El diagnóstico temprano del deterioro cognoscitivo en el adulto mayor permite tratar el problema procurando una mayor eficacia del tratamiento para evitar el avance del deterioro y pueda mantener su autonomía para cumplir sus funciones normales y mantener una mejor calidad de vida.

Para el diagnóstico del nivel de deterioro cognoscitivo y concluir si se encuentra desarrollando un trastorno de demencia se requieren de pruebas

neuropsicológicas que permiten identificar el nivel de las habilidades o capacidades cognoscitivas del adulto mayor.

Existe una lista de pruebas que miden el nivel cognoscitivo según el tipo de deterioro o el área cognoscitiva que es afectada.

Pruebas generales orientadoras o de catastro, orientadas a demencia en general, lenguaje, memoria, capacidad visuoespacial, demencia de la depresión y de gradación de la demencia.

En 1975, Marshal F Folstein junto a sus colaboradores Mc Hugh P. R. y Folstein S. E. elaboraron un breve instrumento (30 ítems) que permitía cuantificar el nivel cognoscitivo de las personas (en particular adultos mayores), con la finalidad de contar con una herramienta de apoyo para estudios clínicos sobre trastornos relacionados con la disminución de la destreza mental como son la demencia por Alzheimer y la depresión aguda. El instrumento conocido como Mini Mental State Examination está estructurado en 5 áreas descritas en la tabla II.1, sus autores diseñaron este instrumento para la detección del deterioro severo de las funciones cognoscitivas aunque también detecta la presencia de un leve deterioro cognoscitivo.

Tabla II.1

Estructura del Mini Mental State Examination de Folstein

Procesos Cognoscitivos	Dimensión
Orientación: Proceso complejo de ubicación de si en relación al entorno, que demanda atención, percepción y memoria, así como conocimiento de categoría relativas a tiempo y a lugar.	<ol style="list-style-type: none"> 1. Orientación en el tiempo (día, mes, año, día de la semana, estación) 5 pts. 2. Orientación en lugar (nombre/tipo, piso/calle, distrito, provincia, departamento) 5 pts.
Aprendizaje verbal: Eficiencia de la memoria a través de la amplitud de registro, retención inmediata y evocación posterior de la palabra relacionado a un objeto.	<ol style="list-style-type: none"> 3. Registro de tres palabras (objetos). 3pts. 4. Evocación de las tres palabras (objetos). 3 pts.
Atención: Focalización del procesamiento y división de la atención, realizando cálculos sucesivos o deletreo inverso.	<ol style="list-style-type: none"> 5. Serie de resta de 7 en 7 a partir de 100. 5 pts.
Lenguaje: Competencia en el uso de los símbolos verbales. Evocación a palabras asociadas a imagen visual. Registro, análisis fonético, retención y articulación verbal. Registro, análisis fonético, retención y comprensión verbal y ejecución motora. Lectura, comprensión verbal y ejecución motora. Generación de ideas, transformación gramatical, representación grafémica y escritura.	<ol style="list-style-type: none"> 6. Denominación de dos objetos. 2 pts. 7. Repetición de una frase. 1 pto. 8. Orden oral. 3 pts. 9. Orden escrita. 1 pto. 10. Escritura de una frase. 1 pto.
Praxis: dibujo de copia de estímulo visual, integrando percepción visual, análisis visual-espacial, planteamiento u ejecución motora.	<ol style="list-style-type: none"> 11. Copia de un dibujo de dos polígonos sobrepuestos. 1 pto.

Apéndice III

USO DEL SOFTWARE R PARA APLICAR MODELO TRI

El análisis estadístico sobre la TRI se realizó con el software R, el procedimiento es explicado a continuación.

1. Se inicia la sesión cargando la base de datos que debe estar guardada en la carpeta con la función *read.table*.

```
amayor<-read.table("G:/TRI, IRT/amayor2002.txt", header=T)
```

Para esto contamos con la base de datos en un archivo de texto con extensión txt, cuyo nombre es amayor2002.txt.

2. Se estimaron estadísticas descriptivas sobre las características del test, coeficiente correlación biserial y alfa de cronbach a través de la función *descript*.

```
descript(amayor)
```

3. Se pide estimar 2 modelos Rasch, el primero con una constante D=1 y el segundo con constante D=1.7.

```
fit1<-rasch(amayor, constraint=cbind(length(amayor)+1,1))
```

```
fit2<-rasch(amayor, constraint=cbind(length(amayor)+1,1.7))
```

4. Obtenemos estadísticos resúmenes del modelo como la estimación de la dificultad de los ítems y el error típico de estimación para cada uno, el logaritmo de la verosimilitud y el índice de Akaike.

```
summary(fit1)
```

```
summary(fit2)
```

5. Posteriormente, se aplicó la prueba de bondad de ajuste para ambos modelos a través *GoF, rasch* para determinar si los modelo se ajusta a los datos que provienen de todo el test. La prueba de bondad de ajuste aplicado es una prueba bootstrap basado en la Chi cuadrado de Pearson.

```
GoF.rasch(fit2,B=199)
```

6. Comparamos ambos modelo a través del test de razón de verosimilitud (LRT) , el cual incluye los índices AIC y BIC Con la función *anova*. También se evaluaron la bondad de ajuste del modelo para cada ítem por medio de la función *ítem.fit*.

```
anova(fit1,fit2)
```

```
item.fit(rasch(amayor, constraint=cbind(length(amayor)+1,1.7)))
```

7. La función *factor.scores* estima las probabilidades condicionales, es decir la probabilidad de acertar el ítem, para luego calcular la puntuación verdadera del test como la suma de las probabilidades.

```
factor.scores(fit2)
```

```
fitted(fit2, type="conditional-probabilities")
```

8. Para estimar el nivel o rasgo latente del individuo, se usa la función *person.fit*.

```
person.fit(rasch(amayor))
```

9. La información del test es calculada por la función *information*, considerando un intervalo de -3 a 3.

```
information(fit2, c(-3, 0))
```

```
information(fit2, c(0, 3), items=NULL)
```

10. La gráfica de la función información del test se realiza por medio de los siguientes comandos

```
plot(fit2, type=c("IIC"), items =0, lwd = 3, cex = 1.4)
```

```
abline(v = -4:4, h = seq(0, 1, 0.2), col = "lightgray", lty = "dotted")
```

11. Asimismo se puede graficar la curva característica para los ítems de la siguiente manera

```
plot(fit2, type=c("ICC"), items =NULL, lwd = 3, cex = 1.4)
```

```
abline(v = -4:4, h = seq(0, 1, 0.2), col = "lightgray", lty = "dotted")
```

Apéndice IV

COMPROBACIÓN DEL SUPUESTO DE UNIDIMENSIONALIDAD

La comprobación de la unidimensionalidad del test se realizó a través del análisis factorial. El análisis factorial se efectuó a través de la matriz de correlación tetracórica, especificando que muestren 7 dimensiones correspondientes al número de áreas que componen al test.

Los resultados presentados en la matriz de correlaciones tetracóricas indican que existen 5 ítems asociados perfectamente ($r_t = 1$), por lo que fueron retirados dentro del análisis factorial y se procedió con el análisis considerando a los 25 ítems restantes.

Los resultados presentados en la Tabla 1 muestran que el 37.87% de la información está siendo explicada por la primera dimensión, mientras que la segunda dimensión aporta tan solo el 16.49%, las dimensiones restantes explican en menor porcentaje la información registradas en los datos.

Tabla IV.1.

Factores y porcentaje de varianza explicada

Factores	Autovalores	% varianza
1	9,467	37,87
2	4,123	16,49
3	2,386	9,542
4	2,164	8,656
5	1,71	6,841
6	1,614	6,457
7	1,355	5,42

1/ Componentes principales

Figura IV.1

Gráfico de la ladera para determinar en número de dimensiones relevantes.

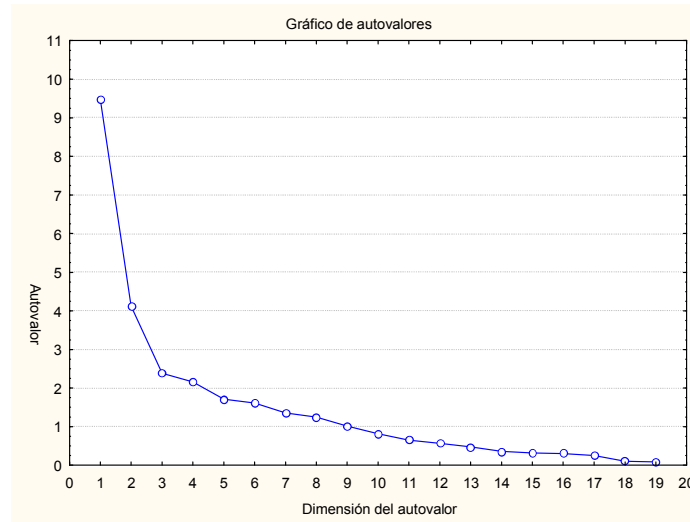


Tabla IV.2

Matriz de correlaciones tetracóricas

	PP1	PP2	PP3	PP4	PP5	PP6	PP7	PP8	PP9	PP10	PP11a	PP11b	PP11c	PP12a	PP12b	PP12c	PP12d	PP12e	PP14a	PP14b	PP14c	PP15	PP16	PP17	PP18	PP20a	PP20b	PP20c	PP21	
PP2	0.59	1.00																												
PP3	0.72	0.70	1.00																											
PP4	0.57	0.55	0.64	1.00																										
PP5	0.39	0.49	0.46	0.53	1.00																									
PP6	0.37	0.54	0.68	0.52	0.49	1.00																								
PP7	0.23	0.30	0.20	0.64	0.47	0.49	1.00																							
PP8	0.45	0.44	0.15	0.43	0.24	0.43	0.71	1.00																						
PP9	0.50	0.37	0.45	0.48	0.39	0.66	0.64	0.47	1.00																					
PP10	0.15	0.48	0.50	0.38	0.29	0.45	0.48	-0.04	0.45	1.00																				
PP11a	0.20	0.45	0.51	0.35	0.27	0.63	-0.79	-0.80	0.23	0.37	1.00																			
PP11b	0.30	1.00	0.62	0.48	0.41	0.95	-0.78	0.28	0.57	0.46	0.94	1.00																		
PP11c	0.01	0.19	0.30	0.34	0.03	0.56	-0.83	0.16	0.44	0.19	0.74	0.85	1.00																	
PP12a	0.52	0.49	0.60	0.58	0.40	0.52	0.61	0.45	0.55	0.46	0.29	1.00	0.25	1.00																
PP12b	0.67	0.42	0.57	0.60	0.30	0.52	0.50	0.45	0.51	1.00	0.87	0.87	0.47	0.88	1.00															
PP12c	0.51	0.33	0.38	0.58	0.20	0.42	0.58	0.40	0.44	0.89	1.00	1.00	0.88	0.95	0.97	1.00														
PP12d	1.00	0.32	0.32	0.52	0.20	0.40	0.53	0.33	0.51	0.88	1.00	1.00	0.87	0.94	0.94	0.99	1.00													
PP12e	0.89	0.24	0.39	0.46	0.21	0.39	0.47	0.26	0.48	0.87	1.00	1.00	0.87	1.00	0.92	0.98	0.99	1.00												
PP14a	0.27	0.37	0.28	0.45	0.42	0.36	0.47	0.36	0.35	0.28	0.25	1.00	0.03	0.43	0.40	0.40	0.42	0.41	1.00											
PP14b	0.11	0.22	0.16	0.28	0.41	0.34	0.22	0.20	0.22	0.19	0.11	1.00	-0.21	0.33	0.32	0.35	0.30	0.31	0.71	1.00										
PP14c	0.23	0.22	0.15	0.32	0.23	0.27	0.17	0.30	0.26	-0.12	0.18	0.88	0.26	0.32	0.26	0.23	0.23	0.21	0.55	0.58	1.00									
PP15	0.24	0.20	0.35	0.18	-0.83	0.40	-0.79	0.23	0.46	0.41	-0.73	-0.72	-0.76	0.55	0.22	0.14	0.09	0.03	0.33	0.40	0.32	1.00								
PP16	-0.80	0.14	-0.78	-0.81	0.06	-0.81	0.22	-0.80	0.02	-0.76	-0.73	-0.73	-0.76	0.29	0.27	0.19	0.13	0.07	0.25	0.11	0.18	0.79	1.00							
PP17	0.32	0.39	0.23	0.40	0.15	0.45	0.50	0.45	0.54	0.44	0.55	0.66	0.47	0.43	0.33	0.24	0.40	0.35	0.32	0.22	0.23	0.74	0.55	1.00						
PP18	0.45	0.40	0.53	0.62	0.31	0.45	0.35	0.28	0.41	0.32	0.36	0.68	0.36	0.46	0.41	0.39	0.32	0.24	0.31	0.28	0.25	0.42	0.36	0.59	1.00					
PP20a	0.06	0.16	-0.06	0.26	0.34	0.07	0.32	0.03	0.33	0.15	0.10	0.20	0.07	0.25	0.28	0.25	0.27	0.28	0.15	0.31	0.05	0.37	0.31	0.28	0.59	1.00				
PP20b	0.03	0.23	0.26	0.27	0.38	0.21	0.15	0.23	0.24	-0.03	0.41	0.53	0.18	0.25	0.30	0.28	0.26	0.24	0.19	0.26	0.11	0.46	0.41	0.53	0.65	0.84	1.00			
PP20c	0.08	0.25	0.37	0.37	0.50	0.23	0.37	0.13	0.37	0.33	0.38	0.51	-0.03	0.38	0.27	0.23	0.20	0.22	0.16	0.27	0.18	0.60	0.38	0.36	0.58	0.85	0.94	1.00		
PP21	0.38	0.41	0.56	0.42	0.55	0.45	0.41	0.36	0.47	0.30	0.36	1.00	0.38	0.63	0.45	0.47	0.46	0.40	0.34	0.35	0.24	0.48	0.36	0.50	0.72	0.31	0.42	0.51	1.00	
PP22b	0.44	0.22	0.52	0.37	0.24	0.38	0.50	0.28	0.39	0.26	-0.03	0.87	0.39	0.55	0.39	0.44	0.45	0.37	0.29	0.20	0.17	0.27	-0.03	0.31	0.69	0.33	0.45	0.45	0.69	

Apéndice V

TABLAS Y GRÁFICOS

Tabla V.1.

Índices de confiabilidad y de dificultad (clásico y Rasch) de los ítems

Ítem	Correlación biserial		Alfa de Cronbach		Índice de dificultad			Criterio de clasificación	
	Incluido	Excluido	Total	Excluido	Clásico D	Rasch b	p	juez	Folstein
P11b	0.27	0.25		0.85	0.99	-3.59	0.76	Severo	Severo
P15	0.17	0.15		0.85	0.99	-3.45	0.00	Severo	Severo
P16	0.10	0.08		0.85	0.99	-3.34	0.31	Severo	Severo
P11a	0.18	0.16		0.85	0.99	-3.34	0.35	Severo	Severo
P11c	0.19	0.16		0.85	0.97	-2.79	0.30	Severo	Severo
P10	0.24	0.21		0.85	0.97	-2.70	0.57	Normal	Severo
P17	0.32	0.28		0.85	0.97	-2.58	0.18	Severo	Severo
P3	0.36	0.32		0.85	0.96	-2.45	0.37	Leve	Severo
P7	0.39	0.35		0.84	0.94	-2.17	0.58	Leve	Severo
P1	0.37	0.32	0.85	0.85	0.93	-2.08	0.66	Normal	Severo
P20b	0.36	0.31		0.85	0.93	-2.06	0.58	Leve	Severo
P8	0.34	0.28		0.85	0.92	-2.02	0.70	Leve	Severo
P20c	0.40	0.34		0.84	0.92	-1.97	0.55	Normal	Severo
P18	0.46	0.41		0.84	0.91	-1.92	0.00	Leve	Severo
P4	0.48	0.43		0.84	0.91	-1.85	0.07	Leve	Severo
P6	0.46	0.40		0.84	0.90	-1.84	0.33	Leve	Severo
P20a	0.34	0.27		0.85	0.89	-1.72	0.29	Severo	Severo
P5	0.42	0.35		0.84	0.87	-1.59	0.61	Normal	Leve
P9	0.50	0.43		0.84	0.85	-1.43	0.79	Normal	Leve
P2	0.46	0.39		0.84	0.79	-1.12	0.76	Normal	Leve
P12a	0.63	0.56		0.84	0.78	-1.06	0.00	Leve	Leve
P14a	0.50	0.42		0.84	0.74	-0.91	0.58	Leve	Leve
P21	0.58	0.51		0.84	0.70	-0.76	0.15	Leve	Leve
P14b	0.46	0.36		0.85	0.61	-0.42	0.00	Normal	Normal
P14c	0.39	0.29		0.85	0.52	-0.09	0.00	Normal	Normal
P22b	0.50	0.41		0.84	0.47	0.07	0.00	Leve	Normal
P12b	0.67	0.60		0.83	0.41	0.28	0.00	Leve	Normal
P12c	0.65	0.59		0.84	0.33	0.61	0.00	Normal	Normal
P12d	0.63	0.56		0.84	0.27	0.83	0.00	Normal	Normal
P12e	0.57	0.50		0.84	0.22	1.07	0.00	Normal	Normal

Tabla V.2.

Características demográficas de los adultos mayores de Lima y Callao

Características demográficas	Frecuencia	Porcentaje
Nivel de instrucción		
Sin nivel	3	0.55
Primaria	266	48.90
Secundaria	157	28.86
Superior no universitario	45	8.27
Superior universitario	73	13.42
Años de estudio		
8 años o menos	285	52.39
más de 8 años	259	47.61
Sexo		
Masculino	285	52.39
Femenino	259	47.61
Edad		
60 - 64	172	31.62
65 - 69	159	29.23
70 - 74	111	20.40
75 - 79	61	11.21
80 a más	41	7.54

Tabla V.3

Clasificación de los ítems

Criterio de juez corregido	Criterio punto de corte de Folstein en ítems ordenados			Total
	Severo	Leve	Normal	
Severo	23,3%	0,0%	0	23,3%
	7	0	0	7
Leve	23,3%	10,0%	6,7%	40,0%
	7	3	2	12
Normal	10,0%	10,0%	16,7%	36,7%
	3	3	5	11
Total	56,7%	20,0%	23,3%	100,0%
	17	6	7	30

Figura V.1

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P15 "¿Qué es esto? (reloj de pulsera)"

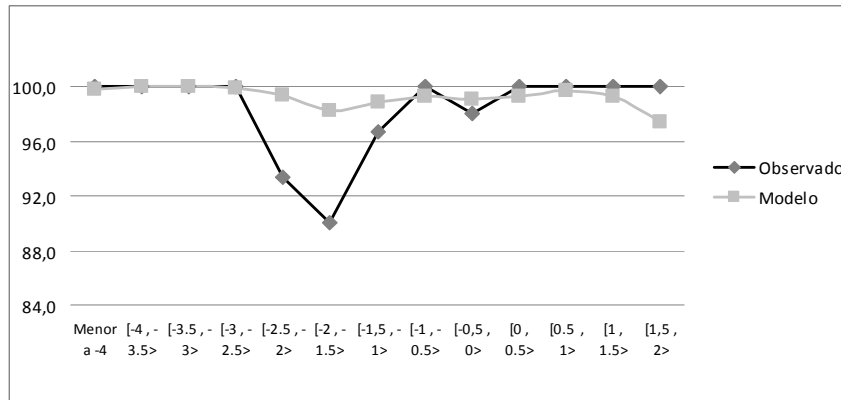


Figura V.2

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P18 "Lea y haga lo que dice el anexo N°1: Cierre los ojos"

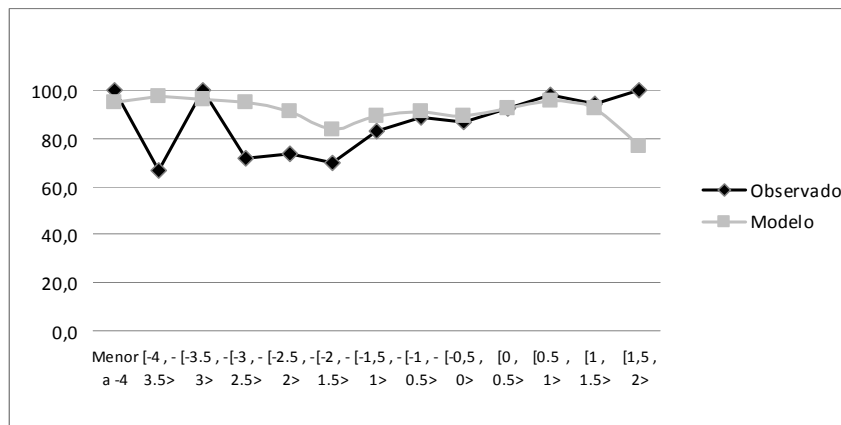


Figura V.3

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P12a “Vamos a restar de 7 en 7 a partir de 100 (93)”

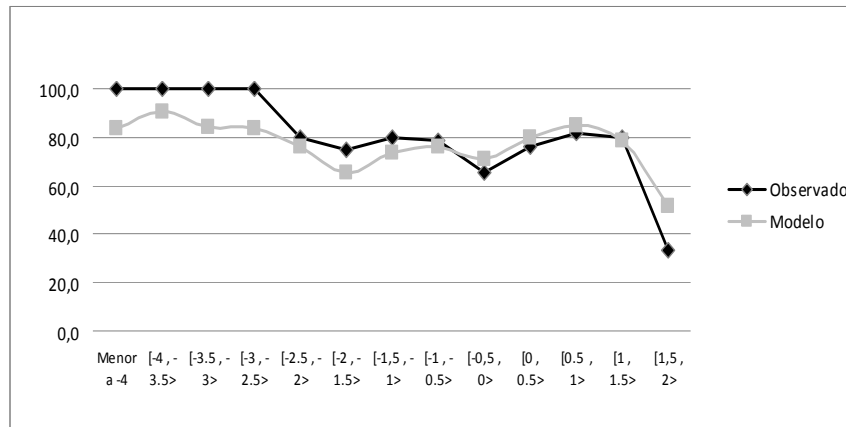


Figura V.4

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P14b “Ahora ¿Cuáles eran los 3 objetos que le pedí que recordara (árbol)?”

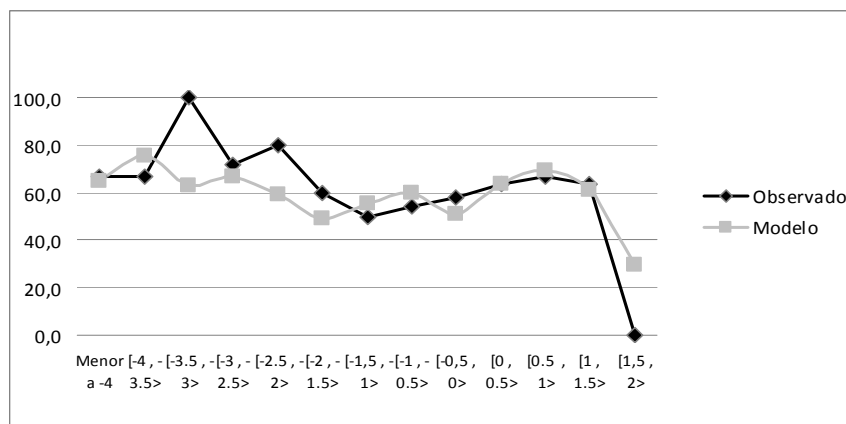


Figura V.5

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P14c “Ahora ¿Cuáles eran los 3 objetos que le pedí que recordara (avión)?”

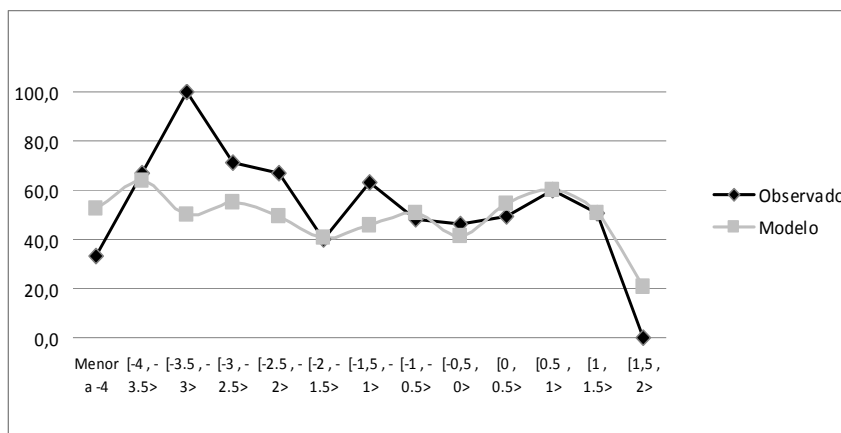


Figura V.6

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P22b “Aquí hay dos dibujos. Por favor copie los dibujos en el mismo papel (pentágonos)”.

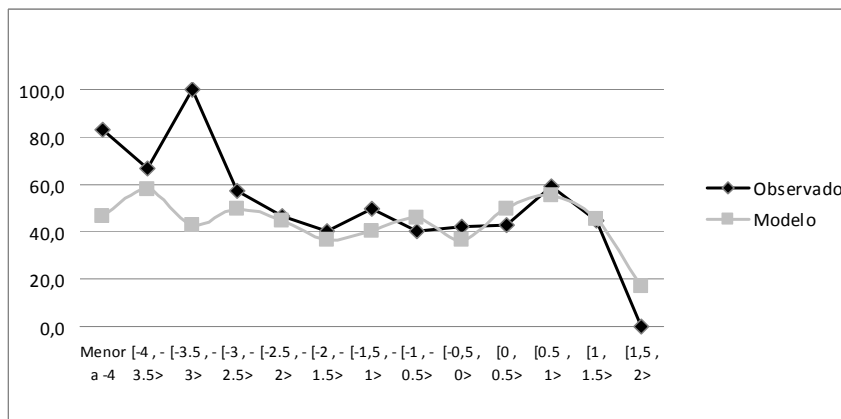


Figura V.7

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P12b “Vamos a restar de 7 en 7 a partir de 100 (86)”

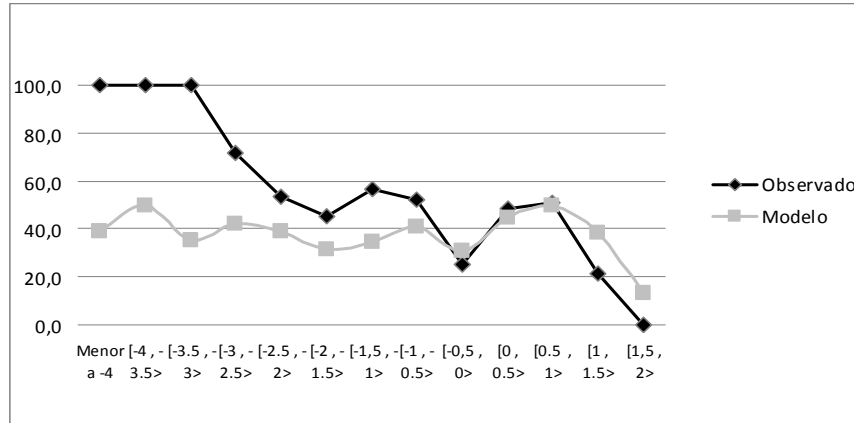


Figura V.8

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P12c “Vamos a restar de 7 en 7 a partir de 100 (79)”

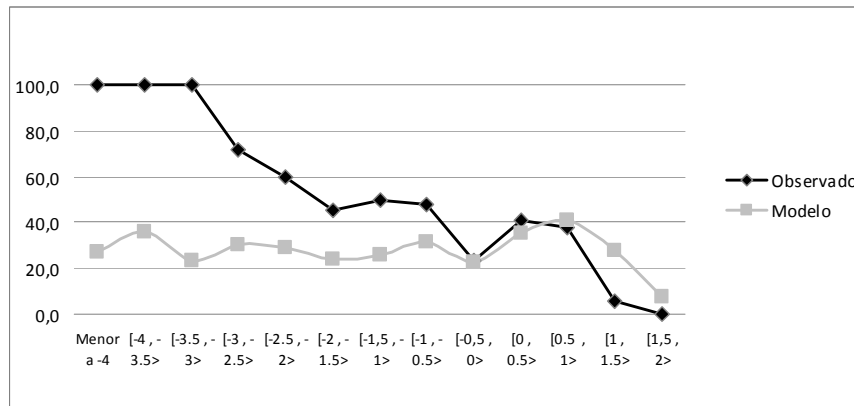


Figura V.9

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P12d “Vamos a restar de 7 en 7 a partir de 100 (72)”

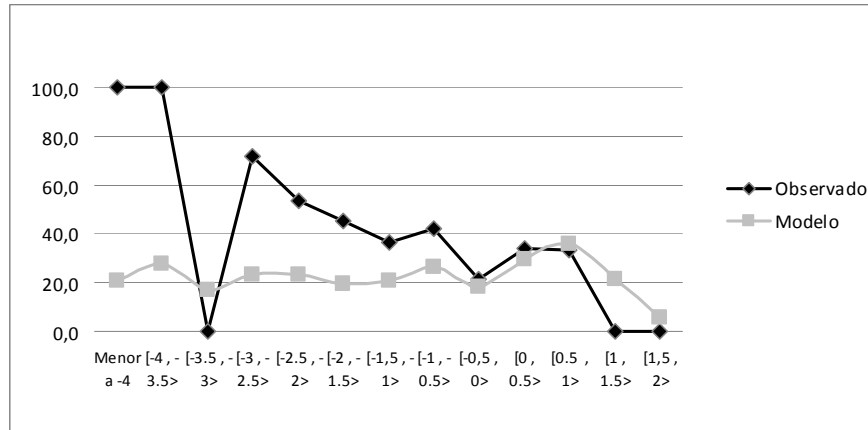


Figura V.10

Porcentaje de acierto empírica y probabilidad de acierto bajo el modelo Rasch para el ítem P12e: “Vamos a restar de 7 en 7 a partir de 100 (65)”

