

UNIVERSIDAD DE SANTIAGO DE CHILE
FACULTAD DE CIENCIA
Departamento de matemática y ciencia de la computación



**Aplicación de un Modelo de Clasificación Diagnóstica en los Atributos del
Test de Hipótesis**

Álvaro Marco Figueroa López

Profesora Guía: Rosa Montaña Espinoza

**Trabajo de graduación presentado a la
Facultad de Ciencia en cumplimiento de los
requisitos exigidos para optar al Título de
Magister en Educación Matemática**

**Santiago – Chile
2016**

© Álvaro Marco Figueroa López

Se autoriza la reproducción parcial o total de esta obra sólo con fines académicos y no comerciales, por cualquier forma, medio o procedimiento, siempre y cuando se incluya la cita bibliográfica del documento.

Hoja de Calificaciones

“Aplicación de un Modelo de Clasificación Diagnóstica en los Atributos del Test de Hipótesis

Álvaro Marco Figueroa López

Este trabajo de graduación fue elaborado bajo la supervisión de la profesora guía Srta. Rosa Montaña Espinoza del Departamento de Matemática y Ciencia de la Computación y ha sido aprobado por los miembros de la Comisión Calificadora: Srta. Daniela Soto Soto y el Sr. Patricio Montero Lagos.

Rosa Montaña Espinoza.
Profesor Guía

Daniela Soto Soto
Profesor Informante

Pedro Marín Álvarez
Director del Departamento

Patricio Montero Lagos
Profesor Informante



Universidad de Santiago de Chile.

Facultad de Ciencia.

Departamento de Matemática y Ciencia de la Computación.

Santiago, abril 2017

Estimado Sr. Director:

Me permito informar a usted, sobre el trabajo de tesis para optar al grado de Magister en Educación Matemática presentado por el Sr. Álvaro Marco Figueroa López titulado "Aplicación de un Modelo de Clasificación Diagnóstica en los Atributos del Test de Hipótesis".

De acuerdo a lo presentado puedo indicar que este trabajo es un aporte al área de la Educación, que representa una aplicación de los modelos diagnósticos en las habilidades involucradas para resolver problemas de test de hipótesis.

En términos específicos, informo respecto a:

- a) El trabajo presenta un desarrollo de los modelos de la teoría de respuesta al ítem y la formulación de un modelo diagnóstico.
- b) La aplicación se sustenta en la creación y aplicación de un test respecto de los test de hipótesis en profesores y estudiantes de pedagogía.
- c) Los resultados se sustentan en la aplicación de los modelos en software de alta complejidad.
- d) Lo realizado demuestra conocimientos y experticias de gran utilidad para el futuro desarrollo profesional de la estudiante.

Considerando lo anteriormente expuesto califico el trabajo con nota siete comas cero (7.0)

Le saluda muy atentamente,

PhD. Rosa Montaña Espinoza

Profesora Guía



INFORME DE SEMINARIO DE TESIS

En relación al trabajo de titulación del Sr. Álvaro Marco Figueroa López, del programa de Magister en Educación Matemática, titulado: "Aplicación de un Modelo de Clasificación Diagnóstica en los Atributos del Test de Hipótesis", informo que el trabajo realizado es un real aporte a la educación matemática, tanto por el planteamiento del problema como por la metodología utilizada, los recursos y la innovación que propone, la cual es plenamente transferible a una práctica docente, los felicito.

En términos específicos, informo respecto a:

1. Relevancia del tema:

La relevancia del tema se expresa en el capítulo I, donde el planteamiento es argumentado a partir de la investigación científica (estudios de Batanero, Vallecillos, entre otros autores). Es interesante leer sobre las problemáticas que ocasiona la enseñanza y el aprendizaje de los test de hipótesis.

2. Fundamento Teórico:

La fundamentación teórica es clara y concisa.

3. Relación entre Objetivos y Logros:

Con respecto a la postulación de los objetivos (página 3) considero que la redacción del primer objetivo responde más a un plan de trabajo.

4. Estructura del Trabajo

Aconsejo no poner la introducción dentro del capítulo I.

5. Relación con los Objetivos de la Carrera

Considero que la tesis es un aporte en el sentido metodológico y del marco teórico que se plantea. Sin embargo, me pregunto acerca de cómo la tesis responde a la problemática de la enseñanza y el aprendizaje de la noción de Test de hipótesis. Si bien, podemos pensar en las habilidades que se deben desarrollar (objetivos II y III) y cómo el análisis que se desarrolla en la tesis muestra los niveles de esas habilidades en esa población, además de las correlaciones que existen entre ellas. Me pregunto ¿Qué medidas se deberían considerar para la enseñanza y el aprendizaje de esta noción (Test de hipótesis)?

6. Resultados e Implementación

Muy Bueno.

7. Conclusiones

Considero que las conclusiones pudieran ser más robustas. Por ejemplo: qué se puede concluir con respecto a las problemáticas que se plantean en el capítulo I (del test de hipótesis). Cuáles son las limitaciones del estudio. Y las prospectivas de la investigación.

8. Bibliografía

Muy Bueno

Muy buen trabajo, felicitaciones

Por lo anteriormente expuesto, califico el trabajo con nota 7,0.

A handwritten signature in black ink, appearing to read 'Daniela Soto Soto', with a large, stylized flourish at the end.

Daniela Soto Soto

Profesora Informante



INFORME DE TRABAJO DE GRADUACIÓN

Estimado Sr. Director:

En relación al trabajo de graduación del Sr. Alvaro Figueroa López titulado: "Aplicación de un Modelo de Clasificación Diagnóstica en los Atributos del Test de Hipótesis" me es grato informar lo siguiente:

- a) El trabajo aborda una temática fundamental para avanzar en la validación de conocimientos científicos, en particular en educación matemática.
- b) Sus objetivos contribuyen a resignificar aspectos teóricos e instrumentales para el uso de test de hipótesis considerando errores frecuentes.
- c) El marco teórico presentado es un muy buen referente como recurso de aprendizaje.
- d) Los procedimientos utilizados satisfacen los criterios de una apropiada parsimonia metodológica.
- e) Los productos obtenidos en su trabajo logran ampliamente los objetivos propuestos.
- f) Las conclusiones son pertinentes, vinculadas con los objetivos del trabajo con una posible subvaloración a las proyecciones del trabajo
- g) Aun cuando el trabajo presenta algunas observaciones menores de estilo, que no desmerecen el destacado trabajo realizado

En atención a lo anterior, califico el trabajo con nota seis como ocho (6.8).

Le saluda muy atentamente,

Patricio Montero Lagos, Ph.D
Profesor informante

Santiago, abril 2017

Agradecimientos

Rendirse nunca fue una opción.

Han sido dos años difíciles, sin embargo, cualquiera que haya gozado la vida sabe lo difícil que se pone a veces, lo importante es que así es más entretenido, así la vida te hace crecer, así sabes el valor de la gente con la que estas y poco a poco aprendes a gozar los momentos que la vida te regala.

El camino llevado hasta hoy comenzó hace más de 8 años, con un cachorro que no sabía el mundo que tenía adelante, en un lugar el cual me enseñó a volar. Pero este camino no ha sido sólo mío, hay mucha gente que ha estado ahí en los buenos y malos momentos, personas que han aguantado mi lejanía y poco tiempo. Almas cálidas que siempre me entregaron su cariño, ánimos y buenas vibras.

Entre ellas está mi madre (Angelina L.) que me ha apoyado y aguantado toda mi lejanía y poco tiempo. Gracias chiquitita por estar ahí y mantener la fe aun cuando no sabías que pasaba. También quisiera agradecer la paciencia de mi familia, en especial a mi abuela la mama Elba (Elba G.), mi hermano (Isaak F.), mi sobrina (Martina F.) y mi viejo (Dagoberto F.) por estar ahí siempre cuando mire hacia el lado. Gracias por estar presente y darme el espacio que necesité por todo este tiempo, gracias por entender mi lejanía y esperarme con los brazos abiertos.

En el camino me perdí, perdí el norte y quise rendirme. Pero cuando más lejos me encontré, cuando las fuerzas se desvanecían, el mundo hizo aparecer a una persona especial, alguien que me acepto, acompañó y cobijó cuantas veces necesite, sin ella no podría estar escribiendo estas líneas, sin ella no sería lo que hoy soy. Aracely, mi compañera, mi polola, gracias. Gracias por el tiempo, por darme tus fuerzas cuando flaqueaban las mías. Gracias por quedarte y abrazarme esas noches de invierno, por amarme en primavera y seguir con caricias en verano y en este otoño.

Además, hay seres fieles, felices y llenos de amor que siempre me han acompañado. Todas las mañanas escucho sus voces, sus ladridos dando ánimos y pidiendo atención; gracias Mojón (Dinki) y Maqui por aparecer en mi vida, gracias por darme ese amor tan puro y sincero que solo los perros saben dar, gracias por seguir aquí creyendo en mí.

También, quisiera agradecer a mis amigos que inquebrantablemente han seguido ahí, dándome su apoyo, esperando una cerveza helada o un tecito lleno de historias y risas como siempre ha sido, gracias Don Wooo (Cristóbal P.) por la amistad, por estar ahí y siempre tener buenas vibras, siempre serás mi mejor amigo. También estas tú, Paulina R. que los años y el poco tiempo sólo nutren nuestra amistad, asimismo pienso en ti Vito (Victoria) que siempre nos tendremos el uno al otro para reírnos de las desgracias, de los amores y de la vida completa.

Además, existen dos personas que admiro profundamente, una de ellas es mi profesora guía, la Srta. Rosa M., usted esa una persona increíble, siempre llena de vida, siempre solucionando la vida, gracias por aceptarme por segunda vez, gracias por ayudarme en este camino y ojalá

viaje mucho, ojalá conozca y recorra cuanto su alma quiera, que sea libre, feliz y nunca cambie, es un ejemplo para todos sus alumnos.

La segunda persona que admiro profundamente es Hank A., un amigo, una persona que desde primer año de universidad nos ha ayudado y enseñado. ¡Eres grande weon, eres el mejor de todos nosotros, una persona increíble y sin ti, no hubiese llegado hasta aquí, Gracias Shoro!

Finalmente, sé que queda mucha gente que falta agradecer, gente que estuvo en algún momento, algunos ya se fueron y otros continúan, pero a todos y cada uno de ustedes quiero agradecer por el simple hecho de aparecer en mi vida, gracias por los momentos, por las risas, por los abrazos, las buenas vibras. Gracias por lo que fue y por lo que aún queda por vivir.

Álvaro (The Lost) Figueroa.

Al hombre se le puede arrebatar todo salvo una cosa, la última de las libertades humanas, la elección de la actitud personal que puede adoptar frente al destino para decidir su propio camino. - Viktor Frankl

Resumen

El Test de Hipótesis es una poderosa herramienta dentro de la Inferencia Estadística y la Ciencia por su utilidad en la toma de decisiones y la confrontación de nuevas teorías en base a una muestra. Sin embargo, su uso no siempre es el adecuado debido a que requiere de varias habilidades para su correcto desarrollo. Estas habilidades, llamadas latentes, no son observables directamente y están definidas a través de otros procesos de raciocinio que si son observables. Dichos procesos, para esta investigación, se dividen en tres habilidades latentes que entenderemos por la capacidad de formular hipótesis, determinar la región de rechazo y las habilidades relacionadas a las probabilidades del error tipo I y II en el Test de Hipótesis.

En el proceso de enseñanza de estas habilidades, se necesita un modelo de medición donde el docente pueda vislumbrar los dominios y los no dominios que tiene cada uno de sus estudiantes. En este contexto los Modelos de Clasificación Diagnóstica y los Modelos Logísticos constituyen una poderosa herramienta en la entrega de información del cuestionario y también clasificando a cada estudiante según su nivel de habilidad y dominio de las habilidades latentes.

Particularmente, esta investigación analiza los dominios de estudiantes y docentes de matemática chilenos en el desarrollo del Test de Hipótesis. Para ello se utiliza el modelo logístico de tres parámetros (3-PL) y el modelo de clasificación diagnóstica Loglineal (LCDM), donde sus resultados entregan la dificultad, discriminación y azar de cada ítem. Además del perfil de dominios, la habilidad y probabilidad de acierto de cada encuestado en relación a los ítems, y así determinar las características de cada grupo de individuos y compararlos según sus habilidades, dominios y preguntas correctas.

Palabras Claves: Test de Hipótesis, Modelo de Clasificación Diagnóstica Loglineal, Modelo Logístico de Tres Parámetros, Habilidades Latentes.

Abstract

The Hypothesis Test is a powerful tool in Statistical Inference and Science because of its usefulness to make decisions and the confrontation of new theories based on a sample. However, its use is not always the most appropriate one due to the different attributes that are required to develop it correctly. These attributes, which are called latent, are not directly observable and are defined through reasoning processes which indeed can be noticeable. For this research, these processes are divided in three latent attributes which are defined as the capacity for formulating hypothesis, establishing the rejection region and skills related to probabilities of errors I and II in to Hypothesis Test.

In the teaching process of these attributes, it is needed a model of measurement where the teacher can notice the mastering of latent attributes of each student. In this context, the Diagnostic Classification Models and Logistic Models are great instruments to acquire information from the questionnaire and to classify each student according to his/her skill level and the mastering of these latent attributes.

Particularly, this study analyzes the Chilean Math students and teachers' mastering in the development of the Hypothesis Test. In order to obtain this information, The Loglinear Cognitive Diagnosis Model and Three Parameter Logistic Model were used. Their analysis show the difficulty, discrimination and guessing of each item. In addition, the results show the mastering profile, the skill level and the probability of the right answer in each item of each respondent. With those results, the respondents' characteristics can be determined and compared with their skill level, mastering and right answers.

Key Words: Hypothesis Test, The Loglinear Cognitive Diagnosis Model, Three Parameter Logistic Model, Attribute Latent.

Tabla de Contenidos

Índice de Tablas	xii
Índice de Ilustraciones	xiii
Capítulo I Introducción	1
1.1 Objetivo General	3
1.1.1 Objetivos Específicos	3
1.2 Antecedentes	4
1.2.1 Inferencia Estadística	4
1.2.1.1 Test de Hipótesis	6
1.3 Medición de Habilidades a través de Instrumentos	10
Capítulo II Marco Teórico	11
2.1 Teoría de Respuesta al Ítem	11
2.1.1 Modelos Logísticos	12
2.2 Modelo de Clasificación Diagnóstica	13
2.2.1 Notación	15
2.2.2 Modelos de Clasificación Diagnóstica Compensatorio	17
2.2.3 Modelos de Clasificación Diagnóstica Conjuntista	18
2.2.3.1 Modelo DINA	18
2.2.3.2 Modelo de Clasificación Diagnóstica Loglineal	20
Capítulo III Metodología de Trabajo	26
3.1 Muestra	26
3.2 Variables	26
3.2.1 Variables Latentes	27
3.3 Hipótesis	28
3.4 Instrumento	29
3.5 Matriz Q	34
Capítulo IV Resultados	35
4.1 Descripción de la Muestra	35
4.2 Análisis Descriptivo	36

4.2.1 Confiabilidad del Instrumento	37
4.2.2 Respuesta según Ítem	37
4.2.3 Encuestados según Cantidad de Preguntas Correctas	38
4.2.4 Comparación de Preguntas Correctas según Género.....	39
4.2.5 Comparación de Preguntas Correctas según Situación Académica	40
4.2.6 Comparación Ítem Correcto según Situación Académica	41
4.3 Análisis según Modelo Logístico de Tres Parámetros.....	42
4.3.1 Función Informativa Modelo 3-PL	44
4.4 Análisis según Modelo de Clasificación Diagnóstica Loglineal	45
4.4.1 Recuento de Clases	46
4.4.2 Parámetros Estimados.....	47
4.4.3 Modelo Logit, Probabilidad y Análisis según Modelo 3-PL.....	48
4.4.4 Perfil Asociado a la Probabilidad de Acierto al Ítem	51
4.4.5 Comparación de Probabilidad según Perfil.....	55
4.4.6 Probabilidad por Atributo.....	56
Capítulo V Conclusiones.....	57
5.1 Conclusiones Respecto a los Resultados.....	58
Bibliografía	60
Anexos	66
Anexo 1: Código SAS	66
Anexo 2: Código LCDM Mplus.....	74
Anexo 3: Función de Probabilidad Según Modelo 3-PL.....	97
Anexo 4: Clase Latente por Patrón de Respuesta	100

Índice de Tablas

Tabla 1: Matriz Q.....	15
Tabla 2: Ejemplo Matriz-Q LCDM	22
Tabla 3: Ejemplo Logit LCDM	22
Tabla 4: Estimador y Función Auxiliar para 4 Parámetros	24
Tabla 5: Q-Matriz	34
Tabla 6: Posibles Patrones de Clase	34
Tabla 7: Confiabilidad según Alpha de Cronbach	37
Tabla 8: Respuestas según Ítem	37
Tabla 9: Análisis según Modelo 3-PL	42
Tabla 10: Patrones de Respuesta y Habilidad θ	43
Tabla 11: Matriz Q para SAS	45
Tabla 12: Clases Latentes de los Encuestados.....	46
Tabla 13: Parámetros Estimados.....	47
Tabla 14: Logit y Función de Probabilidad Ítem 1	48
Tabla 15: Logit y Función de Probabilidad Ítem 2	48
Tabla 16: Logit y Función de Probabilidad Ítem 3	48
Tabla 17: Logit y Función de Probabilidad Ítem 4	49
Tabla 18: Logit y Función de Probabilidad Ítem 5	49
Tabla 19: Logit y Función de Probabilidad Ítem 6	49
Tabla 20: Logit y Función de Probabilidad Ítem 7	50
Tabla 21: Probabilidad de Acierto por Atributo	56

Índice de Ilustraciones

Ilustración 1: Curva Característica Modelo 3-PL	12
Ilustración 2: Función Respuesta Modelo DINO	17
Ilustración 3: Función Respuesta Modelo DINA	19
Ilustración 4: Curva Característica Modelo DINA	19
Ilustración 5: Curva Característica LCDM.....	21
Ilustración 6: Función de Respuesta LCDM	23
Ilustración 7: Ejemplo estimador para 4 parámetros	24
Ilustración 8: Grafico de Caja para Edad Encuestados	35
Ilustración 9: Condición Académica de los Encuestados	36
Ilustración 10: Encuestados según Cantidad de Preguntas Correctas	38
Ilustración 11: Comparación según Preguntas Correctas y Frecuencia Relativa del Género	39
Ilustración 12: Preguntas Correctas según Frecuencia Relativa de Situación Académica	40
Ilustración 13: Porcentaje de Acierto según Situación Académica.....	41
Ilustración 14: Análisis según Modelo 3-PL.....	42
Ilustración 15: Función Informativa Modelo 3-PL	44
Ilustración 16: Clases Latentes de los Encuestados	46
Ilustración 17: Perfil [000].....	51
Ilustración 18: Perfil [100].....	51
Ilustración 19: Perfil [010].....	52
Ilustración 20: Perfil [001].....	52
Ilustración 21: Perfil [110].....	53
Ilustración 22: Perfil [101].....	53
Ilustración 23: Perfil [011].....	54
Ilustración 24: Perfil [111].....	54
Ilustración 25: Probabilidad de Acierto al Ítem según Perfil	55

Capítulo I Introducción

En el mundo de hoy la estadística está presente por doquier, los medios de comunicación, los recursos informáticos disponibles y la formación básica general de los ciudadanos exigen conocimientos estadísticos que les permitan comprender el mundo que les rodea.

Hacking en 1990 afirma que uno de los descubrimientos decisivos del siglo XX fue la constatación de que el mundo no es determinista, donde la importancia de la Inferencia estadística es innegable, tanto por el uso que se hace de ella en la investigación experimental como su utilidad en la toma de decisiones en diversos ámbitos de la actividad humana, en consecuencia, que Rao y Szekely en el 2000 (Barreto, 2012) afirman que la estadística inferencial tiene un gran porvenir en el siglo XXI.

En la actualidad su aplicación se ve facilitada por la existencia de ordenadores y colecciones de programas informáticos de cálculo, accesibles y de fácil manejo incluso para los no especialistas. Por este motivo se incluyen contenidos relacionados en los programas de formación de gran parte de las especialidades universitarias. Sin embargo, esta formación y facilidad de aplicación no siempre implica un uso correcto y eficaz de las nociones estadísticas. Este hecho ha propiciado, sin duda, la aparición de algunos problemas derivados del uso de las técnicas estadísticas por incomprensión o abuso de las mismas (Batanero et al., 1992; Vallecillos & Batanero, 1997), pero también ha permitido, en el ámbito educativo, un mejor acceso al aprendizaje de algunos conceptos cuya comprensión es difícil, si no imposible, sin ellos (Vallecillos, 1995).

En este contexto, la importancia de la Inferencia Estadística es innegable tanto en la toma de decisiones en los diversos ámbitos de la actividad humana, como en la creación de nuevos conocimientos y específicamente, la estadística inferencial es de suma importancia por el rol que cumple en investigaciones experimentales. En particular, en el área de las ciencias de la conducta como es la Pedagogía y la Psicología, en estos campos su utilidad es innegable porque basan sus investigaciones en datos recogidos en muestras de poblaciones, a las que necesita extender sus conclusiones.

Una herramienta fundamental de la estadística inferencial es el contraste de hipótesis tanto por su apoyo en la confirmación de los modelos empleados, como en la demostración de la existencia de efectos de las variables de interés en la investigación (Batanero, 2000). Es por esto que en los últimos años observamos una tendencia creciente a incluir contenidos de inferencia estadística en el currículo de matemáticas en la escuela secundaria y universitaria (Batanero, 2011).

El uso correcto de esta herramienta (Batanero et al., 2012; Vera et al., 2013) requiere el dominio de varias habilidades en el manejo de objetos matemáticos y estadísticos, también del correcto uso del lenguaje, conceptos, propiedades, procedimientos y argumentos. Debido a esta variedad

de conceptos que deben ser comprendidos y relacionados para realizar las pruebas de hipótesis, es que los estudiantes o incluso investigadores se confunden entre sí y lo mal utilizan generando errores conceptuales que han sido discutidos por más de 30 años.

Esta es una de las razones por la cual desde finales de los ochenta existe la constante idea de poder diagnosticar los problemas o errores (Borassi, 1987) con que se encuentran los estudiantes en el aprendizaje del tema, como así también diagnosticar los conceptos que los estudiantes dominan. En consecuencia, desde la década de los 90 surgió un gran interés por parte de la psicometría de estudiar modelos estadísticos con variables latentes como una alternativa para analizar las respuestas de los estudiantes.

Las variables latentes no son observables directamente, sino que se miden a partir de otras variables que sí son observables, por ejemplo, una habilidad latente es encontrar máximos y mínimos de una función real, la cual está definida por reconocer la función y su derivada, simplificar los términos algebraicos extraídos del proceso, calcular los puntos críticos, evaluar en la función original y reconocer los mínimos y máximos relativos y absolutos.

Los modelos que utilizan variables latentes son empleados en diversas disciplinas, tales como la psicología, economía, pedagogía, medicina, física, inteligencia artificial, bioinformática, econometría, gestión y ciencias sociales entre otras. En la actualidad, existen variados modelos que utilizan variables latentes, un conjunto de estos se llaman Modelos de Clasificación Diagnóstica (DCMs) cuyo mayor exponente es Rupp y Templin. Estos modelos se encuentran dentro de un marco más general y algunos de ellos son los Modelos de Clases Latentes sin Restricciones (Haagenars & McCutcheon, 2002), el Análisis Factorial (AF) (McDonald, 1967), la Teoría de Respuesta al Ítem (TRI) y los Modelos Logísticos de Uno, Dos o Tres Parámetros (Embretson & Reise, 2000).

En los últimos años, los Modelos de Clasificación Diagnóstica tienen un desarrollo importante en las mediciones de pruebas educativas y psicológicas ya que, en lugar de una puntuación global de las pruebas, un test diagnóstico proporciona a cada participante un perfil que detalle los conceptos y habilidades o atributos que el individuo domina.

Uniendo lo expuesto anteriormente, se presenta esta investigación la cual utiliza el modelo de clasificación diagnóstica Loglineal y modelo logístico de tres parámetros para clasificar a estudiantes y docentes en sus habilidades latentes en el uso de test de hipótesis. Específicamente, este informe se compone de cinco capítulos donde se informan los objetivos, el marco teórico que sustenta la investigación, la metodología de trabajo, los resultados obtenidos y finalmente, las conclusiones relacionadas a la muestra.

1.1 Objetivo General

Analizar habilidades cognitivas en el uso de problemas de test de hipótesis en estudiantes de pedagogía y docentes del área de educación matemática utilizando modelo de Medición Diagnóstica.

1.1.1 Objetivos Específicos

Los objetivos específicos tienen como propósito, encaminar la investigación hacia el Objetivo General, en este sentido se pretende:

- i) Revisar la literatura acerca de los modelos de clasificación y medición diagnóstica más utilizados.
- ii) Detectar patrones de raciocinio que involucren habilidades y destrezas inmersas en el uso del test de hipótesis.
- iii) Caracterizar los atributos latentes necesarios para resolver ejercicios que involucren el uso de test de hipótesis.
- iv) Determinar la correlación que existe entre las habilidades medidas.
- v) Generar Modelo Loglineal de Clasificación Diagnóstica para encuestados según sus atributos latentes en la resolución de ejercicios que involucren el Test de Hipótesis.

1.2 Antecedentes

1.2.1 Inferencia Estadística

La estadística ha jugado un papel primordial en el desarrollo de la sociedad moderna, al proporcionar herramientas metodológicas generales para analizar la variabilidad, determinar relaciones entre variables, diseñar en forma óptima estudios y experimentos y mejorar las predicciones y toma de decisiones en situaciones de incertidumbre. Ello llevó a Argentina, Chile y otros países a iniciar la formación profesional de estadísticos hace más de 50 años (Haedo, 2001).

Junto con lo anterior, Holmes en el 2002 señala que la enseñanza de la estadística y probabilidad fue ya introducida en 1961 en el currículo de Inglaterra en forma opcional para los estudiantes de 16 a 19 años que querían especializarse en matemáticas, con el fin de mostrar las aplicaciones de las matemáticas a una amplia variedad de materias ya que:

- La estadística es una parte de la educación general deseable para los futuros ciudadanos adultos, quienes precisan adquirir la capacidad de lectura e interpretación de tablas y gráficos estadísticos que con frecuencia aparecen en los medios informativos.
- Es útil para la vida posterior, ya que en muchas profesiones se precisan unos conocimientos básicos del tema.
- Su estudio ayuda al desarrollo personal, fomentando un razonamiento crítico, basado en la valoración de la evidencia objetiva.
- Ayuda a comprender los restantes temas del currículo, tanto de la educación obligatoria como posterior, donde con frecuencia aparecen gráficos, resúmenes o conceptos estadísticos.

Más recientemente, la enseñanza de la estadística se incorpora en forma generalizada a la escuela, institutos y carreras universitarias, ya que se reconoce su carácter instrumental para otras disciplinas, y especialmente se reconoce el valor del desarrollo del razonamiento estadístico en una sociedad caracterizada por la disponibilidad de información y la necesidad de toma de decisiones en ambiente de incertidumbre; también es importante reconocer el papel primordial de las Sociedades e Institutos de Estadística, profesores e investigadores para lograr una cultura estadística para todos.

En una sociedad informada, los métodos estadísticos proporcionan una valiosa y enorme ayuda a las interpretaciones de los resultados de experimentos científicos y ayudan también en el diseño de experimentos que produzcan resultados interpretables. En un sentido muy general, el propósito del análisis estadístico es organizar los datos, revelar su estructura y generalizar conclusiones desde una muestra hasta su población (Trosset, 2006).

Siguiendo lo anterior, la inferencia estadística en el rol de poder generalizar los datos de una muestra a una población juega un papel destacado y su necesidad e importancia han ido en aumento durante los últimos años en diversas ciencias humanas, entre otras la Psicología y Pedagogía, que basan sus investigaciones en datos recogidos en muestras de poblaciones, a las que necesita extender sus conclusiones.

Por lo tanto, la Estadística inferencial se puede definir como aquellos métodos que permiten hacer estimación de una característica de la población o de toma de decisiones respecto a una población, con base solo en los resultados obtenidos de una muestra (Barreto, 2012), es decir, no se requiere el registro de toda la población, sino solo a una parte de esta. Este método utiliza un proceso el cual esta intervenido por dos tipos de razonamientos: el pensamiento deductivo y el pensamiento inductivo.

Estos procesos además de los pensamientos nombrados requieren, por parte del investigador, un conocimiento en el cual intervienen los siguientes objetos (Vera, Díaz & Batanero, 2011):

- *Situaciones-problemas, de donde surge el objeto:* El concepto de hipótesis estadística surgirá de problemas de comparación de dos o más poblaciones, de estimación de parámetros o de toma de decisiones.
- *Lenguaje:* Términos, expresiones, notaciones, gráficos que se usan en el trabajo matemático, por ejemplo, los símbolos usados para denotar los parámetros μ , σ o los usados para hipótesis nula H_0 y alternativa H_1 .
- *Conceptos:* Población y muestra, estadístico y parámetro, región de rechazo y de aceptación.
- *Propiedades:* Por ejemplo, que las hipótesis nula y alternativa son complementarias o que las hipótesis se formulan en función del parámetro.
- *Procedimientos:* Como los requeridos para construir las regiones de rechazo y aceptación.
- *Argumentos:* Usados para justificar o explicar a otra persona las proposiciones y procedimientos.

Es importante recalcar que la Estadística inferencial no proporciona una certeza completa de sus resultados, sino que los mismos están sujetos a una probabilidad de error. Por ejemplo, una selección incorrecta de las unidades muestra puede acarrear la imposibilidad de inferir correctamente las características de la población.

En el área de las ciencias de la conducta es innegable la utilidad de la inferencia estadística, sin embargo, el uso e interpretaciones de esta en las publicaciones no son siempre adecuados, como se muestra en diversas revisiones (Morrison y Henkel, 1970; Menon, 1993; Thompson (1996); Harlow, Mulaik y Steiger, 1997; Ares, 1999; Borges, San Luis, Sánchez, y Cañadas, 2001; Díaz

y de la Fuente, 2004; Batanero y Díaz, 2006; Vera, Diaz, & Batanero, 2011; Batanero, Vera, & Díaz, 2012; Vera & Díaz, 2013). Es más, múltiples autores dedican su investigación al tema de la inferencia estadística y las dificultades que este tema suscita, ganando día a día gran impacto a nivel internacional (Haller & Krauss, 2002).

En este contexto Ward (2015) menciona que la inferencia es la gran idea de las estadísticas, pero es aquí donde la gente tiene un traspié, la mayoría de las personas pueden aceptar el uso de estadísticas descriptivas, entienden por qué se necesitan datos, pueden ver que la forma en que se toma una muestra afecta el resultado, es más, a menudo entienden la necesidad de grupos de control. La mayoría de los conceptos o ideas estadísticas son fácilmente explicables, pero la inferencia es una idea difícil y complicada.

Junto con esto, Harradine, Batanero y Rossman (2011) indican que la comprensión de la inferencia estadística requiere el aprendizaje de tres elementos interrelacionados: el proceso de razonamiento, los cálculos relacionados y los conceptos y vocabularios asociados. Es por esto que mientras la realización de los cálculos en los contrastes de hipótesis es hoy día muy sencilla, gracias al software estadístico, la enseñanza de los conceptos, lenguaje técnico y el razonamiento inferencial es mucho más compleja, lo que explica las muchas dificultades descritas en el uso de la inferencia (Brewer, 1986; Vallecillos, 1999).

En particular, el test de hipótesis es la principal herramienta de la inferencia estadística, por esto a continuación se detallará la importancia y los principales errores relacionados al uso del contraste de hipótesis.

1.2.1.1 Test de Hipótesis

Sea x una variable aleatoria cuya función de probabilidad tiene una distribución que depende de un parámetro; una prueba de hipótesis, contraste o test de hipótesis es un método sistemático que permite evaluar el verdadero valor de un parámetro a partir de una muestra y extrapolar los resultados a la población, dicho método requiere la confrontación de acontecimientos con evidencia real y decidir, en vista de esta evidencia, si dichos sucesos se pueden conservar como razonables o deben ser desechadas y de esta forma, cambiar la imagen de la realidad.

Según Pimienta (2002) uno de los objetivos de las investigaciones científicas es la comprobación de hipótesis, en este contexto, la función de un contraste de hipótesis es probar si alguna relación, la cual es expresa a través de variables medibles, existe o no. Además de esto, se considera a las hipótesis como un instrumento indispensable, por su utilidad como herramientas de trabajo en la teoría ya que son demostrables y representan un poderoso instrumento para el progreso del conocimiento debido a que ayudan a confirmar o negar una teoría en forma independiente de la opinión del investigador.

Una prueba de hipótesis consiste en determinar si un supuesto sobre alguna o algunas características de la población, está ampliamente respaldado por la información obtenida a través de datos muestrales. Se expresa una supuesta hipótesis sobre la población la cual involucra una afirmación del valor de un parámetro y su evaluación y confirmación se basa en la información obtenida de una muestra aleatoria.

El desarrollo del contraste de hipótesis requiere de conocimientos y habilidades específicas que el investigador debe dominar. Estas habilidades pueden ser entendidas como atributos latentes, los cuales son utilizados en la asigna las hipótesis, la diferenciación entre la hipótesis nula y la hipótesis alternativa; también en el análisis de la región de rechazo y el reconocimiento, cálculo e interpretación del p-valor, nivel de significancia, potencia del test y las probabilidades del error tipo I y II.

El mal uso o errores en estas habilidades se intentan explicar, desde la Psicología, por la existencia de heurísticas y sesgos en el razonamiento estocástico. Es importante recalcar que estos errores también se producen en investigadores y estudiantes universitarios, como muestran muchas investigaciones (Birnbbaum, 1982; Vera, Díaz, & Batanero, 2011; Batanero, Vera, & Díaz, 2012; Vera & Díaz, 2013). En España, la investigación didáctica en el tema fue iniciada por Vallecillos con su tesis doctoral (Vallecillos, 1994) y ésta ha sentado bases de una amplia investigación posterior (Díaz, 2007; Alvarado, 2007; Vera & Díaz, 2013).

En este contexto, la mayoría de las investigaciones llevadas a cabo se limitan a en la comprensión del concepto de nivel de significación. Sin embargo, como se describió anteriormente, son muchos los conceptos que se deben comprender para llevar a cabo con éxito un contraste de hipótesis. Vallecillos (1994) describe diversos errores relacionados con la inferencia estadística clasificándolos en varias categorías:

- i) *Parámetro y distribuciones muestrales*: Los estudiantes no dan cuenta de la variabilidad de la distribución muestral del estadístico, ni reconocen su relación con el parámetro poblacional.
- ii) *Confusión entre las hipótesis*: Los estudiantes confunden la hipótesis nula con la hipótesis alternativa. Definen como hipótesis nula aquella que se desea probar. Por otro lado, aunque la teoría estadística se ocupa del último nivel de hipótesis (de las hipótesis estadísticas), alumnos e investigadores confunden los diferentes niveles de abstracción de las hipótesis y cuando encuentran un resultado significativo, lo interpretan en relación a la hipótesis de investigación, aunque el resultado solo se refiere a las hipótesis nula y alternativa (Chow, 1996, Batanero, 2000).
- iii) *Interpretación de las probabilidades de error y sus relaciones*: Los estudiantes confunden las dos probabilidades condicionales que intervienen en la definición del nivel de significación, interpretando α como $P(H_0 \text{ cierta} / \text{Se ha rechazado } H_0)$ lo cual

sería una interpretación bayesiana; en otros casos suprimen la condición en la probabilidad o la confunden con su probabilidad complementaria. Este error fue también descrito por Birnbaum (1982), Rossman (2008) y Falk (1986) quien lo denominó “falacia de la condicional transpuesta”. Este error también se extiende a la confusión entre errores tipo I y tipo II.

- iv) *Nivel de significación*: Se confunde el nivel de significancia con la probabilidad de obtener un resultado correcto, y suelen pensar que esta probabilidad está predeterminados por el tamaño de la muestra. Cohen (1994) indica que la interpretación errónea del p-valor como probabilidad de que la hipótesis nula sea falsa es casi universal y que, además, esta interpretación viene acompañada del error de suponer que su complementario es la probabilidad de que la siguiente replicación del experimento tendrá éxito.
- v) *Criterio de decisión*: Confusión entre región de rechazo y de aceptación.
- vi) *Interpretación de resultados*: Confunden significación estadística con significación práctica, así como un resultado significativo lo ven como el resultado que corrobora la hipótesis nula y en particular al efecto de la variable manipulada por el investigador (Falk, 1986; Granaas, 2002).

Como se mencionó anteriormente, los principales errores relacionados en la comprensión del contraste de hipótesis se refieren al nivel de significación a α . Díaz, Batanero y Wilhelmi (2008), indican que la interpretación incorrecta más extendida de este concepto es cambiar los términos de la probabilidad condicional en la definición del nivel de significación a α , la cual es $P(\text{Rechazar } H_0 | H_0 \text{ Verdadera})$, interpretándolo como la probabilidad de que la hipótesis nula sea cierta, habiendo tomado la decisión de rechazarla (Birnbaum, 1982; Vallecillos, 1994; Lecoutre, Lecoutre y Poitevineau, 2001; Haller y Kraus, 2002; Lecoutre, 2006; Vera, Díaz, & Batanero, 2011; Batanero, Vera, & Díaz, 2012; Vera & Díaz, 2013). El mismo intercambio de condicional se hace en la interpretación del p-valor que se define como la probabilidad de obtener un valor igual o más extremo al dado, si la hipótesis es cierta y que se mal interpreta como probabilidad de que la hipótesis sea cierta si se obtuvo el valor dado por el estadístico.

Esta interpretación incorrecta del nivel de significancia y el p-valor se une, normalmente, a la confusión entre significación estadística y significación práctica, donde la primera se basa en la diferencia, suficientemente grande, del valor del parámetro en función de una cierta variable experimental y en cambio, la significación práctica se refiere a si el tamaño de muestra es suficiente. Ambas dos suelen ir unidas, sin embargo, en la literatura se puede encontrar datos estadísticamente significativos con un pequeño efecto experimental, siempre que la muestra sea suficientemente grande (Lecoutre, 1999).

Otro error de interpretación del p-valor es pensar que este valor indica la probabilidad de que el valor obtenido del estadístico se deba al azar, aunque esto no es cierto en general (Batanero, 2000). Así también existe el error al creer en la conservación del valor del nivel de significancia cuando se realizan contrastes consecutivos en el mismo conjunto de datos (Vera, Díaz & Batanero, 2011).

Además, existe un error frecuente en alumnos e investigadores los cuales confunden la hipótesis nula y la hipótesis alternativa (Vallecillos, 1994). La hipótesis nula o de no efecto es la negación de la hipótesis alternativa (Batanero, 2000), es más, la hipótesis nula se plantea para ser rechazada, mientras que la alternativa sería la negación de la anterior. La hipótesis nula se supone cierta y la distribución muestral del estadístico de contraste se determina aceptando que la hipótesis nula es cierta; cosa que no ocurre con la alternativa.

A pesar de todas estas diferencias, Vallecillos (1994), en su trabajo en una amplia muestra de estudiantes de distintas especialidades (n=436) encontró un 13% aproximadamente de alumnos que confunden la hipótesis nula con la alternativa al plantear las hipótesis en problemas sencillos. Además, en el estudio también encontró que aproximadamente un 20% de los estudiantes de su muestra no saben si la hipótesis del contraste se refiere al parámetro de la población o bien al estadístico muestral. Este error se describe asimismo en las investigaciones de Vallecillos y Batanero (1997) y es reportado en Castro et al. (2007) y Vera et al (2013), de modo que coinciden con Vallecillos en que los alumnos confunden algunas propiedades de las hipótesis nula y alternativa.

Un último aspecto a reseñar (Vallecillos, 1994, 1999) es que los estudiantes pueden confundir el criterio de decisión al aplicar el contraste y, por tanto, llegar a una decisión equivocada. La autora observa una confusión entre región de rechazo y de aceptación, así como sobre la forma en que se deben construir dichas regiones en un contraste unilateral o bilateral. Hay una falta de apreciación de que la hipótesis alternativa determina, junto con el nivel de significación, la región crítica, y que un mayor nivel de significación da una menor área para la región crítica. La comprensión de este punto también se relaciona con la dificultad señalada por Harradine, Batanero y Rossman (2011) de comprensión de la lógica subyacente al contraste de hipótesis.

1.3 Medición de Habilidades a través de Instrumentos

Una manera de medir el tipo de habilidades que se requiere para plantear o desarrollar un problema de test de hipótesis es a través de un instrumento, cuestionario o prueba el cual puede ser analizado a través de la Teoría Clásica de los Test (TCT) o la Teoría de Respuesta el Ítem (TRI).

Ambas teorías son complementarias sin embargo históricamente la TCT es una de las primeras formalizaciones de la teoría de medición para las pruebas educativas y psicología. La TCT descompone la calificación de la prueba observada en un puntaje verdadero y un término de error, y se supone que las variaciones de respuesta se deben a la variación en la capacidad de los sujetos, llamada puntuación verdadera y la variación de otras fuentes externas, llamada error de medición, esta relación cumple la relación matemática $X = T + e$.

La TCT se compone de varios supuestos, estos son:

- La puntuación verdadera coincide con el valor esperado de la puntuación empírica, esto es $V = E(x)$.
- El error se considera una variable aleatoria con distribución normal de media cero y varianza σ_e^2 donde las varianzas de los errores son iguales sin importar la puntuación verdadera a la que este asociado.
- No existe correlación entre las puntuaciones verdaderas de los sujetos y el error.
- No existe correlación entre los errores de medida de dos mediciones diferentes.

Bajo estos supuestos, la Teoría Clásica del Test posee gran simpleza y utilidad por lo tanto es manipulada en diversas áreas de la medición. Sin embargo, esta modelación matemática genera críticas y diversas debilidades debido a que el supuesto $X = T + e$ tiene varias interpretaciones, ya que existen diversas formas de error; y la fiabilidad del test pese a ser un concepto central en la TCT, no es posible definirlo ni estimarlo de forma unívoca. La implementación de múltiples procedimientos para estimarla ha producido confusión en los conceptos de consistencia interna, homogeneidad y unidimensionalidad y, en consecuencia, el valor del coeficiente de fiabilidad depende del método de estimación usado. Ello, unido a que la fiabilidad del instrumento depende de la longitud del test y de la variabilidad de las respuestas de los sujetos a los que se les aplica el test, hace que la fiabilidad como un parámetro característico del instrumento sea difícil de sostener.

Para superar las debilidades del TCT, en esta tesis trabajaremos la Teoría de Respuesta al Ítem en conjunto con la Medición Diagnóstica.

Capítulo II Marco Teórico

2.1 Teoría de Respuesta al Ítem

La TRI es un modelo matemático diseñado para relacionar las capacidades o habilidades, respecto de las características del ítem en relación a su dificultad, discriminación y azar. Estos modelos fueron desarrollados entre los años 1950 y 1960 con el propósito de analizar los ítems de una prueba que fuera puntuada dicotómicamente. Esta teoría se ocupa de la respuesta de los sujetos para cada ítem de la prueba.

Se considera una prueba de respuesta binaria. Además, se define $R_{i,j}$ como la respuesta del sujeto j en el ítem i , donde $R_{i,j} = 1$ si la respuesta es correcta y $R_{i,j} = 0$ si la respuesta es incorrecta.

La TRI utiliza un parámetro latente $\theta \in (-\infty, \infty)$ para denotar la habilidad de un sujeto, y representa la probabilidad de responder correctamente el ítem i –ésimo de acuerdo a una función matemática $p_i(\theta): P(R_{i,j} = 1|\theta)$. Esta función se conoce como la función de respuesta al ítem o curva característica del ítem (Tucker, 1946). La función de respuesta al ítem $p_i(\theta)$ es una función no lineal en θ debido a la restricción $0 \leq p_i(\theta) \leq 1$.

Además, por lo general es una función monótonamente creciente en θ , lo que significa que los sujetos con mayor habilidad tienen también una mayor probabilidad de acierto en el ítem i .

Diferentes funciones de respuesta al ítem, conocidas como modelos, se han propuesto en la literatura para modelar datos de respuesta de los sujetos (Rasch, 1960; Lord & Novick, 1968; Lord, 1980). Entre ellos existe el modelo Logístico de tres parámetros que se presenta a continuación.

2.1.1 Modelos Logísticos

Los modelos logísticos de tres parámetros (3-PL) son comúnmente utilizados para evaluar ítems de alternativas múltiples, ya que su probabilidad de responder correctamente un ítem captura su dificultad, discriminación y azar según el nivel de habilidad del encuestado.

El modelo tiene la siguiente función de respuesta al ítem:

$$p_i(\theta) = c_i + (1 - c_i) \frac{1}{1 + \exp\{-a_i(\theta - b_i)\}}$$

Donde a_i representa la discriminación la cual se entiende como el punto crítico donde se segrega a los sujetos que dominan y los que no dominan las habilidades del ítem; b_i para la dificultad, la cual explica el grado de experticia necesaria para resolver la pregunta, y c_i que es el parámetro de azar entendido como la menor probabilidad de responder correctamente el ítem. La estimación de estos parámetros requiere de métodos numéricos avanzados a través de su función de verosimilitud.

Cuando $c_i = 0$ el modelo es equivalente al modelo logístico de dos parámetros. Por otra parte, si $c_i = 0$ y $a_i = 1$ entonces el modelo es el llamado Modelo de Rasch.

La ilustración 1 muestra tres funciones de respuesta al ítem ajustadas al modelo 3-PL con $(a_i, b_i, c_i) = (2, 1, 0.1)$, $(a_i, b_i, c_i) = (2, 0, 0)$ y $(a_i, b_i, c_i) = (4, 2, 0.15)$.

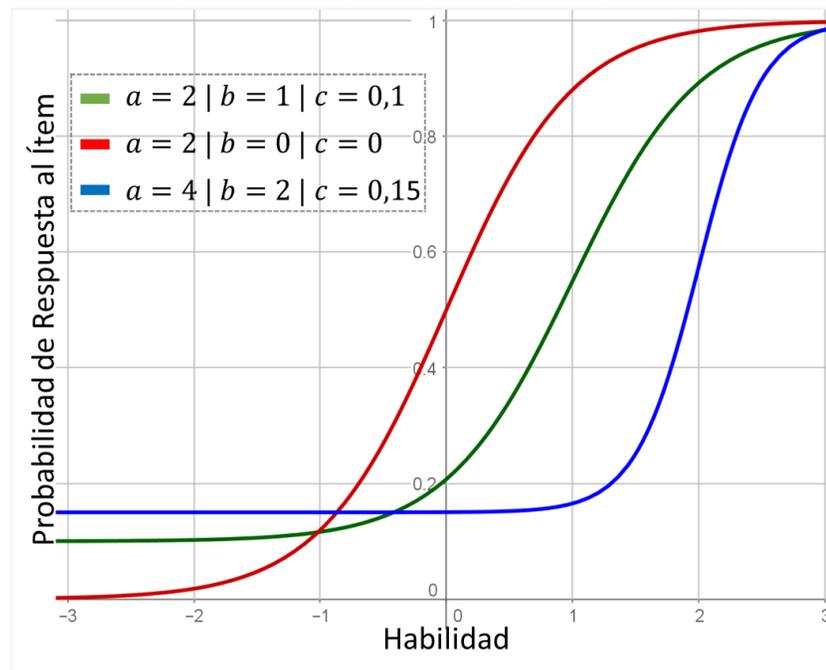


Ilustración 1: Curva Característica Modelo 3-PL

Además de los modelos anteriores que se centran en una variable unidimensional y datos de respuesta dicotómicos, otros desarrollos en la teoría de respuesta al ítem incluyen modelos de

respuesta polinómico. Para mayor información revisar: Samejima (1969); Bock (1972); Van der Linden y Hambleton (1997); Embretson y Reise (2000); Ostini y Nering (2006)); y en los modelos de TRI multidimensionales: Lord y Novick (1968); Samejima (1974); van der Linden y Hambleton (1997); Reckase (2009).

2.2 Modelo de Clasificación Diagnóstica

A partir de la década de los 90 ha surgido un gran interés por parte de la psicometría por estudiar modelos estadísticos con variables latentes como una alternativa para analizar los datos en un contexto donde las clasificaciones multivariadas de los encuestados o participantes de un instrumento de medición se encuentran sobre la base de múltiples habilidades.

Las variables latentes responden a un constructo y no son observables directamente, sino que son deducidas a partir de otras variables que sí son observables y, por lo tanto, medibles utilizando un modelo estadístico. Estos modelos son empleados en diversas disciplinas, tales como la psicología, economía, medicina, física, inteligencia artificial, bioinformática, econometría, gestión y ciencias sociales entre otras. Una de las ventajas de utilizar variables latentes es que reducen la dimensión de los datos, ya que agrupa a las variables que son observables representándolas en un concepto subyacente.

Existen variados modelos que utilizan variables latentes, un conjunto de estos son los Modelos de Clasificación Diagnóstica (DCMs) y se definen como un conjunto de herramientas estadísticas que proporcionan retroalimentación diagnóstica. En los últimos años, los Modelos de Clasificación Diagnóstica tienen un desarrollo importante en las mediciones de pruebas educativas y psicológicas ya que, en lugar de una puntuación global de las pruebas, un test diagnóstico proporciona un perfil a cada participante el cual detalla los conceptos y habilidades, también llamados atributos, que el individuo domina o no domina.

Para los DCMs es fundamental el uso de la Matriz Q, la cual es una matriz de incidencia que especifica la relación ítem-atributo. Una práctica común para la Matriz Q es ser especificada por los expertos cuando los ítems están escritos.

La finalidad de los DCMs es determinar el dominio de un conjunto de atributos o habilidades. Hay dos estadísticos directamente obtenidos de los DCMs que pueden ser utilizados para la clasificación del dominio de los encuestados: las probabilidades marginales “a posteriori” para los atributos, donde se requiere un umbral de una probabilidad para determinar el estado de maestría o el de no dominio para cada atributo y otra categoría llamada “región de la indiferencia” (para las probabilidades cercanas a 0,5), y la probabilidad “a posteriori” para el perfil del atributo, que es más sencilla que la probabilidad marginal a posteriori.

El diagnóstico cognitivo recientemente ha ganado prominencia en la evaluación educativa, psicométrica, y en muchas otras disciplinas (Rupp, Templin & Henson, 2010b) ya que, en lugar de una puntuación global de la prueba, un test de diagnóstico cognitivo proporciona a cada sujeto un perfil detallando de los conceptos y habilidades, también llamados atributos, que él o ella domina. Este tipo de evaluación conlleva interpretaciones las cuales podrían tener un impacto significativo en procesos de aprendizaje e información detallada sobre las fortalezas y debilidades de los estudiantes los cuales serían proporcionados a cada estudiante y a sus profesores.

Las teorías tradicionales de test, tales como la teórica clásica del test y los modelos de la teoría de respuesta al ítem tienen como foco principal el escalamiento y la clasificación de los estudiantes en una competencia latente de recorrido continuo. Sin embargo, la información general como la puntuación de la prueba o estimadores de una competencia obtenidos sobre la base de esos modelos es unidimensional y que no contiene suficiente información para las múltiples habilidades específicas para los atributos en términos de diseño de instrucción efectiva y proporcionar intervención.

Como alternativa, los Modelos de Clasificación Diagnóstica, que también se conocen como modelos de Diagnóstico Cognitivo (Rupp, Templin & Henson, 2010b; Templin 2008), se han desarrollado con el propósito de identificar la presencia o ausencia de múltiples habilidades o atributos detallados.

Estadísticamente hablando, los Modelos de Clasificación Diagnóstica pertenecen a la familia de modelos de estructuras latentes. En particular, los modelos de clase latente pertenecen a una familia más amplia de modelos lineales generalizados y mixtos no lineales que se restringe basados en la denominada Matriz Q (Tatsuoka, 1983), que especifica la relación entre los ítems del test y los atributos latentes.

A continuación, se presentan las nociones básicas de todo Modelo de Clasificación Diagnóstica que utiliza una Matriz Q en su estudio y modelo de variables latentes.

2.2.1 Notación

El número total de encuestados se entenderá por N y la cantidad total de ítems que contiene el test se entenderá por L . Las respuestas dadas son de naturaleza binaria, por lo cual, los datos serán ordenados en una matriz de orden $N \times L$ con valores “0” para respuestas incorrectas y “1” para respuestas correctas. Además, existe una cantidad finita K de atributos o habilidades que están relacionadas tanto a los ítems como a los encuestados el cual es utilizado por el Modelo de Clasificación Diagnóstica.

La **Matriz Q** es el componente clave de los Modelos de Clasificación Diagnostico. En ella se especifica la relación entre los ítems de la prueba y los atributos o habilidades latentes. En particular $Q = \{q_{lk}\}_{L \times K}$ es una matriz de orden $L \times K$ con entradas binarias. Para cada l y k , $q_{lk} = 1$ indica que el ítem l está asociado al atributo k , y $q_{lk} = 0$ en otro caso.

Tomemos la siguiente Matriz Q como un ejemplo:

	Adición	Multiplicación
16 + 1	1	0
5 x 4	0	1
(4 x 7) + 10	1	1

Tabla 1: Matriz Q

Hay dos atributos y tres ítems. El primer ítem está asociado a la suma, el segundo a la multiplicación y el tercero requiere habilidades tanto para la suma como para la multiplicación.

La respuesta a los ítems se define por el vector:

$$\mathbf{R} = (R^1, \dots, R^L)^T$$

Donde R es el arreglo de respuesta asociado a los J ítems de una prueba. Estas respuestas son de naturaleza binaria, es decir, para cada l , R^l es una variable binaria que toma los valores 0 o 1, y el superíndice T indica la traspuesta. El valor de L indica la cantidad de vectores de respuesta posibles, por ejemplo, en la tabla 1, $L = 3$ por lo tanto existen $2^3 = 8$ posibles vectores de respuestas.

El **perfil de Atributos** que relaciona los K atributos existentes asociados a un individuo se define como:

$$\boldsymbol{\alpha} = (\alpha^1, \dots, \alpha^K)^T$$

Donde α es el vector de atributos, también llamado vector de habilidades, donde α^k posee naturaleza binaria, es decir, puede tomar valores 0 o 1 que indica la ausencia o presencia del k –ésimo atributo, con $k = 1, \dots, K$. Para la tabla 1, $K = 2$, por lo tanto, tiene $2^2 = 4$ posibles

perfiles de atributos, por extensión son: $(0,0), (0,1), (1,0), (1,1)$, donde el primer elemento representa la suma y el segundo la multiplicación.

Es importante notar que α y R , están asociados a las respuestas y habilidades de un sujeto, por lo tanto, se utilizarán subíndices para asociarlos a cada participante de la prueba. Por ejemplo, $R_i = (R_i^1, \dots, R_i^L)^T$ es el vector de respuesta del sujeto i y análogamente $\alpha_i = (\alpha_i^1, \dots, \alpha_i^K)^T$ es el vector de habilidades para el participante i .

Con N sujetos, se observarán R_1, \dots, R_N respuestas, pero no $\alpha_1, \dots, \alpha_N$ atributos ya que el propósito principal del diagnóstico cognitivo es estimar con precisión el perfil de atributos $\alpha_1, \dots, \alpha_N$ a partir de los datos de respuesta R_1, \dots, R_N . A veces, también es interesante la proporción de sujetos con ciertos perfiles de atributo, que se especifica de la siguiente manera.

Para la **proporción** se asume que los perfiles de atributos son independiente e idénticamente distribuidos (en inglés Independent and Identically Distributed ó i.i.d.) y la proporción de sujetos con perfil de atributo α está dada por:

$$p_\alpha = P(\alpha_i = \alpha)$$

Escribiremos $p = (p_\alpha: \alpha \in \{0,1\}^K)^T$. El perfil completo se descompone en el primer elemento que posee cero atributos dominados y se expresado por: $p_0 = P(\alpha_i = (0, \dots, 0)_{K \times 1}^T)$, y el segundo elemento el cual contempla los perfiles con algún tipo de atributo dominado.

Se ocupan los vectores $\mathbf{0}_K$ y $\mathbf{1}_K$ para representar a los vectores columna K-dimensionales cero y uno, respectivamente. Cuando no hay ninguna otra posible interpretación, omitimos el índice de la longitud y sólo se escribe $\mathbf{0}$ y $\mathbf{1}$.

Finalmente, la **respuesta ideal** se denota por $\xi^j(\alpha, Q)$ la cual indica si un sujeto posee el perfil del atributo o habilidad α y es capaz de proporcionar una respuesta positiva al ítem j si la relación sujeto-atributo está especificada por la matriz Q .

En la actualidad existen numerosos modelos de clasificación diagnóstica debido a sus diversas estructuras de respuesta ideal, en general existen dos categorías: modelos compensatorios y modelos conjuntivos; a continuación, se presenta una descripción de estos y se presentan varios modelos de diagnóstico populares en la literatura. Una revisión más exhaustiva de esos modelos se puede encontrar en Rupp et al. (2010).

2.2.2 Modelos de Clasificación Diagnóstica Compensatorio

El modelo de clasificación diagnóstica compensatorio necesita atributos específicos para su interpretación. La falta de un atributo puede ser compensada por la presencia de otro, por ejemplo, existen casos donde diferentes atributos representan diferentes estrategias en el encuentro de la solución de un ítem; por lo tanto, una respuesta acertada al ítem sólo requiere del éxito de una estrategia, es decir, basta con un solo atributo exitoso. La respuesta ideal en el modelo compensatorio está dada por:

$$\xi^J(\alpha, Q) = I(\alpha^k \geq q_{jk}; \forall k = 1, \dots, k) = 1 - \prod_{k=1}^K (1 - \alpha^k)^{q_{jk}}$$

El modelo DINO (Deterministic Input, Noisy output “OR” gate) es un modelo diagnóstico compensatorio (Templin y Henson 2006; Templin 2008). La respuesta ideal $\xi_{DINO}^J(\alpha, Q)$ tiene la forma planteada anteriormente. Por lo tanto, necesita dominar sólo un atributo de la matriz Q para ser capaz de responder adecuadamente el ítem. Utilizando las definiciones del modelo DINA, que se mostrará en el capítulo 2.2.3.1, para c_j y g_j , la respuesta ideal bajo el modelo DINO tiene la forma:

$$P(R^j = 1 | \alpha, c_j, g_j) = c_j^{\xi_{DINO}^J(\alpha, Q)} g_j^{1 - \xi_{DINO}^J(\alpha, Q)}$$

Gráficamente:

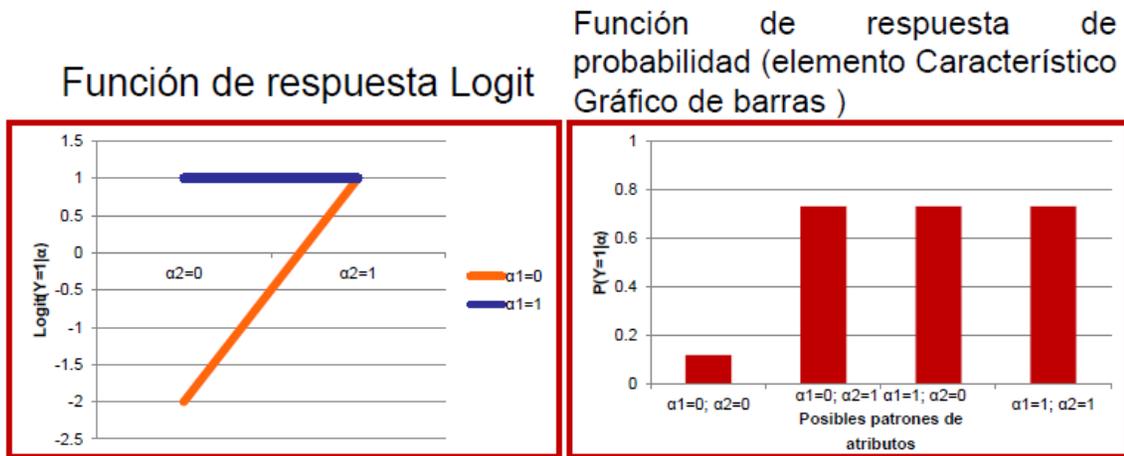


Ilustración 2: Función Respuesta Modelo DINO

2.2.3 Modelos de Clasificación Diagnóstica Conjuntista

Para los modelos conjuntistas, la respuesta ideal estará dada por:

$$\xi^J(\alpha, Q) = I(\alpha^k \geq q_{jk}; \forall k = 1, \dots, k)$$

Donde $I(\cdot)$ es la función característica habitual. Esta ecuación muestra que se requieren todas las habilidades especificadas por la matriz Q para un desempeño exitoso en el ítem correspondiente ($\xi = 1$) y la falta de estas competencias en cualquier atributo requerido conduce a una respuesta ideal negativa ($\xi = 0$). En otras palabras, sin importar la cantidad de atributos adicionales innecesarios, estos no compensarán la falta de atributos necesarios.

2.2.3.1 Modelo DINA

Un modelo conjuntivo de diagnóstico básico y popular es el modelo DINA (Deterministic Input, Noisy output “AND” gate) (Junker y Sijtsma, 2001). El modelo DINA introduce los llamados parámetros de “deslizamiento” y “azar” (Gongjun, 2013). El concepto se debe a Macready y Dayton (1977) para las pruebas de domino, asimismo Van der Linden en 1978 hace referencia a este concepto. El parámetro de “deslizamiento” es la probabilidad de que un sujeto con perfil de atributo α responde negativamente a un ítem si la respuesta ideal para ese ítem está dada por $\xi^J(\alpha, Q) = 1$; de forma similar, el parámetro de “azar” se refiere a la probabilidad de que un sujeto responda positivamente a si su respuesta ideal está dada por $\xi^J(\alpha, Q) = 0$.

Se utiliza $s = (s_1, \dots, s_J)^T$ para denotar la probabilidad de *deslizamiento* y $g = (g_1, \dots, g_J)^T$ para las probabilidades de *azar* para una cantidad J de ítems. En la discusión es más conveniente trabajar con el complemento del parámetro de *deslizamiento*. Por lo tanto, definimos $c = 1 - s$ como la probabilidad de responder correctamente, con c_{ij} como la notación correspondiente al ítem específico.

Dado un específico perfil de un sujeto α , la respuesta al ítem j y bajo el modelo DINA sigue una distribución Bernoulli:

$$P(R^j = 1 | Q, \alpha, c_j, g_j) = c_j^{\xi^j(\alpha, Q)} g_j^{1 - \xi^j(\alpha, Q)}$$

Además, condicionado por α , $R = (R^1, \dots, R^J)$ son conjuntamente independientes.

Gráficamente:

Función de respuesta Logit

Función de respuesta de probabilidad (elemento Característico Gráfico de barras)

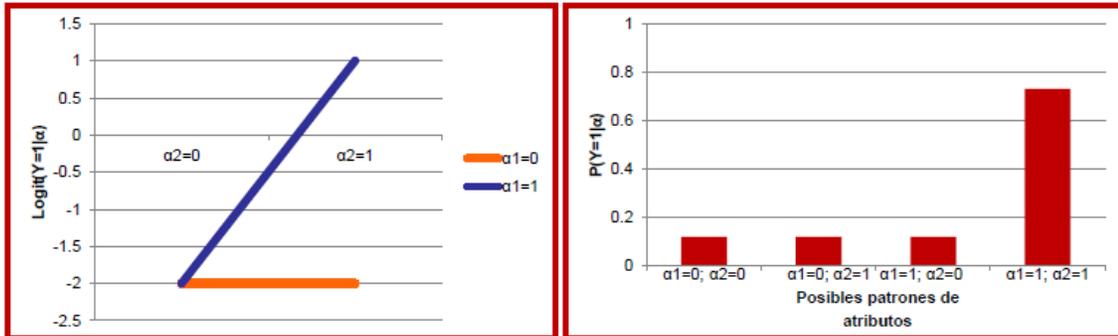


Ilustración 3: Función Respuesta Modelo DINA

En el caso particular de $K = 1$, la curva característica del ítem bajo el modelo DINA puede ser tomada como una versión discretizada de la función característica del ítem de un modelo de TRI.

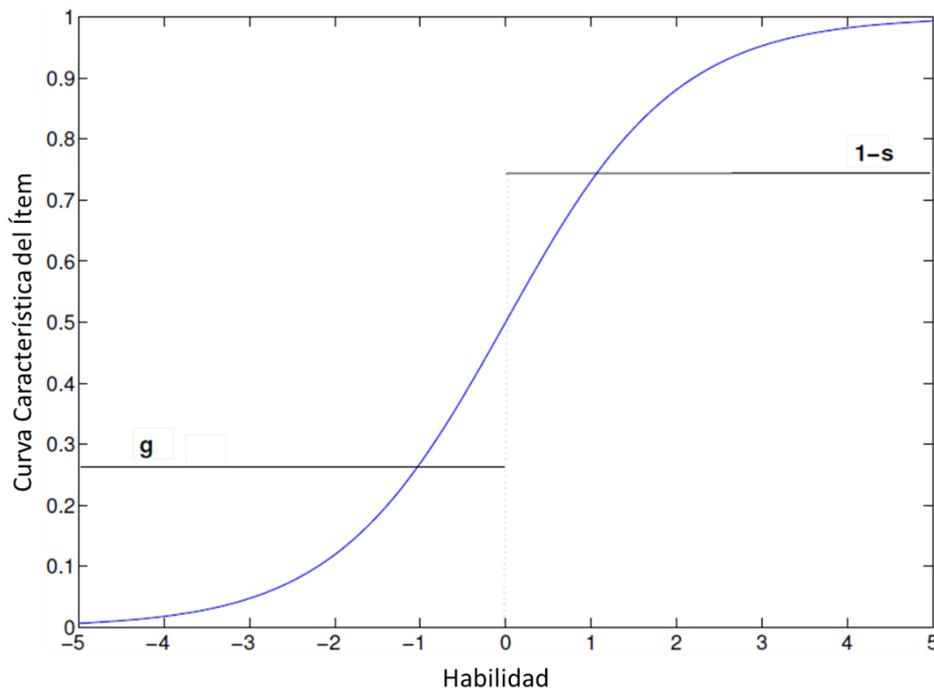


Ilustración 4: Curva Característica Modelo DINA

Se puede observar que los sujetos con baja capacidad se toman como no dominantes de la capacidad necesaria y tienen probabilidad de respuesta correcta igual a g , mientras que aquellos con alta capacidad están considerados como dominantes de la habilidad y tienen probabilidad de respuesta correcta $1 - s$. En la literatura, esto también se conoce como la prueba de dominio (Ver Macready y Dayton (1977); Van der Linden (1978)).

Una generalización del modelo DINA es una versión reducida del Modelo Unificado Reparametrizado (RUM) (Hartz, 2002). Bajo el modelo RUM Reducido, la probabilidad toma la forma para un ítem j :

$$P(R^j = 1 | \alpha, \pi_j, r_{jk}) = \pi_j \prod_{k=1}^K r_{jk}^{\alpha_j k (1 - \alpha^k)}$$

Donde π_j es la probabilidad de respuesta correcta para un sujeto que posee todos los atributos requeridos. Además, r_{jk} es el parámetro de penalización por no poseer el atributo k , donde $0 < r_{jk} < 1$. El modelo RUN Reducido es también un modelo conjuntivo y generaliza el modelo DINA permitiendo que los parámetros de *deslizamiento* y *azar* varíen a través de variados perfiles de atributos.

Además, recientemente se ha demostrado matemática que el modelo DINA plenamente conjuntivo el cual combina todas las habilidades requeridas en una función binaria, puede ser reformulado como un caso especial de compensación de la modelación general diagnóstica compensatoria/aditiva. Esto se puede lograr en más de una forma, tal que las definiciones transformadas habilidades-espacio resultantes y matrices de diseño (Matriz Q) son diferentes entre sí, pero matemáticamente equivalentes al modelo DINA, produciendo idénticas probabilidades de respuesta basado en estos modelos.

2.2.3.2 Modelo de Clasificación Diagnóstica Loglineal

El Modelo de Clasificación Diagnóstica Loglineal (LCDM) pertenece a los modelos diagnósticos de múltiples variables latentes binarias combinadas con el fin de producir una estructura latente que proporciona más información sobre el rendimiento de los examinados que los modelos de variables latentes unidimensionales.

Los acontecimientos recientes en el modelo diagnóstico hacen hincapié en la posibilidad de que múltiples habilidades pueden interactuar de manera conjunta dentro de la función del ítem, mientras que las habilidades individuales todavía pueden conservar efectos aditivos separables.

La extensión de cualquier modelo DINA a la versión generalizada (G-DINA) o el modelo general de diagnóstico compensatorio/aditivo (GDM) para la modelación diagnóstica cognitiva Loglineal está dirigido a la integración de los modelos con habilidades conjuntivas y las que asumen el funcionamiento compensatorio de múltiples variables de habilidad.

En los modelos LCDM, los log-odds (logit) de una respuesta correcta condicionada a un patrón de atributo de un encuestado α_r están dados por:

$$\text{Logit}(Y_{ri} = 1 | \alpha_r) = \ln \left(\frac{P(Y_{ri} = 1 | \alpha_r)}{1 - P(Y_{ri} = 1 | \alpha_r)} \right)$$

El logit se utiliza porque las respuestas son binarias, es decir, los ítems se respondieron correctamente (1) o incorrectamente (0). El modelo lineal con un enlace de identidad y error gaussiano es inapropiado para datos categóricos debido a que pueden conducir a predicciones imposibles, existen casos donde las probabilidades son mayores que uno o menores a cero.

Se puede apreciar el comportamiento de la función para distintos valores:

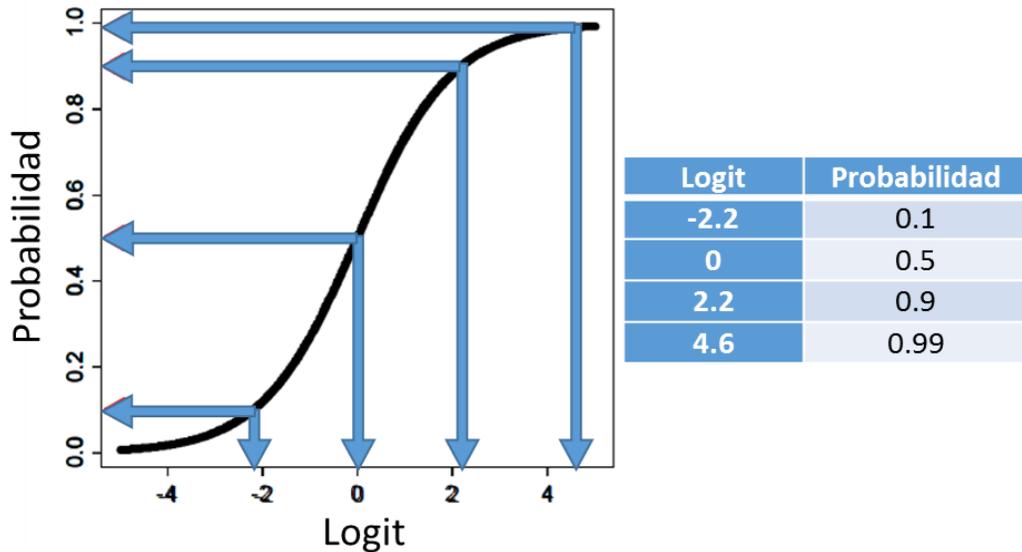


Ilustración 5: Curva Característica LCDM

El LCDM proporciona el logit de una respuesta correcta para el ítem i el cual está relacionado a dos habilidades latentes de un examinado r :

$$\text{Logit}(Y_{ri} = 1|\alpha_r) = \lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{r1} + \lambda_{i,1,(2)}\alpha_{r2} + \lambda_{i,2,(1,2)}\alpha_{r1}\alpha_{r2}$$

Donde:

- ✓ $\text{Logit}(Y_{ri} = 1|\alpha_r)$: es el logit de una respuesta correcta al elemento i por el encuestado r .
- ✓ $\lambda_{i,0}$: Es el intercepto, es decir, el logit para los examinados que no dominan ningún atributo ($\alpha_{r1} = \alpha_{r2} = 0$), también llamado grupo de referencia.
- ✓ $\lambda_{i,1,(1)}$: Efecto principal condicional para el atributo 1, marca el Incremente del logit cuando el examinado domina el primer atributo (independiente del dominio del segundo atributo).
- ✓ $\lambda_{i,1,(2)}$: Efecto principal Condicional para el atributo 2, marca el Incremente del logit cuando el examinado domina el segundo atributo (independiente del dominio del primer atributo).
- ✓ $\lambda_{i,2,(1,2)}$: Es la interacción entre los dos atributos. Marca el cambio en el logit para la maestría tanto en el primer como en el segundo atributo.

Mientras que los logit son útiles para las variables continuas sin límites, los datos categóricos se basan en los análisis de probabilidades estimadas. La función inversa logit convierte el logit sin límites en una probabilidad:

$$P(Y_{ri} = 1|\alpha_r) = \frac{\exp(\text{Logit}(P(Y_{ri} = 1|\alpha_r)))}{1 + \exp(\text{Logit}(P(Y_{ri} = 1|\alpha_r)))}$$

En donde esta es también la forma de un modelo TRI y de regresión logística.

Supongamos el siguiente ejemplo, Ítem 1.- Calcule los valores tipificados de $\alpha = 0,1$ en un contraste bilateral para la μ , con varianza poblacional conocida.

EL LCDM contiene los siguientes cuatro parámetros para un ítem i : un parámetro de intercepción ($\lambda_{i,0}$), dos parámetros de efectos principales ($\lambda_{i,1,(1)}$; $\lambda_{i,1,(2)}$) y un último parámetro para la interacción entre los atributos 1 y 2 ($\lambda_{i,2,(1,2)}$). Supongamos, además, que los parámetros poseen los siguientes valores:

Parámetro	Logit	Efecto
$\lambda_{i,0}$	-2	Intercepto
$\lambda_{i,1,(1)}$	2	Efecto Principal del Atributo 1: Región crítica y de aceptación
$\lambda_{i,1,(2)}$	1	Efecto Principal del Atributo 2: Contraste unilateral y bilateral
$\lambda_{i,2,(1,2)}$	0	Interacción entre los atributos

Tabla 2: Ejemplo Matriz-Q LCDM

Por lo tanto:

α_{r1}	α_{r2}	Función Logit de LCDM	Logit	Probabilidad
0	0	$\lambda_{i,0} + \lambda_{i,1,(1)} \cdot 0 + \lambda_{i,1,(2)} \cdot 0 + \lambda_{i,2,(1,2)} \cdot 0 \cdot 0$	-2	0,12
0	1	$\lambda_{i,0} + \lambda_{i,1,(1)} \cdot 1 + \lambda_{i,1,(2)} \cdot 0 + \lambda_{i,2,(1,2)} \cdot 1 \cdot 0$	-1	0,27
1	0	$\lambda_{i,0} + \lambda_{i,1,(1)} \cdot 0 + \lambda_{i,1,(2)} \cdot 1 + \lambda_{i,2,(1,2)} \cdot 0 \cdot 1$	0	0,5
1	1	$\lambda_{i,0} + \lambda_{i,1,(1)} \cdot 1 + \lambda_{i,1,(2)} \cdot 1 + \lambda_{i,2,(1,2)} \cdot 1 \cdot 1$	1	0,73

Tabla 3: Ejemplo Logit LCDM

Gráficamente:

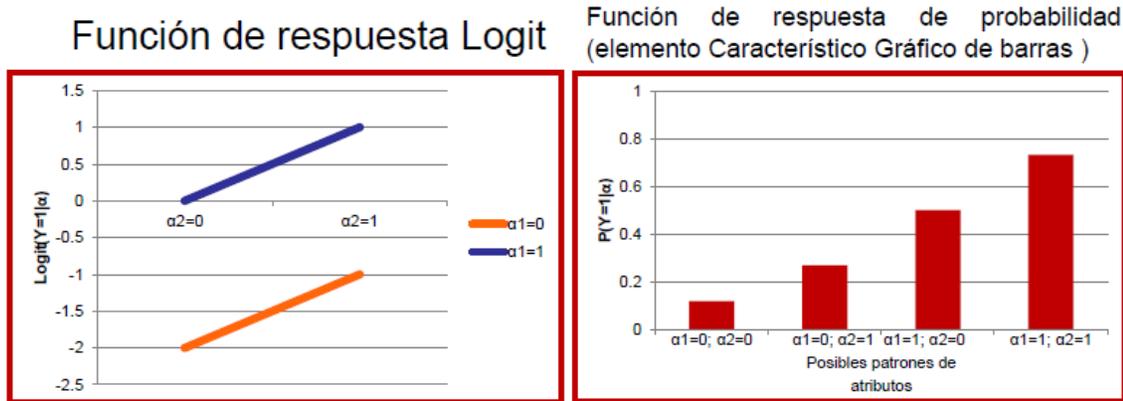


Ilustración 6: Función de Respuesta LCDM

En este caso, la tabla 3, muestra la conversión en una probabilidad de una respuesta correcta de acuerdo con la función definida por el LCDM.

Modelo General

El LCDM tiene como base al Modelo Diagnostico General de Von Davier (2014), en donde se permite utilizar variables latentes categóricas y continuas. Para ítems que miden más de dos atributos, las interacciones de alto nivel son posibles pero difíciles de estimar. La forma general del LCDM aparece en la literatura de la psicometría en una forma general (Henson, Templin & Willse, 2009; Templin, 2016).

El MCDL especifica la probabilidad de una respuesta correcta como una función de un conjunto de atributos y una matriz Q:

$$P(Y_{ri} = 1|\alpha_r) = \frac{\exp(\lambda_i^T h(q_i, \alpha_r))}{1 + \exp(\lambda_i^T h(q_i, \alpha_r))}$$

La forma general del MCDL es una ecuación que hace posible que cualquier número de atributos sea explicado por un ítem. Para que esto ocurra, se debe saber el vector correspondiente al perfil de atributos ($r: \alpha_r = [\alpha_{r1} \ \alpha_{r2} \ \alpha_{r3} \ \dots]$) y el vector fila para cada ítem de la matriz Q ($i: q_i = [a_1 \ a_2 \ a_3 \ \dots]$) tq $a_i \in \{0,1\}$).

Además, a partir de la notación general del LCDM, λ_i es un vector de todos los posibles parámetros de los ítems para el ítem i . Cabe destacar que si todas las entradas A en la matriz Q en q_i son equivalentes a 1, el tamaño es $2^A \times 1$. Es importante aclarar que no todos los parámetros se pueden estimar si algún $q_{ia} = 0$.

A continuación, un ejemplo para cuatro atributos:

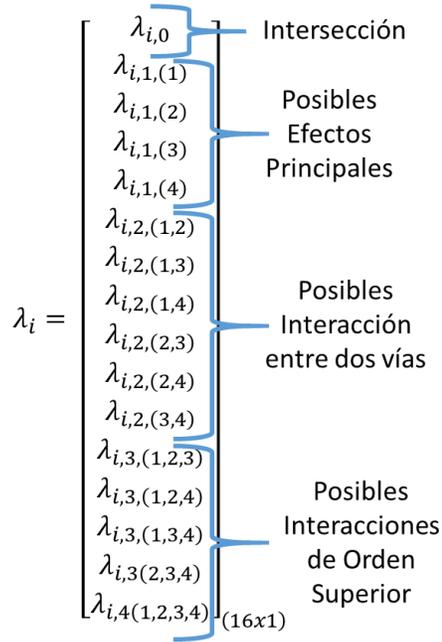


Ilustración 7: Ejemplo estimador para 4 parámetros

Al mismo tiempo, la forma general para cuatro parámetros en el MCDL posee la Función auxiliar $h(q_i, \alpha_r)$ la cuál es el vector evaluador de la función y tiene la forma:

$\lambda_i =$	$h(q_i, \alpha_r) =$
$\begin{bmatrix} \lambda_{i,0} \\ \lambda_{i,1,(1)} \\ \lambda_{i,1,(2)} \\ \lambda_{i,1,(3)} \\ \lambda_{i,1,(4)} \\ \lambda_{i,2,(1,2)} \\ \lambda_{i,2,(1,3)} \\ \lambda_{i,2,(1,4)} \\ \lambda_{i,2,(2,3)} \\ \lambda_{i,2,(2,4)} \\ \lambda_{i,2,(3,4)} \\ \lambda_{i,3,(1,2,3)} \\ \lambda_{i,3,(1,2,4)} \\ \lambda_{i,3,(1,3,4)} \\ \lambda_{i,3,(2,3,4)} \\ \lambda_{i,4,(1,2,3,4)} \end{bmatrix}_{(16 \times 1)}$	$\begin{bmatrix} 1 \\ (q_{i1}, \alpha_{r1}) \\ (q_{i2}, \alpha_{r2}) \\ (q_{i3}, \alpha_{r3}) \\ (q_{i4}, \alpha_{r4}) \\ (q_{i1}, \alpha_{r1})(q_{i2}, \alpha_{r2}) \\ (q_{i1}, \alpha_{r1})(q_{i3}, \alpha_{r3}) \\ (q_{i1}, \alpha_{r1})(q_{i4}, \alpha_{r4}) \\ (q_{i2}, \alpha_{r2})(q_{i3}, \alpha_{r3}) \\ (q_{i2}, \alpha_{r2})(q_{i4}, \alpha_{r4}) \\ (q_{i3}, \alpha_{r3})(q_{i4}, \alpha_{r4}) \\ (q_{i1}, \alpha_{r1})(q_{i2}, \alpha_{r2})(q_{i3}, \alpha_{r3}) \\ (q_{i1}, \alpha_{r1})(q_{i2}, \alpha_{r2})(q_{i4}, \alpha_{r4}) \\ (q_{i1}, \alpha_{r1})(q_{i3}, \alpha_{r3})(q_{i4}, \alpha_{r4}) \\ (q_{i2}, \alpha_{r2})(q_{i3}, \alpha_{r3})(q_{i4}, \alpha_{r4}) \\ (q_{i1}, \alpha_{r1})(q_{i2}, \alpha_{r2})(q_{i3}, \alpha_{r3})(q_{i4}, \alpha_{r4}) \end{bmatrix}_{(16 \times 1)}$

Tabla 4: Estimador y Función Auxiliar para 4 Parámetros

Por lo tanto, el logit buscado se obtiene a través de la matriz producto: $\lambda_i^T h(q_i, \alpha_r)_{(1 \times 1)}$.

Extendiendo esta expresión, se obtiene:

$$\begin{aligned}
 \text{Logit}(Y_{ri} = 1|\alpha_r) &= \lambda_{i,0} + \sum_{a=1}^4 \lambda_{i,1,(a)}(q_{ia}\alpha_{ra}) + \sum_{a=1}^3 \sum_{b=a+1}^4 \lambda_{i,2,(a,b)}(q_{ia}\alpha_{ra})(q_{ib}\alpha_{rb}) \\
 &+ \sum_{a=1}^2 \sum_{b=a+1}^3 \sum_{c=b+1}^4 \lambda_{i,3,(a,b,c)}(q_{ia}\alpha_{ra})(q_{ib}\alpha_{rb})(q_{ic}\alpha_{rc}) \\
 &+ \lambda_{i,4(1,2,3,4)}(q_{i1}, \alpha_{r1})(q_{i2}, \alpha_{r2})(q_{i3}, \alpha_{r3})(q_{i4}, \alpha_{r4})
 \end{aligned}$$

Finalmente, la expresión general es:

$$\begin{aligned}
 \text{Logit}(Y_{ri} = 1|\alpha_r) &= \lambda_i^T h(q_i, \alpha_r) \\
 &= \lambda_{i,0} + \sum_{a=1}^A \lambda_{i,1,(a)}(q_{ia}\alpha_{ra}) + \sum_{a=1}^{A-1} \sum_{b=a+1}^A \lambda_{i,2,(a,b)}(q_{ia}\alpha_{ra})(q_{ib}\alpha_{rb}) \\
 &+ \sum_{a=1}^{A-2} \sum_{b=a+1}^{A-1} \sum_{c=b+1}^A \lambda_{i,3,(a,b,c)}(q_{ia}\alpha_{ra})(q_{ib}\alpha_{rb})(q_{ic}\alpha_{rc}) + \dots
 \end{aligned}$$

Capítulo III Metodología de Trabajo

Interesado por lo expuesto en el Capítulo II y por el hecho de que la comprensión y correcto manejo en el contraste de hipótesis es un requisito para generar nuevos conocimientos, el objetivo de esta investigación es estudiar las competencias en la formulación y asignación de hipótesis nula y alternativa, el análisis de la región de rechazo y el cálculo e interpretación de los errores tipo I y II.

Para generar el análisis se utilizó los softwares Microsoft Excel, R-Commander, Mplus y SAS manipulando las macros facilitadas por Jonathan Templin en su página web.

3.1 Muestra

La muestra está compuesta por 300 individuos que cumplan con ser estudiantes de pedagogía en matemática o docentes de matemática de Chile y que estén cursando o tengan aprobada la asignatura de Estadística Inferencial.

3.2 Variables

Edad: Se define como los años cumplidos del encuestado hasta la fecha en la cual resolvió el cuestionario.

Género: Es el sexo del participante, respetando la declaración del encuestado.

Condición Académica: Es la situación académica del encuestado al momento de desarrollar el cuestionario, esta situación puede diferir entre estudiante, profesor o la combinación de ambas para aquellos docentes que sigan sus estudios de postgrado.

Ítem: Se refiere a cada una de las preguntas que forman parte del cuestionario.

3.2.1 Variables Latentes

Es importante recordar que las variables latentes no son observables directamente, sino que son deducidas a partir de variables que si son medibles. Además, las variables latentes son de gran importancia puesto que son utilizadas en la construcción de la matriz Q (ver capítulo 3.5) que relaciona las habilidades con cada ítem.

Las variables latentes trabajada en esta investigación son:

Hipótesis: Se entiende por la capacidad que tiene el estudiante por formular, plantear, establecer, asignar o diferenciar entre la hipótesis nula y la alternativa. Esta decisión también es crucial para determinar el criterio de rechazo o aceptación de la hipótesis nula, entendiendo que el estadístico de contraste se determina aceptando que la hipótesis nula es cierta.

Región de Rechazo: Es la capacidad que tiene el estudiante para establecer, calcular, discernir o analizar la región crítica de rechazo o aceptación de la hipótesis nula en un contraste de hipótesis.

Errores tipos I y tipo II: Se entiende como la capacidad del estudiante para interpretar, calcular o diferenciar el nivel de significación α , p-valor y la probabilidad del error tipo I. Así también se entiende como las habilidades para analizar, calcular e interpretar la probabilidad del error tipo II y la potencia del test de hipótesis.

3.3 Hipótesis

Las hipótesis de trabajo se confrontarán en base a los patrones de respuesta de los encuestados y sus perfiles de clases según el LCDM mostrado en el Capítulo IV. Las hipótesis de esta investigación son:

- H1.** Un individuo puede presentar un menor número de ítems correcto en comparación a otro individuo y aun así poseer una mayor habilidad.
- H2.** Los sujetos que tienen un mayor dominio en la región de rechazo y los errores tipo I y II tienen habilidades más altas en relación a los sujetos que dominan las hipótesis y el error tipo I y II en un contraste de hipótesis.
- H3.** Los sujetos que dominan las habilidades relacionadas a la generación de hipótesis y el error tipo I y II en un contraste de hipótesis, poseen una probabilidad de acierto mayor a 0,6 en cada ítem.
- H4.** En un test de hipótesis, los encuestados que dominan las habilidades relacionadas a la región de rechazo y el error tipo I y II poseen mayor probabilidad de acierto a los ítems que los encuestados con un dominio en las habilidades relacionadas a las hipótesis y la región de rechazo.
- H5.** Los docentes que continúan sus estudios presentan perfiles de respuesta con mayor cantidad de dominios en comparación con el resto de los encuestados.
- H6.** Los estudiantes presentan perfiles de respuesta con menor cantidad de dominios en comparación con el resto de los encuestados.

3.4 Instrumento

El instrumento de recolección de datos es un cuestionario presentado a través de la plataforma *Google Formular*, este examen posee siete ítems de selección múltiple con cuatro alternativas donde sólo una es correcta.

Para elaborar el cuestionario en forma rigurosa, se comenzó con una definición de las “habilidades cognitivas para el desarrollo de test de hipótesis”, delimitando las variables latentes que se deseaba evaluar. Esta información tabular se encuentra en la Tabla 5, la cual entendemos por Matriz Q, en ella se muestra la relación que existe entre los ítems del cuestionario y los dominios que se quieren diagnosticar en los encuestados.

Los ítems que se muestran a continuación están seleccionados a partir del trabajo de Batanero, Vera y Díaz (2011) y Vallecillos & Batanero (1997b), los cuales han sido modificados para contextualizarlos a la realidad chilena. Estos ítems, según los autores, fueron seleccionados a partir de un banco de ítems previamente construido mediante pruebas piloto de ítems y valoración mediante el juicio de expertos.

Cabe destacar que las alternativas correctas están ennegrecidas y las justificaciones de los ítems se encuentran al final de cada uno.

Ítem 1 La puntuación típica correspondiente a un $\alpha = 0,05$ en un contraste derecho es aproximadamente:

- a) 1,64**
- b) -2,57
- c) -1,64
- d) 2,57

Justificación

El ítem evalúa la comprensión de la región de rechazo dada una puntuación crítica, también se requiere del manejo de tablas estadísticas. La respuesta correcta es la (a), pues si el nivel de significación es 0,05 para un test donde importan el extremo derecho, ese es el área que deja a la derecha el percentil 1,64 de la distribución normal. En el distractor (c) se confunde con el test izquierdo. El distractor (d) y (b) implica errores en el manejo de las tablas, especialmente el último distractor ocurre cuando el estudiante también confunde con un test izquierdo.

Ítem 2 Un laboratorio está produciendo un jarabe contra la tos utilizando distintas combinaciones de los componentes. De acuerdo a la norma debe tener un etiquetado de 15 mg/ml. En este contexto, se ha extraído una muestra de 110 envases elegidos al azar y se estudian sus concentraciones. Se juzga que el proceso está fuera de la ley cuando su concentración media es menor a 14,99 o mayor a 15,01 mg/ml. Suponiendo que la cantidad en cada recipiente se encuentra aproximada por una distribución normal con una desviación estándar de 0,05 mg/ml y que la hipótesis nula es que el jarabe cumple con lo indicado en su etiqueta. Obtenga la probabilidad del error de tipo I.

- a) 0,16
- b) 0,04**
- c) 0,84
- d) 0,96

Justificación

Este ítem evalúa la comprensión del error tipo I, el cálculo de su probabilidad y la utilización del teorema del límite central (TLC). La respuesta correcta es la (b) porque si definimos:

\bar{x} : *Concentración Media de Medicamento*

Tenemos: $H_0: \mu = 15$; $H_1: \mu \neq 15$ con $\sigma = 0,05$ y $n = 110$

Utilizando el TLC, la va. $\bar{x} \sim N\left(\mu, \left(\frac{\sigma}{n}\right)^2\right)$

En caso de que la hipótesis nula sea cierta $\bar{x} \sim N(15, (0,0048)^2)$

Por lo tanto, la probabilidad del error tipo I es:

$$P(\text{Rechazar } H_0 | \mu = 15) = 1 - P(14,99 \leq \bar{x} \leq 15,01) = 1 - P(-2,083 \leq z \leq 2,083) = 0,04$$

El estudiante elegirá el distractor (d) cuando calcule el complemento de la probabilidad deseada, es decir: $P(14,99 \leq \bar{x} \leq 15,01)$. Mientras que si elige el distractor (a) o (c) no estará utilizando el teorema del límite central, es decir: Calculan: $P(14,99 \leq \bar{x} \leq 15,01) = P(-0,2 \leq z \leq 0,2)$. En este caso los encuestados no toman en cuenta la alta probabilidad que arroja el cálculo de esta.

Ítem 3 Supongamos un contraste sobre la media, siendo la variable estudiada la calidad de vida en escala 0 a 100. Para $H_0: \mu = 100$, $H_1: \mu = 110$, $\alpha = 0,05$ y $\beta = 0,4406$. ¿Cuál es la probabilidad de rechazar H_0 cuando no es cierta?

- a) 0,05
- b) 0,5594**
- c) 0,4406
- d) 0,95

Justificación

El ítem permite evaluar la comprensión de la potencia de una prueba y probabilidad de error tipo II. La respuesta correcta es la (b), ya que la potencia se define como la probabilidad de rechazar la hipótesis nula siendo falsa. El alumno que elige el distractor (c) confunde la potencia con la probabilidad del error tipo II. Los estudiantes que eligen el distractor (a) confunden la potencia con la probabilidad de cometer un error tipo I y finalmente, los estudiantes que eligen el distractor (d) confunden la potencia con la confiabilidad.

Ítem 4 Queremos conocer si difiere la proporción de asignaturas aprobadas por estudiantes que son de la Región Metropolitana (P_M) y los que pertenecen a otras regiones (P_O) y no disponemos de ninguna información previa. La hipótesis nula que se debe plantear es:

- a) $P_M = P_O$
- b) $P_M \neq P_O$
- c) $P_M \leq P_O$
- d) $P_M \geq P_O$

Justificación

Este ítem evalúa la formulación de las hipótesis estadísticas, particularmente valora la comprensión de la asignación de hipótesis estadísticas, midiendo la capacidad del alumno en la elección de la hipótesis nula adecuada partiendo de un contexto de aplicación. La respuesta correcta es la (a), puesto que la hipótesis nula es la contraria a la que desea probar el investigador (Chow, 1996). El distractor (b) lo elegiría un estudiante que confunde las hipótesis nula y alternativa, confusión descrita por Vallecillos (1994). El distractor (b) y (c) lo elegiría el estudiante que no comprende la importancia de no disponer de información previa.

Ítem 5 ¿Cuál de las siguientes hipótesis está bien formulada?

- a) $H_0: \mu_1 - \mu_2 = 3$ $H_1: \mu_1 - \mu_2 \neq 4$
- b) $H_0: \mu = 10$ $H_1: \mu \geq 10$
- c) $H_0: \sigma = 2$ $H_1: \sigma \neq 2$
- d) $H_0: P = 0,55$ $H_1: P \leq 0,5$

Justificación

El ítem, al igual que el anterior, también mide el primer atributo latente, específicamente, evalúa el conocimiento de las reglas para establecer la hipótesis nula y la hipótesis alternativa. La respuesta correcta es la (c), ya que los conjuntos donde se define el parámetro para cada una de las hipótesis deben ser excluyentes y cubrir el espacio paramétrico (Vallecillos, 1994). Estas son las condiciones exigidas al plantear correctamente las hipótesis, que no se cumplen en los otros distractores, es decir, en los distractores las hipótesis no son complementarias o excluyentes.

Ítem 6 Supongamos que conocemos la verdad absoluta sobre la eficacia de dos tratamientos (A y B), y sabemos que existen diferencias en la efectividad de ambos para curar el H1N1. Un investigador que realice un estudio y parta de la hipótesis “*no existen diferencias en la efectividad de los tratamientos A y B para curar el H1N1*” cometerá un error tipo II cuando:

- a) Concluya que A y B no son efectivos para curar el H1N1
- b) Concluya que A y B no difieren en su efectividad para curar el H1N1**
- c) Concluya que A y B difieren en su efectividad para curar el H1N1
- d) Ninguna de las Anteriores

Justificación

Este ítem evalúa si el alumno es capaz de diferenciar entre los errores tipo I y tipo II, y analiza la capacidad de diferenciar entre hipótesis estadísticas e hipótesis de investigación (Chow, 1996). La respuesta correcta es la (b), ya que se cometerá un error de tipo II si decide que no existe diferencia entre los dos tratamientos para curar la depresión, cuando en realidad si existen. El alumno que elige el distractor (a) no logra interpretar el enunciado, ya que el problema afirma se conoce que ambos medicamentos han sido probados, resultando efectivo. Quien elige el distractor (c) estaría confundiendo el error de tipo I con el de tipo II. Finalmente, el distractor (d) será elegido si el estudiante no interpreta el enunciado y además está confundido entre el error tipo I y tipo II.

Ítem 7 Cuando realizamos un test de hipótesis, la regla de decisión nos lleva a rechazar la hipótesis nula siempre que:

- a) El estadístico de prueba caiga en la región de rechazo
- b) El p-valor sea menor que el valor de alfa
- c) a) y b) son correctas**
- d) Ninguna de las Anteriores

Justificación

Finalmente, este ítem evalúa la comprensión en la formulación de hipótesis y las condiciones en las que se tomará la decisión de rechazar la hipótesis nula en base a la región crítica. La respuesta correcta es la (c), puesto que la hipótesis nula se rechaza tanto si se produce la situación ilustrada en (a) como en (b). El distractor (a) la respuesta es verdadera, pero no es el único argumento para rechazar la hipótesis nula en un contraste, por eso quien lo elige no conoce la posibilidad que brinda el distractor (b). En el distractor (b) los estudiantes no asocian el cálculo de una probabilidad con la regla de decisión, no logran comprender ni son capaces de relacionar estos conceptos. Para el distractor (d), los estudiantes no comprenden cuando se ha de rechazar la hipótesis nula.

3.5 Matriz Q

La siguiente tabla es la Matriz-Q de la investigación, las coordenadas de la matriz poseen orden binario, donde 1 implica la asociación entre el ítem y la habilidad o variable latente, en cambio un 0 implica la no relación entre el ítem y la variable latente.

Variables Latentes			
	Hipótesis	Región de Rechazo	Errores tipos I y tipo II
Ítem 1	0	1	0
Ítem 2	0	0	1
Ítem 3	0	0	1
Ítem 4	1	0	0
Ítem 5	1	0	0
Ítem 6	1	0	1
Ítem 7	0	1	1

Tabla 5: Q-Matriz

Cabe mencionar que, para el tratamiento de los DCMs, los datos o patrones de respuesta faltantes en la muestra obtenida se reemplazaran por ceros debido a la naturaleza binaria de estos. Los patrones de respuesta dan paso a los patrones de clase y estos últimos están definidos por una permutación con repetición relacionada a la cantidad de variables latentes que se desea medir, la cual está dada por $2^3 = 8$ y son los siguientes:

0 0 0
1 0 0
0 1 0
0 0 1
1 1 0
1 0 1
0 1 1
1 1 1

Tabla 6: Posibles Patrones de Clase

Capítulo IV Resultados

En el siguiente capítulo se presentan, analizan e interpretan los datos desde un punto de vista descriptivo considerando la confiabilidad del test y la cantidad de preguntas correctas, continuando con el estudio de la discriminación, dificultad y azar del cuestionario según el modelo 3-PL, culminando con la modelación diagnóstica Loglineal de los parámetros presentados en el capítulo anterior.

4.1 Descripción de la Muestra

Los 300 individuos encuestados manifestaron su género, edad y condición académica para dar lugar al siguiente estudio descriptivo y la comparación entre algunas de estas variables con su porcentaje de acierto en los ítems o sus perfiles de clase y patrones de respuesta. Es importa recalcar que para el estudio no importará el nivel socioeconómico o perfil biológico, sólo importará que las capacidades cognitivas del encuestado sean normales.

El 57,67% de los encuestados pertenecen al género masculino y el 42,33% restante corresponde al género femenino, además, la edad promedio de los encuestados es cercana a 28 años, con una desviación estándar de 6,4 años. En la Ilustración 8 se aprecia que la amplitud intercuartil está comprendida entre los 23 y 32 años, con dos casos atípicos los cuales muestran 52 y 53 años respectivamente.

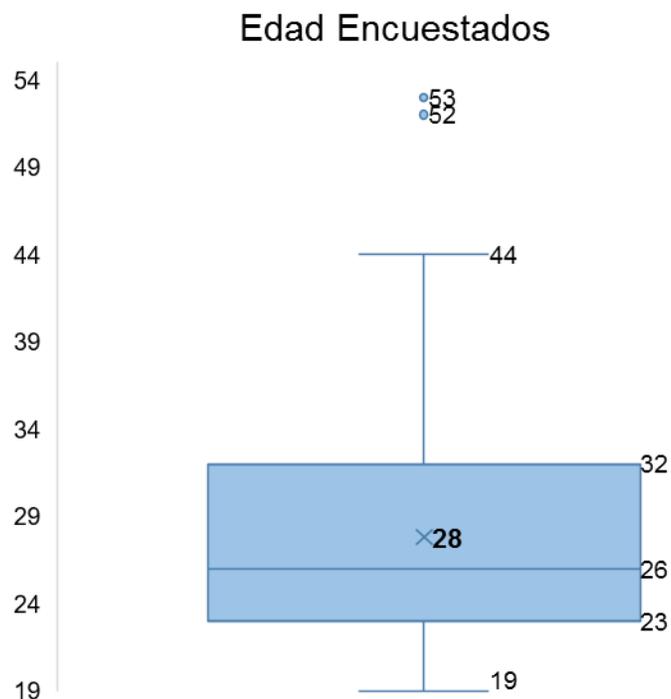


Ilustración 8: Gráfico de Caja para Edad Encuestados

Finalmente, en relación a la condición académica de los encuestados, el 41,33% está cursando una carrera de pregrado y el 58,67% está compuesto por docentes y una pequeña parte de estos últimos, correspondiente al 19,89% (11,67% del total) está cursando un postgrado en alguna casa de estudio universitario. En la Ilustración 9 se muestra la Condición Académica de los Encuestados.

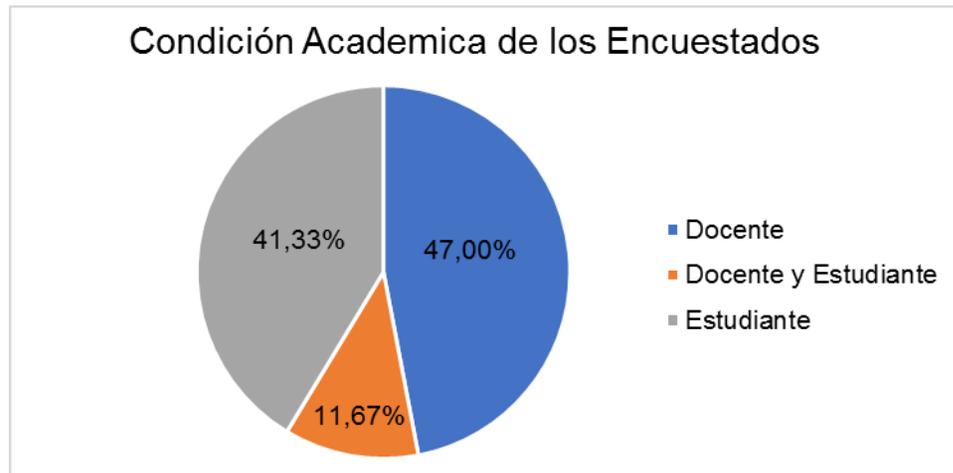


Ilustración 9: Condición Académica de los Encuestados

Para el análisis posterior, además de analizar los dominios de sus variables latentes, se estudiará la cantidad de preguntas correctas por género y también la frecuencia de respuestas correctas según la condición académica del encuestado.

4.2 Análisis Descriptivo

El análisis comenzará con el estudio de la confiabilidad del test según el Alpha de Cronbach, luego se hará una descripción porcentual de cada ítem respecto a las alternativas elegidas por los encuestados, finalizando con el estudio de la cantidad de respuestas correctas y su comparación con respecto al género y a la condición académica de los encuestados.

Es importante recalcar que la muestra no presenta igual cantidad de individuos por género y lo mismo ocurre con la cantidad de personas según condición académica, donde sólo un reducido porcentaje corresponde a Docentes y Estudiantes (ver Ilustración 9), por lo tanto, en la comparación se utilizará la frecuencia relativa en ambos casos.

4.2.1 Confiabilidad del Instrumento

Los datos fueron procesados a través del software R Commander, obteniendo los siguientes resultados:

N	300
Alpha de Cronbach	0,593

	Alpha si el ítem es removido	Ítem-total correlación sin el ítem	Ítem-total correlación	Proporción respuesta correcta	Discriminación del Ítem
Ítem 1	0,509	0,446	0,636	0,71	0,62
Ítem 2	0,591	0,202	0,439	0,297	0,47
Ítem 3	0,553	0,317	0,555	0,45	0,66
Ítem 4	0,623	0,109	0,367	0,653	0,43
Ítem 5	0,526	0,395	0,601	0,687	0,63
Ítem 6	0,6	0,183	0,441	0,583	0,5
Ítem 7	0,457	0,562	0,735	0,583	0,85

Tabla 7: Confiabilidad según Alpha de Cronbach

Debido a la escasa cantidad de ítem, se espera que su confiabilidad no sea alta, pese a esto, Templin asegura que los Modelos de Clasificación Diagnóstica proporcionan un alto nivel de confiabilidad para sus estimaciones con respecto a los modelos tradicionales (Templin & Bradshaw, 2013).

4.2.2 Respuesta según Ítem

A continuación, se muestra la distribución de las frecuencias relativas de cada alternativa según ítem donde las alternativas correctas están ennegrecidas:

	Alternativa			
	a	b	c	d
Ítem 1	71,33%	3,39%	7,89%	17,39%
Ítem 2	27,46%	30,33%	25,1%	17,11%
Ítem 3	18,1%	46,00%	12,67%	23,23%
Ítem 4	66,67%	23,61%	3,74%	5,98%
Ítem 5	2,5%	24,39%	70,33%	2,78%
Ítem 6	15,3%	60,33%	17,34%	7,03%
Ítem 7	19,1%	14,06%	60,67%	6,17%

Tabla 8: Respuestas según Ítem

Todos los ítems muestran la mayor frecuencia en la alternativa correcta, pese a esto, sólo cinco poseen una alta frecuencia en él. El ítem 1 es el que muestra una mayor frecuencia, por lo tanto, se espera que sea el más fácil. Luego el segundo ítem con mayor frecuencia en la alternativa correcta es el 5, seguido por el ítem 4, 7 y el último que muestra un porcentaje mayor al 60% de acierto es el ítem 6. En cambio, los ítems 2 y 3 son respondidos correctamente por menos de la mitad de los encuestados.

4.2.3 Encuestados según Cantidad de Preguntas Correctas

Inspirados en la tabla anterior, se presenta la distribución de encuestados según la cantidad de preguntas correctas.

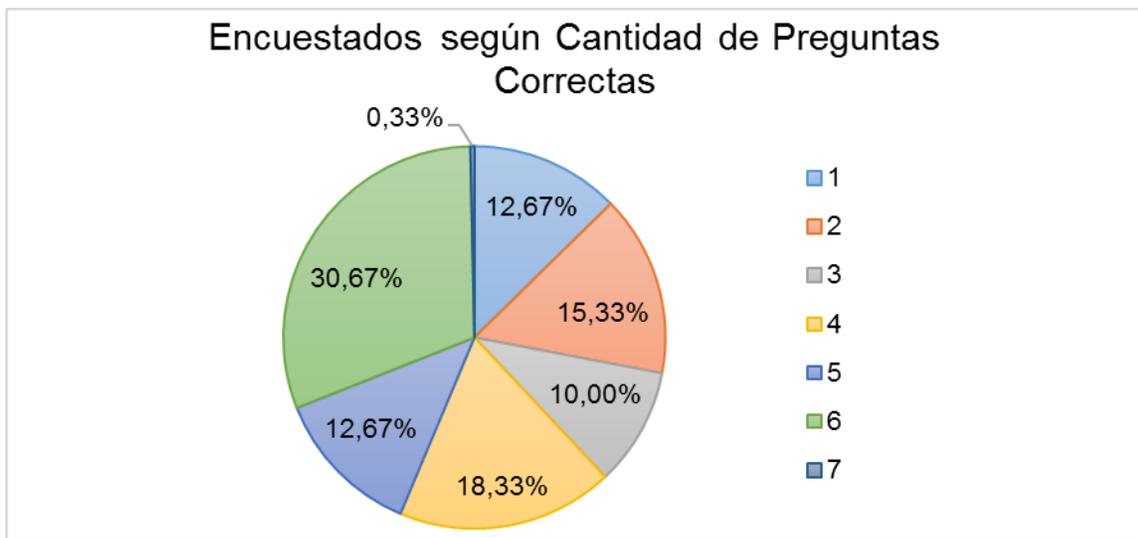


Ilustración 10: Encuestados según Cantidad de Preguntas Correctas

Las menores frecuencias se obtienen en los casos extremos, es más, no existen encuestados con cero respuestas correctas y se presenta un ínfimo porcentaje de encuestados con todas las alternativas correctas. En cambio, la mayor frecuencia corresponde a individuos con 6 respuestas correctas, finalmente, la mediana pertenece a los encuestados con 4 aciertos en el cuestionario.

4.2.4 Comparación de Preguntas Correctas según Género

A continuación, se compara la cantidad de preguntas correctas según el género del encuestado.

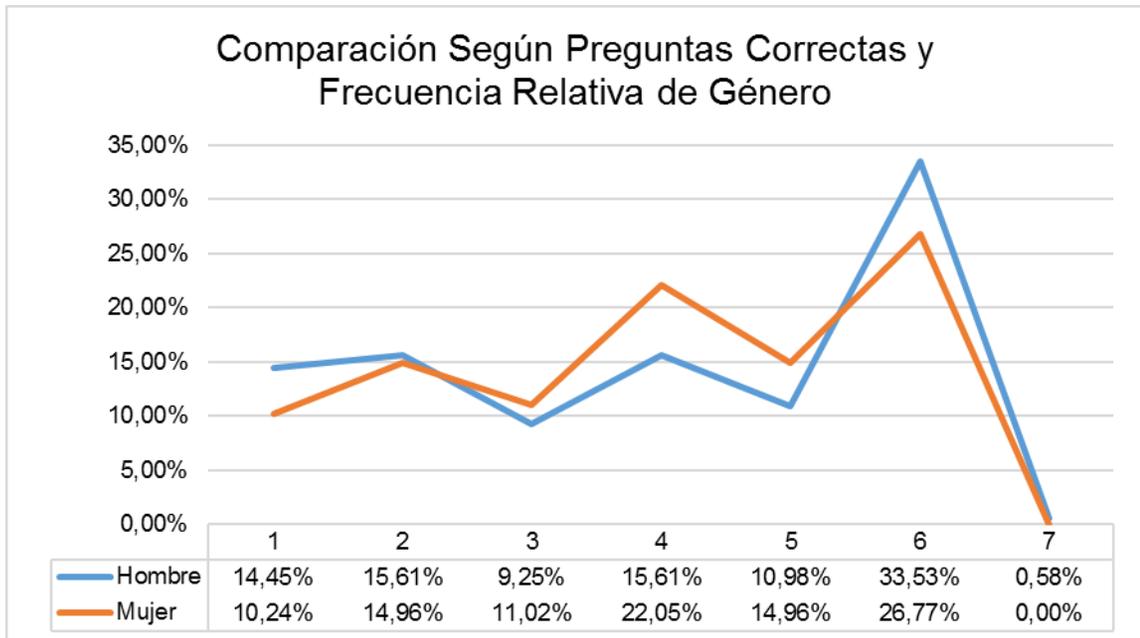


Ilustración 11: Comparación según Preguntas Correctas y Frecuencia Relativa del Género

No existe gran diferencia entre la cantidad de preguntas correctas y el género del encuestado, ambos poseen una mediana igual a cuatro y su moda está marcada por las seis preguntas correctas.

Existe una pequeña diferencia en los extremos de la distribución, cuando se analiza el porcentaje de encuestados con sólo 1 o sólo 2 preguntas correctas donde existe un mayor porcentaje de hombres. De manera semejante, existe una mayor cantidad de hombres con todas las respuestas correctas o sólo 6 ítems correctos.

4.2.5 Comparación de Preguntas Correctas según Situación Académica

En la siguiente imagen se presenta la distribución de respuestas correctas y su relación con la situación académica de los encuestados.

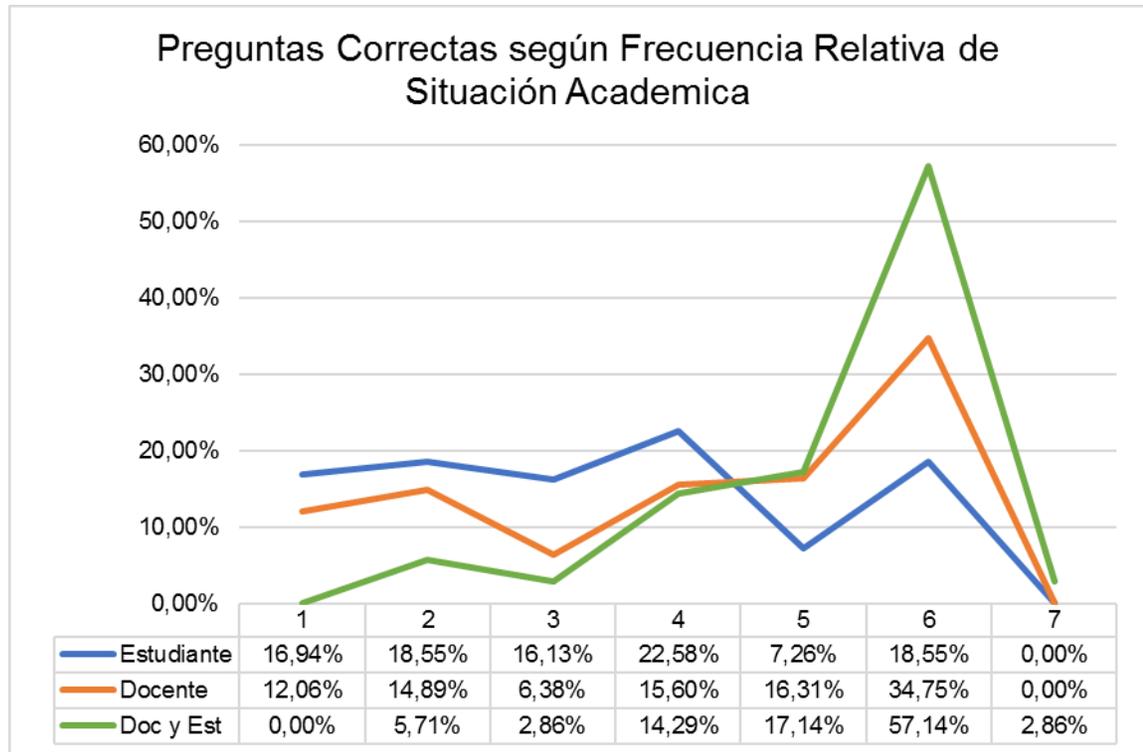


Ilustración 12: Preguntas Correctas según Frecuencia Relativa de Situación Académica

Como muestra la Ilustración 12, existe una gran diferencia entre las diversas situaciones académicas y la cantidad de preguntas correctas. Como primer análisis, la mediana de las respuestas correctas por estudiantes de pregrado es 3, en cambio, para docentes se sitúa en 5 y los docentes que participan en un post-título poseen una mediana igual a 6 respuestas correctas.

Una situación semejante ocurre cuando se desea estudiar la moda de cada situación académica, por una parte, los estudiantes poseen una moda de 4 preguntas correctas en cambio los docentes, tanto los que siguen un post-título como los que no, muestran una moda de 6 preguntas correctas.

Finalmente, este último grupo es el único que posee individuos con todas sus respuestas correctas y su gráfica muestra un crecimiento considerable, en detalle, son el único grupo donde todos sus integrantes poseen por lo menos 2 preguntas correctas, asimismo, este grupo muestra el mayor porcentaje de individuos con 5, 6 y 7 preguntas correctas, y más del 55% contestó correctamente seis preguntas.

4.2.6 Comparación Ítem Correcto según Situación Académica

El siguiente gráfico, muestra la relación entre, el porcentaje de individuos que responden acertadamente cada ítem según la situación académica a la cual pertenecen.

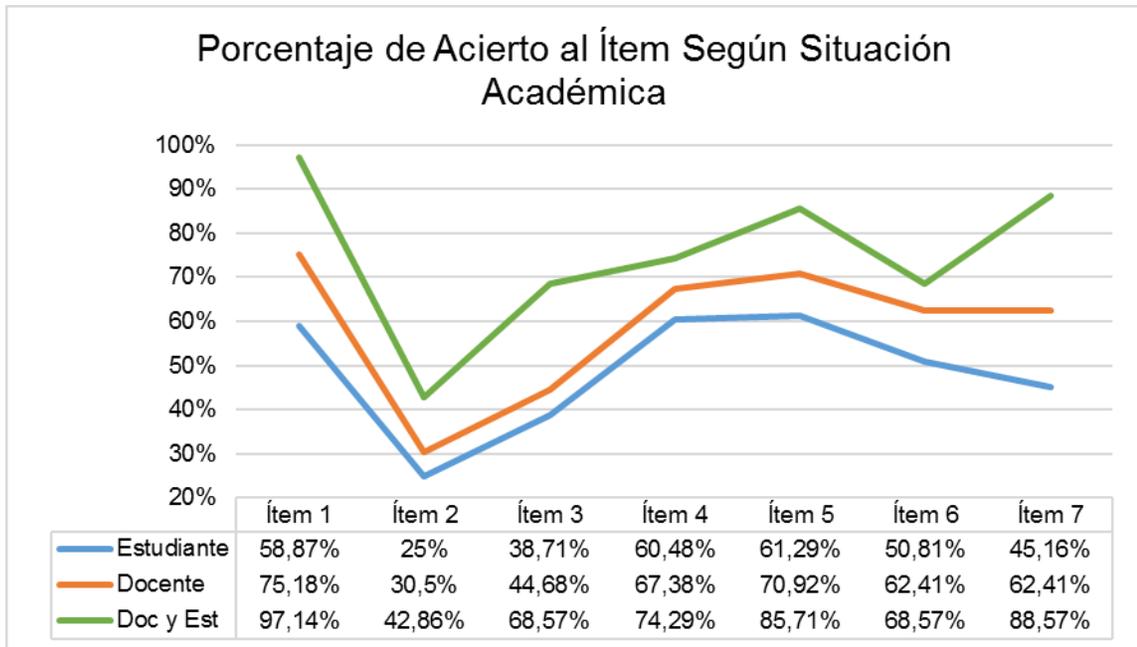


Ilustración 13: Porcentaje de Acierto según Situación Académica

Los docentes que siguen sus estudios son los que presentan mejores resultados y en la Ilustración 13 se verifica esto, ya que este grupo muestra los mayores porcentajes de respuestas acertadas en cada ítem, en otras palabras, este grupo tiene mayor porcentaje de individuos con respuestas correctas en cada ítem.

En contraparte al grupo de docentes que siguen un post-grado, están los estudiantes de pregrado que exhiben los peores resultados en cada ítem, el peor resultado de este grupo está en la segunda pregunta donde sólo el 25% de los estudiantes responden correctamente el ítem.

Finalmente, no existe grupo donde el 100% de sus integrantes tenga resuelto correctamente algún ítem.

4.3 Análisis según Modelo Logístico de Tres Parámetros

Para estudiar la discriminación, dificultad y azar de cada ítem en conjunto con los patrones de respuesta y el nivel de habilidad relacionada a cada uno se utilizó el software R Commander con la librería IRTShyni, obteniendo los siguientes resultados:

	Discriminación	Dificultad	Azar
Ítem 1	5.913	-0.0508	3.84e-01
Ítem 2	0.644	1.4605	3.04e-05
Ítem 3	1.293	0.2115	1.39e-06
Ítem 4	1.855	0.8985	5.34e-01
Ítem 5	1.280	-0.8011	7.44e-04
Ítem 6	0.588	-0.6178	3.04e-04
Ítem 7	3.221	-0.2272	1.26e-04

Tabla 9: Análisis según Modelo 3-PL

Gráficamente, las curvas características según el modelo 3-PL para cada ítem son:

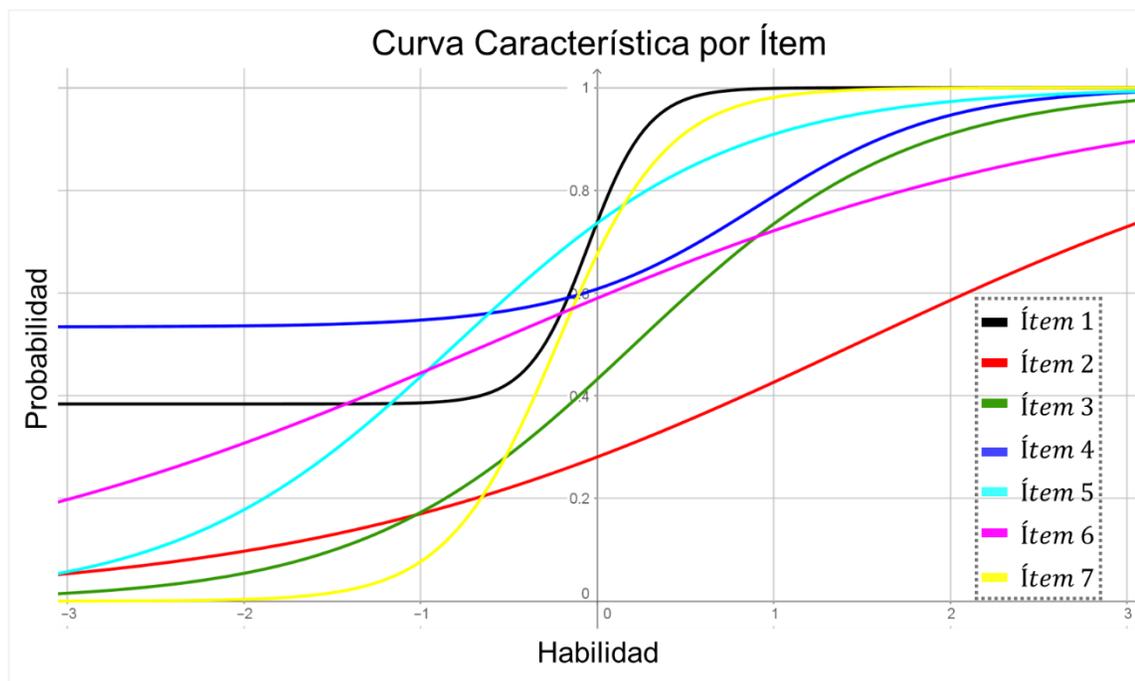


Ilustración 14: Análisis según Modelo 3-PL

El análisis y comparación de estos indicadores se llevarán a cabo en el capítulo 4.4.3 donde se confrontarán con los parámetros del modelo de clasificación diagnóstica Loglineal.

Se observan 82 patrones de respuesta, con una moda de 38 repeticiones iguales a 1011111. Bajo el modelo logístico 3-PL cada patrón está relacionado con un nivel de habilidad θ y el error estándar, estos datos se presentan a continuación (Revisar Anexo 3 para visualizar la probabilidad de acierto al ítem de cada patrón según modelo 3-PL):

Patrón	Theta	ES	Patrón	Theta	ES	Patrón	Theta	ES
0000001	-0,405	0,354	0101010	-0,725	0,512	1010111	0,505	0,494
0000010	-0,96	0,589	0101100	-0,569	0,439	1011000	-0,564	0,75
0000100	-0,752	0,507	0101110	-0,468	0,392	1011001	0,268	0,434
0000110	-0,619	0,448	0101111	-0,132	0,283	1011011	0,395	0,498
0000111	-0,216	0,291	0110010	-0,502	0,396	1011100	-0,037	0,407
0001000	-1,159	0,675	0110111	-0,065	0,272	1011101	0,604	0,603
0001001	-0,371	0,35	0111111	-0,03	0,28	1011110	0,053	0,383
0001010	-0,922	0,595	1000000	-1,183	0,68	1011111	0,848	0,682
0001100	-0,713	0,506	1000001	-0,01	0,377	1100011	0,157	0,373
0001101	-0,236	0,303	1000010	-0,931	0,632	1100101	0,26	0,402
0001110	-0,58	0,445	1000011	0,07	0,365	1100110	-0,126	0,45
0001111	-0,185	0,291	1000100	-0,659	0,641	1100111	0,364	0,441
0010000	-0,749	0,505	1000101	0,164	0,374	1101010	-0,597	0,722
0010001	-0,266	0,305	1000110	-0,35	0,794	1101100	-0,174	0,547
0010010	-0,616	0,447	1000111	0,251	0,399	1101101	0,406	0,503
0010011	-0,214	0,291	1001000	-1,149	0,694	1101110	-0,048	0,413
0010100	-0,495	0,393	1001001	0,062	0,383	1101111	0,582	0,593
0010101	-0,158	0,28	1001010	-0,884	0,654	1110010	-0,124	0,447
0010110	-0,412	0,357	1001011	0,149	0,392	1110011	0,366	0,442
0011000	-0,709	0,505	1001100	-0,571	0,743	1110110	0,069	0,365
0011100	-0,459	0,388	1001101	0,266	0,433	1110111	0,68	0,547
0011101	-0,127	0,282	1001110	-0,192	0,579	1111001	0,409	0,505
0011111	-0,081	0,279	1001111	0,391	0,496	1111011	0,586	0,595
0100001	-0,331	0,326	1010000	-0,654	0,643	1111100	0,061	0,383
0100011	-0,272	0,306	1010011	0,253	0,399	1111101	0,875	0,688
0100100	-0,608	0,443	1010101	0,376	0,445	1111110	0,149	0,392
0100111	-0,163	0,281	1010110	-0,019	0,38	1111111	1,172	0,731
0101000	-0,902	0,587						

Tabla 10: Patrones de Respuesta y Habilidad θ

4.3.1 Función Informativa Modelo 3-PL

En la siguiente ilustración se presenta la gráfica de la Función Informativa según el modelo 3-PL, la cual tiene la siguiente expresión algebraica:

$$I(\theta) = \left(a_i \frac{1 - P_i(\theta)}{P_i(\theta)} \right) \left(\frac{P_i(\theta) - c_i}{1 - c_i} \right)^2$$

Esta función refleja el nivel de habilidad donde se obtiene la mayor información del cuestionario, en ella se espera que el punto máximo se encuentre en el punto con abscisa cercana a cero. Con esto se entiende que el test posee una dificultad aceptable; del mismo modo, si el punto máximo se encuentra en extremos negativos o positivos se entiende que el test presenta una dificultad muy baja o muy alta respectivamente.

La representación gráfica de la función informativa para el cuestionario es:

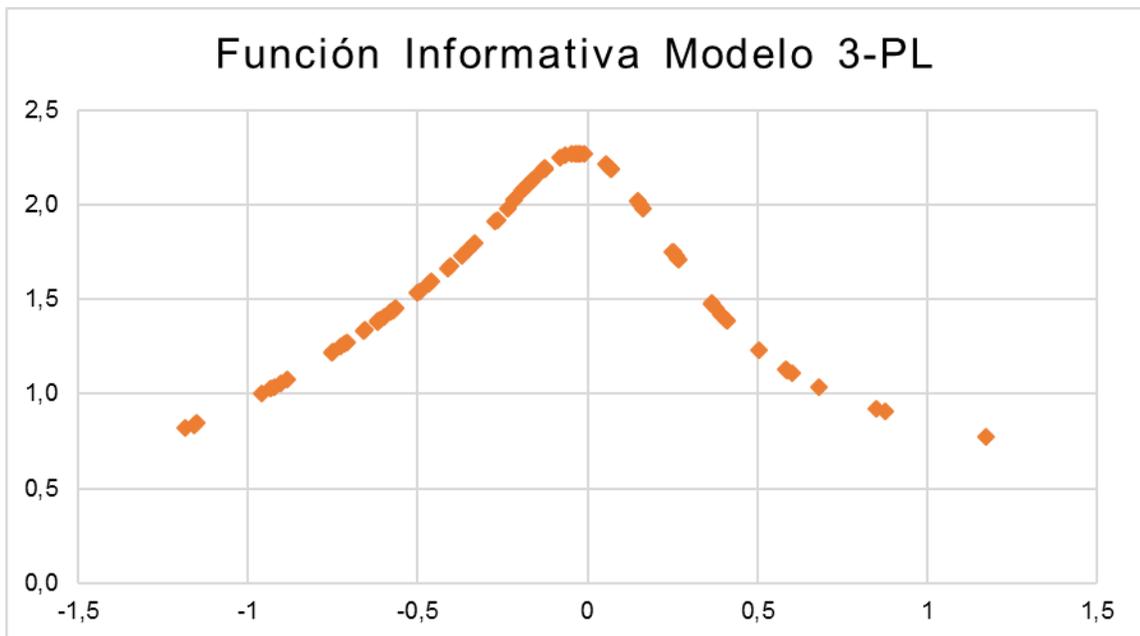


Ilustración 15: Función Informativa Modelo 3-PL

4.4 Análisis según Modelo de Clasificación Diagnóstica Loglineal

Para implementar el modelo de clasificación diagnóstica Loglineal se ocupan los softwares SAS y Mplus. El primer ayudará a generar el código que se ejecutará posteriormente en Mplus, obteniendo lo archivos que darán lugar a interpretaciones acerca del modelo, sus parámetros y la probabilidad de acierto al ítem para cada perfil de encuestados. El desarrollo de estos códigos está en Anexo 1 y Anexo 2.

La matriz Q que se definirá e implementará debe contener el ítem, los atributos latentes y las interacciones entre ellos, es decir:

	Hipótesis	Región de Rechazo	Errores tipos I y tipo II	Interacciones
<i>Ítem 1</i>	0	1	0	1
<i>Ítem 2</i>	0	0	1	1
<i>Ítem 3</i>	0	0	1	1
<i>Ítem 4</i>	1	0	0	1
<i>Ítem 5</i>	1	0	0	1
<i>Ítem 6</i>	1	0	1	2
<i>Ítem 7</i>	0	1	1	2

Tabla 11: Matriz Q para SAS

4.4.1 Recuento de Clases

El recuento final de clases y proporciones para las clases latentes basadas en el modelo de clasificación diagnóstica Loglineal es:

Clases Latentes	Proporción
[000]	20,04%
[100]	14,04%
[010]	2,39%
[001]	4,23%
[110]	2,02%
[101]	4,57%
[011]	7,5%
[111]	45,22%

Tabla 12: Clases Latentes de los Encuestados

En las Clases latentes el primer dígito trata el dominio o no dominio del primer atributo y así sucesivamente hasta el tercer dígito. Con esta información podemos estimar la probabilidad de elegir un sujeto al azar que haya realizado el test y posea un determinado perfil de atributos. La representación gráfica de la Tabla 12 es:

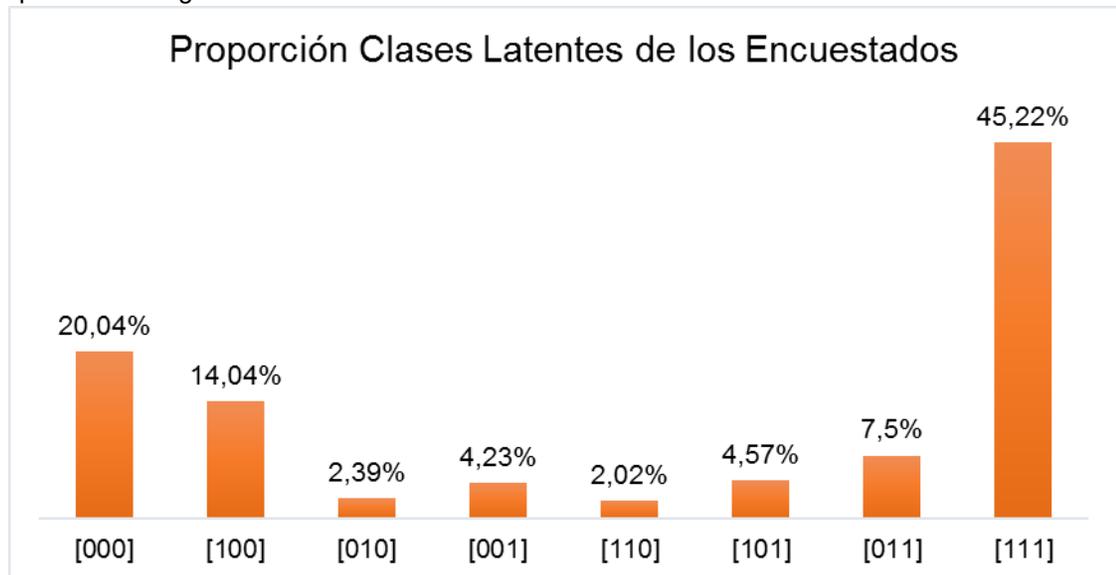


Ilustración 16: Clases Latentes de los Encuestados

Como muestra la Tabla 12 e Ilustración 16, gran parte de los encuestados posee un perfil [111], es decir, dominan las habilidades ligadas a la *Hipótesis*, *Región de Rechazo* y *Error tipo I y tipo II* (ver capítulo 3.2.1 para su definición), así también es importante mencionar que un gran porcentaje de encuestados no dominan ninguna variable latente (perfil [000]) o bien dominan sólo el primer atributo latente (perfil [100]). Finalmente, es trascendental informar que cada patrón de respuesta está relacionado a una clase latente, para ver el detalle se invita a revisar el Anexo 4).

4.4.2 Parámetros Estimados

En este apartado se procede a mostrar los parámetros estimados según el LCDM, además del estadístico, el error estándar y el p-valor ligado al parámetro estimado.

<i>Parámetros</i>	<i>Estimación</i>	<i>Error Estándar</i>	<i>Estadístico</i>	<i>P-Valor</i>
L1_0	-0,737	0,247	-2,984	0,003
L1_11	13,631	2,063	6,606	0
L2_0	-1,943	0,34	-5,719	0
L2_11	1,554	0,389	3,996	0
L3_0	-1,89	0,352	-5,375	0
L3_12	2,507	0,392	6,401	0
L4_0	0,064	0,207	0,308	0,758
L4_13	1,209	0,358	3,372	0,001
L5_0	-0,336	0,23	-1,461	0,144
L5_13	3,049	1,353	2,254	0,024
L6_0	-0,477	0,225	-2,123	0,034
L6_11	1,317	0,301	4,382	0
L6_13	1,317	0,301	4,382	0
L6_213	-1,317	0,301	-4,382	0
L7_0	-2,417	0,684	-3,532	0
L7_12	0	0	100	0
L7_13	4,583	0,764	6,002	0
L7_223	0	0	-42,994	0

Tabla 13: Parámetros Estimados

4.4.3 Modelo Logit, Probabilidad y Análisis según Modelo 3-PL

Con la información de la Tabla 13 se puede construir el Logit y la probabilidad asociada a él para cada ítem, además como se mencionó anteriormente, confrontaremos la discriminación, la dificultad y el azar de cada uno de estos ítems.

Ítem 1

α_1	α_2	α_3	Modelo	Logit	Probabilidad
0	0	0	$-0.737 + 13.631 \cdot (0)$	-0.737	0.324
0	1	0	$-0.737 + 13.631 \cdot (1)$	12.894	1

Tabla 14: Logit y Función de Probabilidad Ítem 1

Este ítem presenta una alta discriminación para los encuestados con un nivel de habilidad diferente a $\theta = -0.0508$, es más, es el ítem con mayor discriminación dentro del cuestionario. Su dificultad es media, sin embargo, si el estudiante domina el atributo medido su probabilidad de acierto es elevada. Finalmente, la mínima probabilidad de responder correctamente el ítem es cercana a 0,3 tanto en el LCDM como en el modelo 3-PL.

Ítem 2

α_1	α_2	α_3	Modelo	Logit	Probabilidad
0	0	0	$-1.943 + 1.554 \cdot (0)$	-1.943	0.125
0	0	1	$-1.943 + 1.554 \cdot (1)$	-0.389	0,404

Tabla 15: Logit y Función de Probabilidad Ítem 2

El ítem muestra una alta dificultad, particularmente es el más difícil del cuestionario lo cual puede relacionarse con la baja probabilidad de acierto de los encuestados que si dominan el atributo medido. Por otra parte, el ítem presenta una baja discriminación y la mínima probabilidad de responder correctamente es 0,0000304 en el modelo 3-PL y 0,125 en el LCDM.

Ítem 3

α_1	α_2	α_3	Modelo	Logit	Probabilidad
0	0	0	$-1.89 + 2.507 \cdot (0)$	-1.89	0.131
0	0	1	$-1.89 + 2.507 \cdot (1)$	0.617	0.65

Tabla 16: Logit y Función de Probabilidad Ítem 3

Este ítem presenta una discriminación y dificultad media. Analizando desde el LCDM posee una baja probabilidad de acierto aun cuando el encuestado domine el atributo necesario para su desarrollo. Finalmente, el azar en el modelo 3-PL es cercana a 0,000001 y en el LCDM la probabilidad de acierto sin dominar el atributo es de 0,131.

Ítem 4

α_1	α_2	α_3	Modelo	Logit	Probabilidad
0	0	0	$0.064 + 1.209 \cdot (0)$	0.064	0.516
1	0	0	$0.064 + 1.209 \cdot (1)$	1.273	0.781

Tabla 17: Logit y Función de Probabilidad Ítem 4

El ítem muestra una dificultad y discriminación media. Pese a esto, su probabilidad de responder correctamente el ítem sin dominar el atributo medido o haciendo uso de otras habilidades no medidas es la más alta del test en ambos modelos, en el modelo 3-PL es de 0,534 y para el LCDM es de 0,516.

Ítem 5

α_1	α_2	α_3	Modelo	Logit	Probabilidad
0	0	0	$-0.336 + 3.049 \cdot (0)$	-0.336	0,417
1	0	0	$-0.336 + 3.049 \cdot (1)$	2.713	0,938

Tabla 18: Logit y Función de Probabilidad Ítem 5

Este ítem presenta una baja dificultad, particularmente es la más baja del test, lo cual podría estar relacionada con la alta probabilidad de acierto al ítem para los encuestados que dominen el atributo medido. Por otra parte, su discriminación es media y su probabilidad de azar es baja en el modelo 3-PL, en cambio en el LCDM la probabilidad de responder correctamente el ítem sin dominar el atributo necesario es cercana a 0,4.

Ítem 6

α_1	α_2	α_3	Modelo	Logit	Probabilidad
0	0	0	$-0.477 + 1.317 \cdot (0) + 1.317 \cdot (0) - 1.317 \cdot (0) \cdot (0)$	-0.477	0,383
1	0	0	$-0.477 + 1.317 \cdot (1) + 1.317 \cdot (0) - 1.317 \cdot (1) \cdot (0)$	0.84	0,698
0	0	1	$-0.477 + 1.317 \cdot (0) + 1.317 \cdot (1) - 1.317 \cdot (0) \cdot (1)$	0.84	0,698
1	0	1	$-0.477 + 1.317 \cdot (1) + 1.317 \cdot (1) - 1.317 \cdot (1) \cdot (1)$	0.84	0,698

Tabla 19: Logit y Función de Probabilidad Ítem 6

El ítem muestra una baja discriminación y dificultad. La probabilidad de responder correctamente sin utilizar el atributo medido es de 0,383 en el LCDM y el azar es de 0,000304 según el modelo 3-PL.

La distribución de la probabilidad de acierto al ítem posee un comportamiento semejante al modelo de clasificación diagnóstica compensatorio (Ver Capítulo 2.2.2), es decir, basta que el encuestado domine uno de los dos atributos para obtener la máxima probabilidad de acierto. Asimismo, el dominio de ambos atributos, no logra una mayor probabilidad de acierto en comparación con sólo utilizar una estrategia para desarrollar el ítem.

Ítem 7

α_1	α_2	α_3	Modelo	Logit	Probabilidad
0	0	0	$-2.417 + 0 \cdot (0) + 4.583 \cdot (0) + 0 \cdot (0) \cdot (0)$	-2.417	0,082
0	1	0	$-2.417 + 0 \cdot (1) + 4.583 \cdot (0) + 0 \cdot (1) \cdot (0)$	-2.417	0,082
0	0	1	$-2.417 + 0 \cdot (0) + 4.583 \cdot (1) + 0 \cdot (0) \cdot (1)$	2.166	0,897
0	1	1	$-2.417 + 0 \cdot (1) + 4.583 \cdot (1) + 0 \cdot (1) \cdot (1)$	2.166	0,897

Tabla 20: Logit y Función de Probabilidad Ítem 7

Este ítem muestra una alta discriminación y una dificultad media sin embargo la probabilidad de acertar al ítem es alta cuando el estudiante domina los atributos medidos. Finalmente, la mínima probabilidad de responder correctamente al ítem es 0,082 en el LCDM y el azar es de 0,000126 según el modelo 3-PL.

Analizando el logit obtenido del estudio bajo el Modelo de Clasificación Diagnóstica Loglineal, se observa que para desarrollar correctamente el ítem 7 sólo se necesita del tercer atributo (Ver Tabla 5), mientras que el segundo no aporta a la probabilidad, es decir, pese a que a priori se planteó como un ítem que mide dos atributos, bajo este análisis, sólo necesita de uno para su correcto desarrollo.

En general el ítem con mayor discriminación es el 1, seguido por el ítem 7 y el tercer lugar lo ocupa el ítem 4. Por otra parte, al analizar la dificultad del test, se muestra que el ítem más complejo es el 2, el segundo lugar lo ocupa el ítem 4 y el tercer lugar es ocupado por el ítem 3. Finalmente, estudiando el azar de los ítems, la pregunta con mayor probabilidad de acierto sin utilizar el atributo medido es la 4, el segundo y tercer lugar lo ocupa el ítem 1 y el ítem 5 respectivamente.

4.4.4 Perfil Asociado a la Probabilidad de Acierto al Ítem

A continuación, se presenta la probabilidad de acierto al ítem para cada perfil de clases latentes, esta probabilidad esta ordenada de forma ascendente en todos los casos:

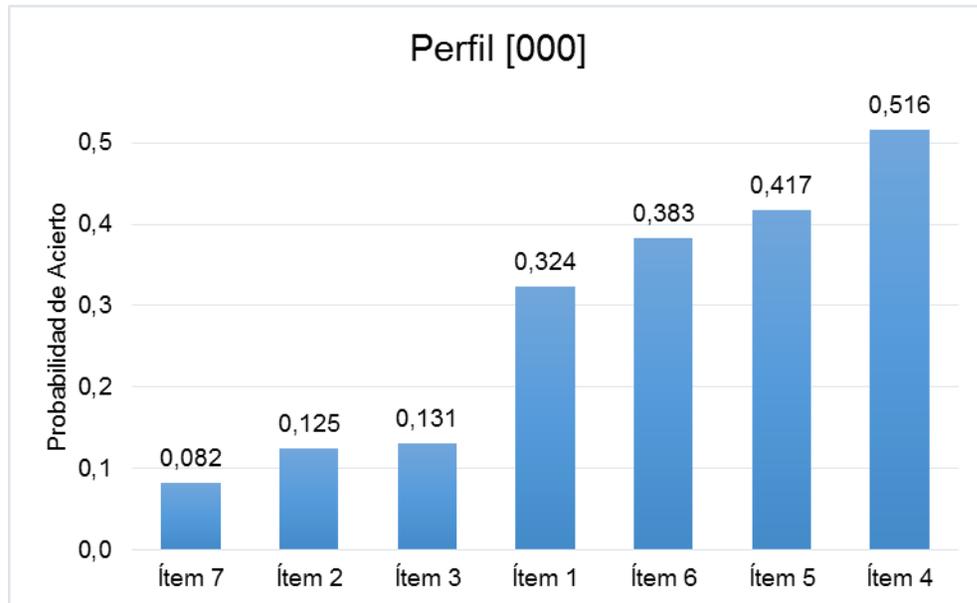


Ilustración 17: Perfil [000]

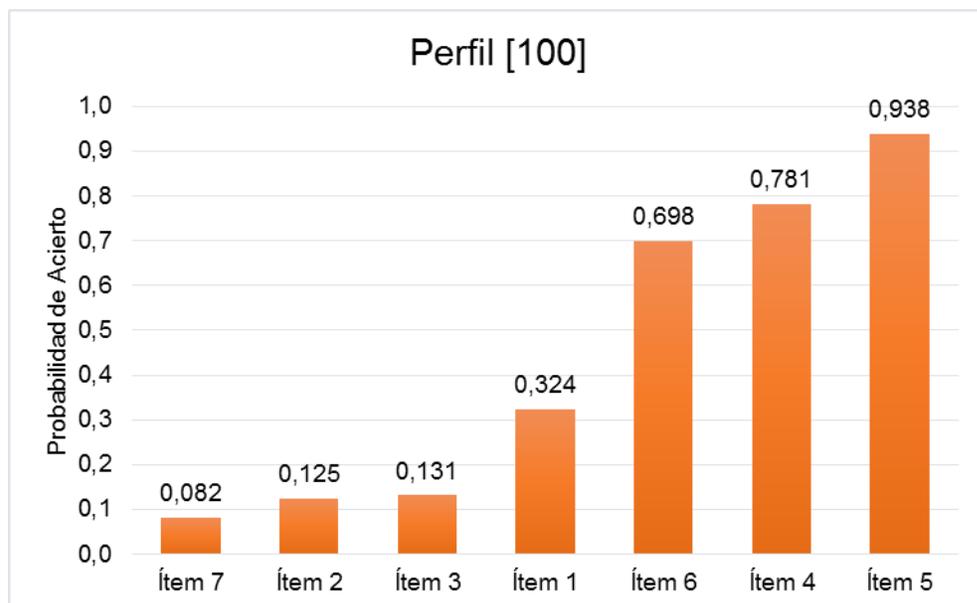


Ilustración 18: Perfil [100]

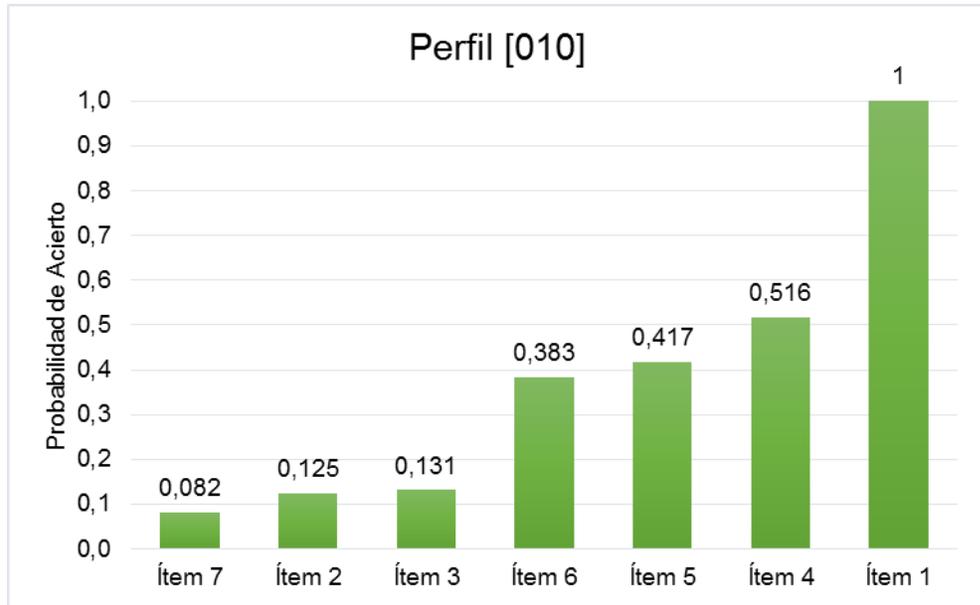


Ilustración 19: Perfil [010]

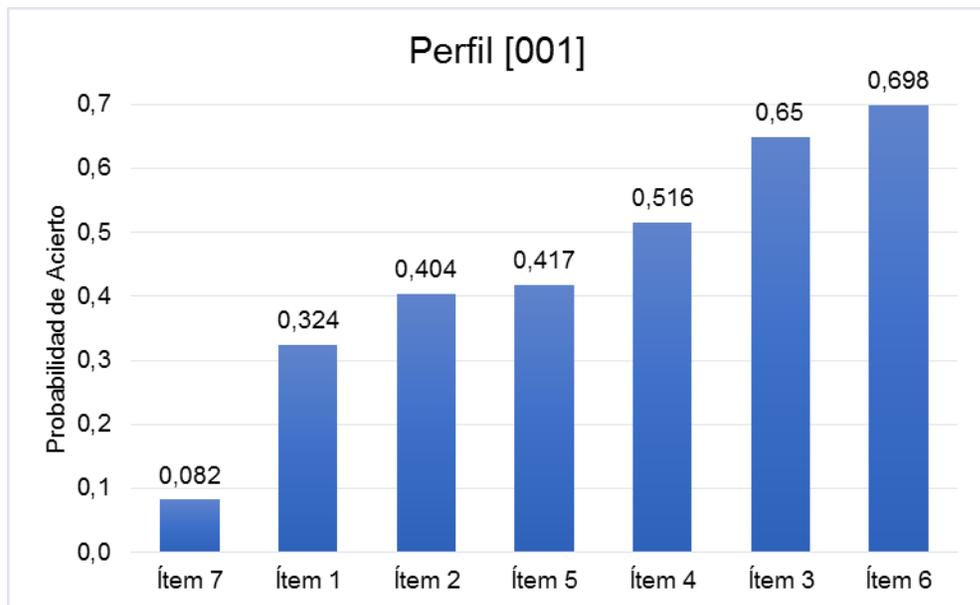


Ilustración 20: Perfil [001]

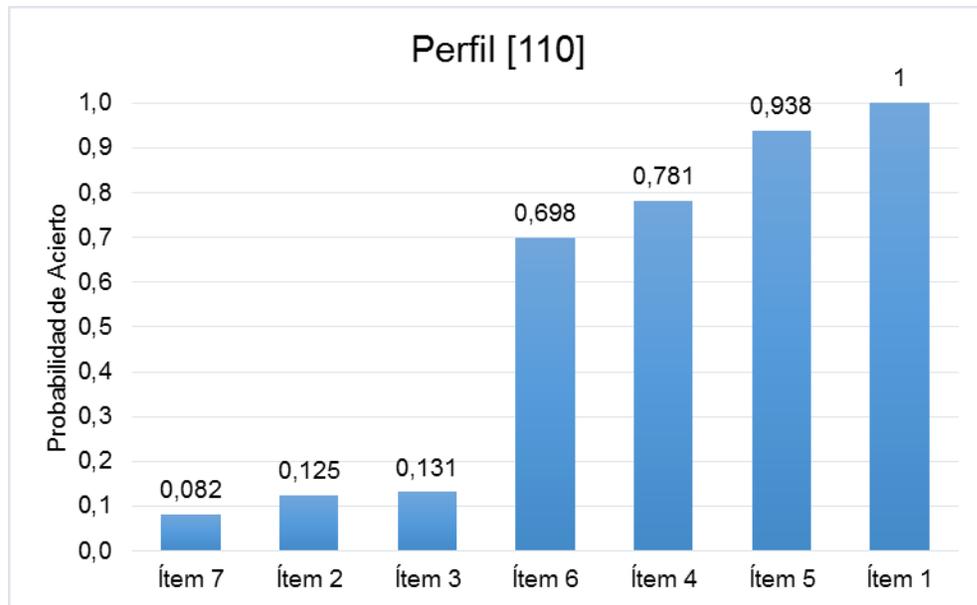


Ilustración 21: Perfil [110]

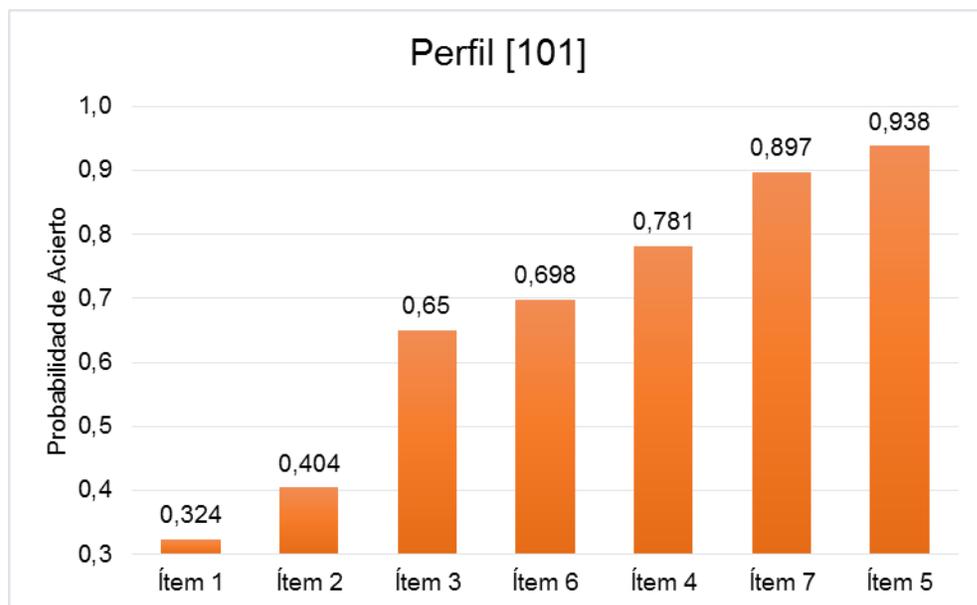


Ilustración 22: Perfil [101]

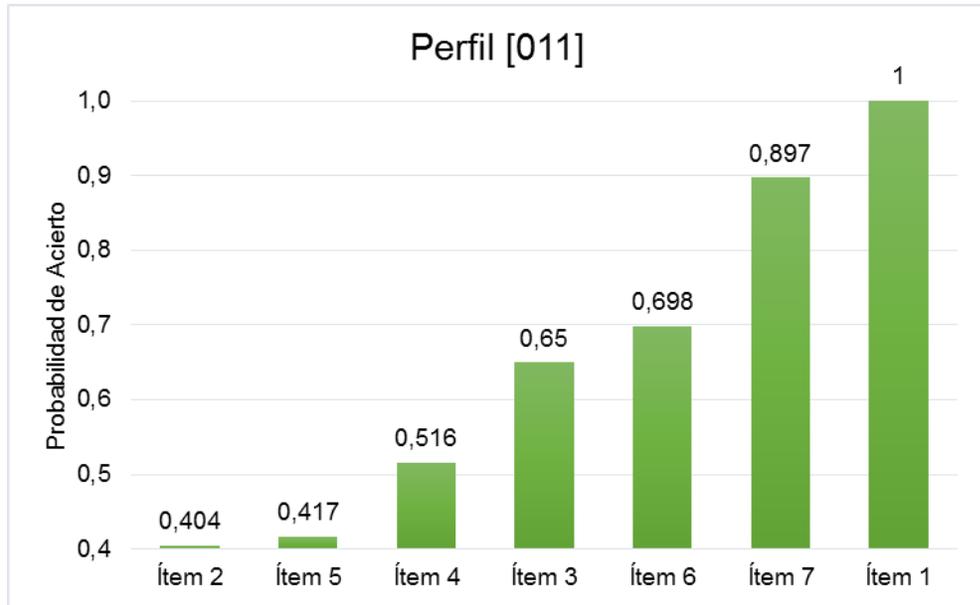


Ilustración 23: Perfil [011]

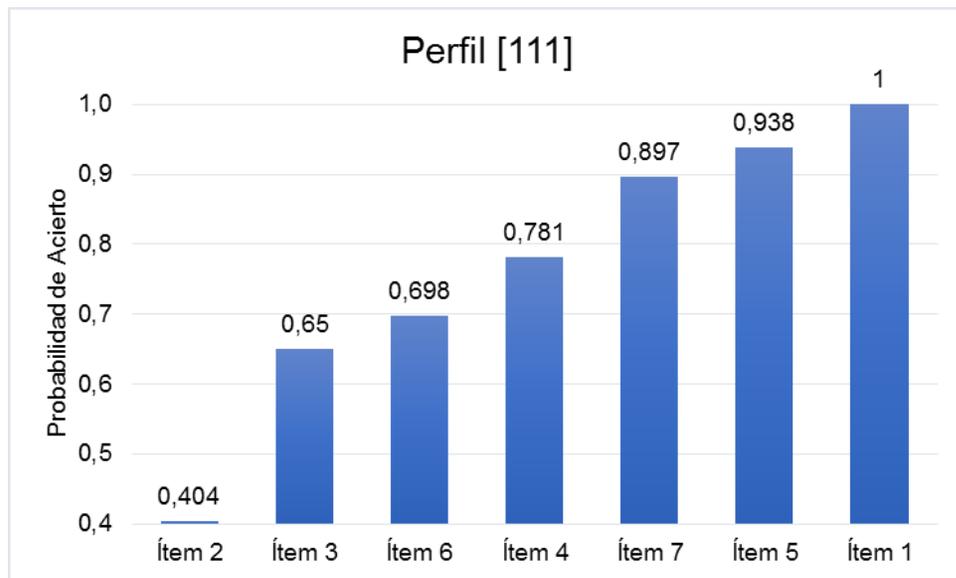


Ilustración 24: Perfil [111]

4.4.5 Comparación de Probabilidad según Perfil

Finalmente, se muestra la comparación entre los distintos perfiles y la probabilidad de responder correctamente cada ítem:

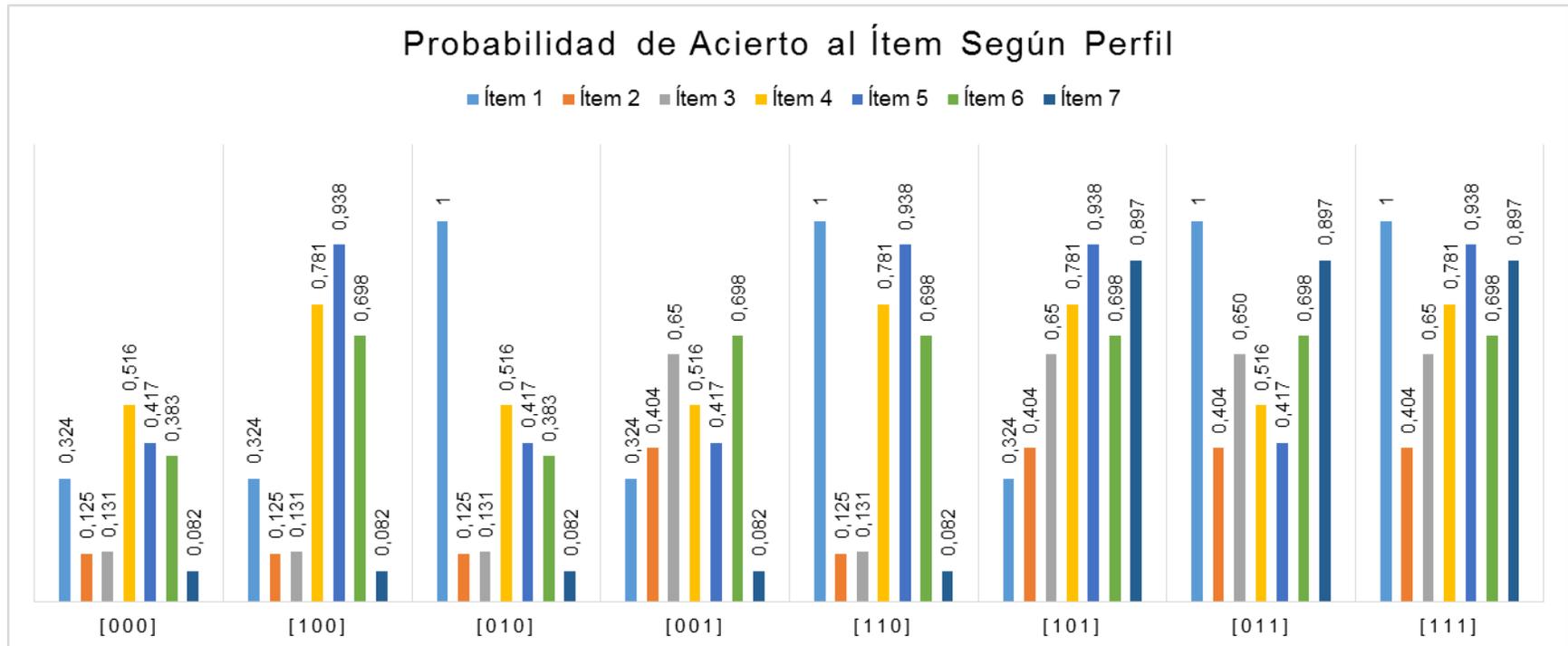


Ilustración 25: Probabilidad de Acierto al Ítem según Perfil

4.4.6 Probabilidad por Atributo

La siguiente tabla presenta los resultados de las probabilidades de dominio de los atributos medidos por las variables latentes y su confiabilidad:

Atributo	Probabilidad	Confiabilidad
Hipótesis	0,658	0,781
Región de Rechazo	0,571	0,823
Error tipo I y tipo II	0,615	0,874

Tabla 21: Probabilidad de Acierto por Atributo

Analizando los resultados con respecto al manejo del contraste de hipótesis, el 65,8% de los encuestados domina las habilidades para formular, plantear, establecer, asignar o diferenciar entre hipótesis nula e hipótesis alternativa. Semejante situación ocurre con el dominio de las habilidades para calcular, diferenciar o interpretar el nivel de significación α , p-valor, la probabilidad del error tipo I y del error tipo II y el cálculo de la potencia del test de hipótesis, donde el 61,5% de los encuestados dominan esta habilidad latente. Finalmente, un menor porcentaje, equivalente al 57,1% de la muestra son capaces de establecer, calcular, discernir o analizar la región de crítica de rechazo o aceptación de la hipótesis nula en un contraste de hipótesis.

Es importante mencionar que estos datos poseen una alta confiabilidad, en definitiva, la menor confiabilidad es mayor al 75% y la máxima es aproximadamente 88%.

Capítulo V Conclusiones

El test de Hipótesis es, innegablemente, una poderosa herramienta de la estadística inferencial, sin embargo su uso no siempre es el adecuado, esto es debido a que su desarrollo requiere de variadas habilidades o atributos latentes, donde es crucial la capacidad para formular, plantear, establecer, asignar o diferenciar entre la hipótesis nula y la alternativa, también es muy importante, la destreza de investigador para determinar el criterio de rechazo o aceptación de la hipótesis nula o establecer, calcular, discernir y analizar la región de crítica de rechazo en un contraste de hipótesis y, finalmente, es de suma importancia la maestría para interpretar, calcular o diferenciar el nivel de significación α , p-valor, la potencia y la probabilidad del error tipo I y II en el desarrollo del test de hipótesis.

El proceso de aprendizaje de estas habilidades es complejo y el docente necesita una herramienta que permita medir el dominio de sus estudiantes en base a un cuestionario, en este contexto, recientemente se están desarrollando los Modelos de Clasificación Diagnóstica y los Modelos Logísticos ya que son una poderosa herramienta en la entrega de información acerca del cuestionario y también permiten clasificar a cada individuo según su nivel de competencia en las categorías de *maestro* o de *no maestro* respecto de algún tipo de constructo o habilidad latente. Sin embargo, estos modelos poseen una dificultad al momento de estimar sus parámetros ya que se requiere de métodos numéricos avanzados los cuales son trabajados a través de softwares avanzados.

No obstante, en las ciencias de la conducta como es la pedagogía, es innegable la utilidad y necesidad de estos modelos ya que se precisa analizar variadas habilidades latentes y las relaciones entre ellas; en este contexto los modelos diagnósticos son una herramienta eficaz en la entrega de información puesto que analizan a cada estudiante y lo clasifican según el perfil de dominios de variables latentes al cual pertenece, como así también indican su nivel de habilidad y la probabilidad de responder acertadamente al ítem.

Finalmente, haciendo uso de esta información, el docente logra evaluar a cada uno de sus estudiantes y vislumbrar sus debilidades y dominios, lo que puede generar un cambio en su planificación orientada a la retroalimentación de cada uno de sus estudiantes según sus necesidades y, en consecuencia, el profesor mejorará sus metodologías de enseñanza alcanzando el aprendizaje esperado por él y por todos sus estudiantes.

En línea con lo anterior y con el fin de mostrar el acontecer chileno en relación a las habilidades latentes en el desarrollo del test de hipótesis, se presentan las conclusiones con respecto a los resultados obtenidos, también las conclusiones relacionadas al cuestionario y a los perfiles de clase derivados del análisis de la muestra.

5.1 Conclusiones Respecto a los Resultados

Entre los encuestados, la mayor probabilidad de dominio se consigue en el atributo asociado al planteamiento de hipótesis, cabe señalar que, a pesar de tener una alta confiabilidad, esta resulta ser la menor con respecto a los otros dominios. Por otra parte, la menor probabilidad de dominio, que aun así esta sobre el 50%, corresponde al concepto de la región de rechazo, es decir, más de la mitad de los individuos poseen una alta capacidad para establecer, calcular, discernir o analizar la región de crítica de rechazo o aceptación en un contraste de hipótesis.

Analizando con base en el modelo logístico de tres parámetros y el modelo de clasificación diagnóstica Loglineal es que se reflejan casos donde un estudiante puede responder acertadamente menos ítems que otro y aun así presentar mayor habilidad, esto se obtiene en base a que el encuestado responde correctamente los ítems más complejos. Esta situación sería un error si nos basamos en la puntuación habitual en los test de alternativas múltiples donde sólo una es correcta, sin embargo, bajo estos dos modelos, los estudiantes son medidos según sus habilidades y no sólo por la cantidad de ítems correctos.

Un ejemplo de lo anterior son los encuestados que responden correctamente el primer, cuarto y séptimo ítem en comparación de los que erran sólo en el inicial, los primeros tienen una habilidad $\theta = 0,062$ mientras que los últimos tienen una habilidad $\theta = -0,03$, lo que implica probabilidades de acierto más altas por parte de los individuos que presentan el primer patrón de respuesta (Ver Anexo 3 para la comparación general de habilidad y probabilidades de acierto).

Siguiendo el estudio con lo expuesto en la Ilustración 25, los individuos pertenecientes a los perfiles que dominan por lo menos una habilidad tienen como mínimo una probabilidad de acierto al ítem mayor a los encuestados que no dominan ningún atributo latente. Además, existen individuos que dominan dos habilidades latentes y la combinación de estas implican cuatro o seis probabilidades de acierto mayores a las de los individuos que pertenecen al grupo de referencia.

Particularmente, los sujetos que dominan todas las habilidades relacionadas al desarrollo del test de hipótesis, poseen probabilidades de acierto bastante altas en casi todos los ítems, exceptuando la probabilidad ligada al segundo ítem, este bajo éxito es ocasionado por la complejidad del ítem, particularmente su probabilidad es de 0,4, por lo tanto, no importa cuántas habilidades domine una persona, existe un ítem tan complejo que necesitará más habilidades latentes las cuales no son estudiadas en este trabajo, sin embargo, una proyección a esta investigación sería la asimilación de estas habilidades en el estudios.

Finalmente, estos resultados son razonables, debido al gran número de conceptos que deben discriminar y relacionar los estudiantes en el correcto desarrollo del test de hipótesis. En detalle y con base en la Ilustración 12 y 13, la diferencia en el porcentaje de acierto entre estudiantes, profesores titulados y docentes que decidieron seguir sus estudios, queda firmemente expresada, donde este último grupo en todos los casos presenta un mayor porcentaje de respuestas correctas, lo que se relaciona con perfiles de respuesta con mayores probabilidades de acierto en cada ítem. En caso contrario están los estudiantes que presentan bajos logros y perfiles de respuesta con más desaciertos y por consecuencia, menores habilidades y menor probabilidad de acierto asociada a sus patrones de respuesta y clases latentes.

Esta diferencia puede estar justificada desde un punto de vista temporal, ya que los estudiantes han trabajado las habilidades ligadas al test de hipótesis sólo en algunos cursos, en cambio los docentes que persiguen un título de post-grado han utilizado estas habilidades a lo largo de su carrera profesional y las han pulido en sus cursos de post-grado. En esta línea, se comparte lo planteado por Harradine, Batanero y Rossman (2011) los cuales dicen que el razonamiento inferencial debe desarrollarse en un espacio de tiempo amplio y, además, sería importante comenzar a introducirlo de forma informal desde la enseñanza secundaria y trabajar su formalidad y aplicaciones en sus años posteriores, mejorando y avanzando en su razonamiento y campos de trabajo.

Bibliografía

- 1.- Alvarado, H. (2007). *Significados institucionales y personales del teorema central del límite en la enseñanza de estadística en ingeniería. Tesis Doctoral*. Granada: Universidad de Granada.
- 2.- Ares, M. (1999). La prueba de significación de la «hipótesis cero» en las investigaciones por encuesta. *Metodología de Encuestas*, 1, 47-68.
- 3.- Barreto, A. (2012). El progreso de la Estadística y su utilidad en la evaluación del desarrollo. *Papeles de Población*, 1-31.
- 4.- Batanero, C. (2000). Controversies around the role of statistical tests in experimental research. *Mathematical Thinking and Learning*, 2, 75-91.
- 5.- Batanero, C. (2011). Del análisis de datos a la inferencia: Reflexiones sobre la formación del razonamiento estadístico. *Conferencia Internacional de Educación Matemática*. Recife, Brasil.
- 6.- Batanero, C., & Díaz, C. (2006). Methodological and didactical controversies around statistical inference. *Actes du 36ièmes Journées de la Société Française de Statistique [CD-ROM]*. Paris: Société Française de Statistique.
- 7.- Batanero, C., Godino, J., & Vallecillos, A. (1992). El análisis de datos como útil y como objeto en didáctica de la matemática. *Educación Matemática*, 4, 46-52.
- 8.- Batanero, C., Vera, O., & Díaz, C. (2012). Dificultades de estudiantes de psicología en la comprensión del contraste de hipótesis. *Números*, 80, 91-101.
- 9.- Birnbaum, I. (1982). Interpreting statistical significance. *Teaching statistics*, 4, 24-27.
- 10.-Bock, D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37, 29-51.
- 11.-Borassi, R. (1987). Exploring Mathematics through the Analysis of Errors. *For the Learning of Mathematics*, 7, 2-9.
- 12.-Borges, A., San Luis, C., Sánchez, A., & Cañadas, I. (2001). El juicio contra la hipótesis nula: muchos testigos y una sentencia virtuosa. *Psicothema*, 13, 174-178.
- 13.-Brewer, K. (1986). Behavioural statistics textbooks: source of myths and misconceptions?. En *Proceedings of the ICOTS II* (págs. 127-131). Victoria, Canada: University of Victoria.

- 14.-Caceres, M. (2016). *Modelos de Diagnósticos de Clasificación (DCMs): Una aplicación al área educacional*. Universidad de Santiago de Chile.
- 15.-Castro, E., Vanhoof, S., Van den Nororgate, W., & Onghena, P. (2007). Student's misconceptions of statistical inference: A review of the empirical evidence from research on statistical education. *Educational Research Review*, 2, 98, 113.
- 16.-Chow, S. (1996). *Statistical significance: Rational, validity and utility*. London: Sage.
- 17.-Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003.
- 18.-Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología. Tesis Doctoral*. Granada: Universidad de Granada.
- 19.-Díaz, C., & De la Fuente, I. (2004). Controversias en el uso de la inferencia en la investigación experimental. *Metodología de las Ciencias del Comportamiento*, Edición especial 2004, 161-167.
- 20.-Díaz, C., Batanero, C., & y Wilhelmi, R. (2008). Errores frecuentes en el análisis de datos en Educación y Psicología. *Publicaciones*, 35, 109-133.
- 21.-Embretson, S., & Reise, S. (2000). *Item Response Theory for Psychologists*. Mahwah, New Jersey: Erlbaum.
- 22.-Falk, R. (1986). Misconceptions of statistical significance. *Journal of Structural Learning*, 9, 83-96.
- 23.-Gongjun, X. (2013). *Statistical Inference for Diagnostic Classification Models*. Columbia University.
- 24.-Granaas, M. (2002). Hypothesis testing in psychology: throwing the baby out with the bathwater? En B. Phillips, *Proceedings of the Sixth International Conference on Teaching of Statistics*. Ciudad del Cabo: International Association for Statistical Education.
- 25.-Haagenars, J., & McCutcheon, A. (2002). *Applied latent class analysis*. Cambridge: Cambridge University Press.
- 26.-Hacking, I. (1990). *The taming of chance*. Cambridge: Cambridge University Press.
- 27.-Haedo, S. (2001). An overview of the teaching of statistics at schools and University in Argentina. *Proceedings of the 53rd Session of the International Statistical Institute, Bulletin of ISI* (págs. 165-167). Seoul: International Statistical Institute.
- 28.-Haller, H., & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research*, 7, 1-20.

- 29.-Harlow, L., Mulaik, A., & Steiger, H. (1997). *What if there were no significance tests?* Mahwah, New Jersey: Erlbaum.
- 30.-Harradine, A., Batanero, C., & Rossman, A. (2011). *Students' and teachers' knowledge of sampling and inference. Teaching statistics in school mathematics- Challenges for teaching and teacher education.* New York: Springer, en prensa.
- 31.-Hartz, S. (2002). *A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality.* University of Llinois.
- 32.-Henson, R., Templin, J., & Wilse, J. (2009). Defining a Family of Cognitive Diagnosis Models Using Log-Linear Models with Latent Variables. *Psychometrika*, 74, 191.
- 33.-Holmes, P. (2002). Some lessons to be learnt from curriculum developments in statistics. *Proceedings of the Sixth International Conference on Teaching of Statistics*, Ciudad del Cabo.
- 34.-Junker, B., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25, 258-272.
- 35.-Lecoutre, B. (1999). Beyond the significance test controversy: Prime time for Bayes? *Bulletin of the International Statistical Institute: Proceedings of the Fifty-second Session of the International Statistical Institute.* Helsinki: International Statistical Institute.
- 36.-Lecoutre, B. (2006). Training students and researchers in Bayesian methods for experimental data analysis. *Journal of Data Science*, (págs. 207-232).
- 37.-Lecoutre, B., Lecoutre, P., & Pointevineau, J. (2001). Uses, abuses and misuses of significance tests in the scientific community: Won't the Bayesian choice be unavoidable? *International Statistical Review*, 69, 339-418.
- 38.-Lord, F. (1980). *Applications of item response theory to practical testing problems.* Mahwah, New Jersey: Erlbaum.
- 39.-Lord, F., & Novick, M. (1968). *Statistical Theories of Mental Test Scores.* Addison-Wesley: Reading, MA.
- 40.-Macready, G., & Dayton, C. (1977). The use of probabilistic models in the assessment of mastery. *Journal of Educational Statistics*, 2, 99-120.
- 41.-McDonald, R. (1967). Non-linear Factor Analysis. *Psychometric Monographs*, 15.
- 42.-Menon, R. (1993). Statistical significance testing should be discontinued in mathematics education research. *Mathematics Education Research Journal*, 5, 4-18.

- 43.-Morrison, E., & Henkel, E. (1970). *The significance tests controversy - a reader*. Chicago: Aldine.
- 44.-Ostini, R., & Nering, M. (2006). *Polytomous Item Response Theory Models*. Sara Miller McCune: Sage Publication, Inc.
- 45.-Pimienta, R. (2002). *Prueba Estadística de Hipótesis*. México: Universidad Autónoma Metropolitana.
- 46.-Rao, R., & Széleli, G. (2000). *Statistics for the 21st century*, Marcel Dekker. Nueva York.
- 47.-Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research.
- 48.-Reckase, M. (2009). *Multidimensional Item Response Theory*. Springer.
- 49.-Rossman, A. (2008). Reasoning about informal statistical inference: one statistician's view. *Statistics Education Research Journal*, 7, 5-19.
- 50.-Rupp, A., & Jonathan, T. (2010a). Unique Characteristics of Diagnostic Classification Models: A Comprehensive Review of the Current State-of-the-Art. En *Measurement: Interdisciplinary Research and Perspectives* (págs. 219-262).
- 51.-Rupp, A., Templin, J., & Henson, R. (2010b). *Diagnostic Measurement: Theory, Methods, and Applications*. Guilford Press.
- 52.-Samejima, F. (1968). Estimation of Latent Ability Using a Response. *Psychometric Monographs*, 17.
- 53.-Samejima, F. (1974). Normal ogive model on the continuous response level in the multidimensional latent space. *Psychometrika*, 39, 111-121.
- 54.-Tatsuoka, K. (1983). Rule space: an approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20, 345-354.
- 55.-Templin, J. (2008). *CDM: cognitive diagnosis modeling with Mplus*. Obtenido de http://jonathantemplin.com/files/dcm/dcm14pre932/dcm14pre932_lecture04.2.pdf
- 56.-Templin, J. (14 de Noviembre de 2016). *Psychometric Models: The Loglinear Cognitive Diagnosis Model*. Obtenido de Psychometrics and Statistics as a Way of Life: http://jonathantemplin.com/files/dcm/dcm16ncme/dcm16ncme_section03.pdf.
- 57.-Templin, J., & Bradshaw, L. (2013). Diagnosing Teachers' Understandings of Rational Numbers: Building a Multidimensional Test Within the Diagnostic Classification Framework. *Educational Measurement*, 33, 2-14.

- 58.-Templin, J., & Henson, R. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, 11, 287-305.
- 59.-Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: three suggested reforms. *Educational Researcher*, 25, 26-30.
- 60.-Trosset, M. (2006). *An Introduction to Statistical Inference and Its Applications*. Virginia.
- 61.-Tucker, L. R. (1946). Maximum validity of a test with equivalent items. En *Psychometrika* (págs. 1-13).
- 62.-Vallecillos, A. (1994). *Estudio teórico-experimental de errores y concepciones sobre el contraste estadístico de hipótesis en estudiantes universitarios. Tesis Doctoral*. Universidad de Granada.
- 63.-Vallecillos, A. (1995). *Sugerencias metodológicas para la enseñanza del teorema central del límite en la enseñanza secundaria*. Granada: Fundación Educación y Futuro.
- 64.-Vallecillos, A. (1999). *Evidencias empíricas sobre dificultades en el aprendizaje de los test de hipótesis*. Granada: Universidad de Granada.
- 65.-Vallecillos, A., & Batanero, C. (1997a). Aprendizaje y enseñanza del contraste de hipótesis: concepciones y errores. *Enseñanza de las Ciencias*, 15, 189-197.
- 66.-Vallecillos, A., & Batanero, C. (1997b). *Conceptos activados en el contraste de hipótesis estadísticas y su comprensión por estudiantes universitarios*. Obtenido de Universidad de Granada: <http://www.ugr.es/~batanero/pages/ARTICULOS/Recherches.pdf>
- 67.-Van der Linden, W. (1978). Forgetting, guessing, and mastery: The macready and dayton models revisited and compared with a latent trait approach. *Journal of Educational Statistics*, 3, 305-317.
- 68.-Vera, O., & Díaz, C. (2013). Dificultades de estudiantes de psicología en relación al contraste de Hipotesis. *Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria*, (págs. 197-203). Granada.
- 69.-Vera, O., Diaz, C., & Batanero, C. (2011). Dificultades en la formulación de hipótesis estadísticas por estudiantes de Psicología. *Union*, 27, 41-61.
- 70.-Von Davier, M. (2014). *The Log-Linear Cognitive Diagnostic Model (LCDM) as a Special Case of the General Diagnostic Model (GDM)*. Obtenido de ETS Research Report Series: <http://onlinelibrary.wiley.com/doi/10.1002/ets2.12043/abstract>

71.-Ward, N. (9 de Noviembre de 2015). Obtenido de Learnand Teach Statistics:
<https://learnandteachstatistics.wordpress.com/2015/11/09/understanding-statistical-inference/>

72.-Wim, v. d., & Ronald, H. (1997). *Handbook of Modern Item Response Theory*. Springer.


```

* Import Q-matrix into SAS;
DATA &saslibname..&Qname.;
INPUT &itemstem. attribute1-attribute3 &ordervar.;
DATALINES;
1 0 1 0 1
2 0 0 1 1
3 0 0 1 1
4 1 0 0 1
5 1 0 0 1
6 1 0 1 2
7 0 1 1 2;
RUN;

```

```

* Import original data into SAS dataset;
DATA &saslibname..&dataname.;
INPUT &IDname. &itemlist.;
datalines;

```

1	1	1	1	1	1	0	1
2	1	1	1	1	1	0	1
3	1	1	1	1	0	1	1
4	1	0	0	1	1	1	0
5	1	0	1	1	0	0	0
6	0	0	0	1	1	0	0
7	0	0	0	1	0	0	0
8	0	0	1	1	1	1	1
9	0	0	1	0	1	1	0
10	1	0	0	1	1	0	1
11	1	0	0	0	1	0	0
12	1	0	0	0	0	0	0
13	1	0	1	1	1	1	0
14	1	0	0	1	1	1	0
15	0	1	0	0	1	1	1
16	0	1	0	0	0	1	1
17	1	0	1	1	1	0	1
18	1	0	0	1	1	1	0
19	1	0	1	0	1	1	1
20	1	0	0	1	1	1	1
21	1	1	0	1	1	1	1
22	1	0	1	1	0	0	1
23	1	0	1	1	1	1	1
24	1	1	1	1	1	0	1
25	0	1	0	1	1	1	0
26	1	0	0	1	0	0	0
27	1	1	1	0	1	1	1
28	1	0	1	1	1	1	1
29	0	0	0	1	1	1	0
30	1	0	1	1	1	0	1
31	0	0	0	0	0	1	0
32	1	0	1	1	1	1	1
33	0	0	0	1	1	0	0
34	1	0	1	1	1	1	1
35	1	0	1	0	0	0	0
36	1	1	0	1	1	1	1
37	1	0	0	1	1	1	1
38	1	1	1	1	1	0	1
39	1	1	1	0	1	1	1

40	1	1	1	1	1	0	1
41	1	1	1	0	1	1	1
42	1	1	1	1	1	0	1
43	1	0	1	0	1	1	1
44	0	1	1	0	0	1	0
45	1	1	1	0	1	1	1
46	1	0	1	1	1	1	1
47	1	0	1	0	1	0	1
48	0	1	0	1	0	1	0
49	1	0	1	1	1	1	1
50	1	1	0	1	1	1	1
51	1	0	0	1	1	1	0
52	1	1	1	1	1	0	1
53	1	0	1	1	1	1	1
54	1	0	1	1	1	1	1
55	1	1	0	1	1	1	1
56	1	1	1	0	1	1	1
57	1	0	0	0	0	0	0
58	1	1	0	0	0	1	1
59	1	0	1	1	1	1	1
60	1	0	0	0	0	0	0
61	0	0	0	1	0	0	0
62	1	1	0	1	1	1	1
63	1	1	1	1	1	0	1
64	1	0	1	0	1	0	1
65	1	0	1	1	1	1	1
66	1	1	0	1	1	1	1
67	1	0	1	1	1	1	1
68	1	0	0	1	0	0	0
69	0	1	0	0	0	1	1
70	1	0	0	0	0	1	1
71	1	1	1	1	1	0	1
72	0	0	0	1	1	1	0
73	1	1	1	0	1	1	1
74	1	1	0	1	1	1	1
75	1	1	1	0	1	1	1
76	1	0	1	1	1	1	1
77	0	0	1	1	1	0	0
78	0	1	0	0	1	0	0
79	1	0	1	1	1	1	1
80	1	0	0	1	1	0	1
81	1	1	0	0	1	0	1
82	1	1	1	1	1	1	0
83	0	0	0	1	0	0	0
84	1	1	0	1	1	1	1
85	0	0	0	0	1	1	0
86	1	0	1	1	1	0	0
87	0	0	0	0	0	1	0
88	1	0	0	1	0	0	0
89	0	1	0	1	1	1	1
90	0	0	0	1	0	0	0
91	1	0	0	1	1	1	0
92	1	1	0	1	1	1	1
93	1	0	0	1	1	1	0
94	1	1	1	1	1	0	1
95	0	0	0	0	1	1	0

96	0	1	0	0	1	0	0
97	1	0	0	0	1	1	1
98	1	0	1	1	1	1	1
99	1	0	1	1	1	0	1
100	0	0	0	1	0	0	0
101	1	0	1	1	1	1	1
102	1	1	0	1	1	1	1
103	1	0	1	1	1	1	1
104	1	0	0	1	0	1	1
105	1	0	1	1	1	1	1
106	0	0	0	0	1	1	0
107	1	1	0	1	1	1	1
108	0	0	0	0	1	1	0
109	1	0	0	1	1	1	1
110	0	0	0	1	0	1	0
111	1	0	1	1	1	1	1
112	1	0	0	1	1	0	1
113	1	0	1	1	0	0	1
114	1	0	0	0	1	0	0
115	0	0	0	1	0	0	0
116	1	0	0	1	0	1	1
117	0	1	0	0	0	0	1
118	1	0	0	1	0	0	0
119	1	1	0	1	1	1	1
120	1	0	1	1	0	1	1
121	0	0	0	0	1	1	1
122	0	0	0	1	0	0	0
123	1	0	1	1	1	1	0
124	1	0	0	0	1	0	0
125	1	1	1	1	1	0	1
126	1	0	0	0	1	1	0
127	0	0	1	1	0	0	0
128	0	0	0	1	1	0	1
129	0	0	0	1	1	0	0
130	1	0	0	1	1	1	1
131	1	0	0	1	0	0	0
132	0	0	1	0	1	0	1
133	1	0	0	0	0	0	0
134	1	0	1	1	1	1	1
135	1	0	1	0	1	1	0
136	1	1	0	1	1	1	1
137	1	1	0	1	1	0	0
138	1	0	0	1	1	1	0
139	1	0	0	1	1	1	1
140	0	1	0	1	1	0	0
141	1	0	1	1	1	0	1
142	0	0	1	0	1	1	0
143	1	0	0	1	1	1	0
144	1	1	1	1	1	0	1
145	1	0	0	0	0	1	0
146	1	0	1	1	0	0	1
147	1	0	1	1	1	1	1
148	0	1	0	1	0	0	0
149	0	0	1	1	1	1	1
150	0	0	1	0	0	0	1
151	1	0	0	1	1	0	1

152	0	0	0	0	0	0	1
153	1	0	0	1	1	1	1
154	1	1	0	1	1	1	1
155	0	0	0	1	1	1	1
156	1	0	1	0	0	1	1
157	1	0	1	1	1	1	1
158	0	0	0	0	1	0	0
159	0	0	0	1	0	0	0
160	1	0	1	1	1	1	1
161	1	0	1	1	1	1	1
162	1	0	0	0	1	0	1
163	1	0	1	0	1	0	1
164	1	1	1	1	1	1	1
165	1	0	0	0	0	1	1
166	1	1	1	0	1	1	1
167	1	1	0	1	1	1	0
168	0	0	1	0	1	0	1
169	1	0	1	0	1	0	1
170	1	0	1	0	1	1	1
171	1	0	0	0	1	1	1
172	1	0	1	1	1	1	0
173	1	0	1	0	1	1	0
174	1	0	0	0	0	0	1
175	0	0	1	0	0	1	0
176	0	0	0	0	1	0	0
177	1	1	1	0	1	1	0
178	0	0	0	0	1	0	0
179	1	0	0	1	0	1	0
180	1	0	0	1	0	1	0
181	1	0	1	1	0	0	1
182	1	0	0	0	1	1	1
183	0	0	0	0	0	1	0
184	0	1	0	0	1	0	0
185	1	0	0	1	0	1	1
186	1	1	0	1	1	1	1
187	1	0	0	1	0	0	1
188	1	0	1	1	1	1	1
189	1	0	0	0	0	1	0
190	1	0	1	1	1	0	1
191	1	0	0	0	1	0	0
192	1	0	0	0	1	1	1
193	1	1	0	1	1	0	0
194	0	0	0	1	0	0	0
195	1	1	1	1	0	0	1
196	0	0	0	0	0	1	0
197	1	0	1	1	1	1	1
198	1	0	0	1	0	1	0
199	0	0	0	0	0	1	0
200	1	0	0	0	0	1	0
201	1	0	1	0	1	0	1
202	0	0	0	1	1	1	0
203	1	1	1	1	1	0	0
204	1	1	0	1	1	0	1
205	1	1	0	1	0	1	0
206	1	0	0	1	0	1	1
207	1	1	1	0	0	1	1

208	0	0	0	1	0	0	0
209	0	1	0	1	1	0	0
210	1	0	1	1	1	1	1
211	1	1	0	1	1	1	1
212	1	0	0	0	0	1	1
213	1	0	1	1	1	1	1
214	0	0	0	0	1	1	0
215	0	0	0	0	1	1	0
216	1	1	1	1	1	0	1
217	1	1	1	1	1	0	1
218	1	0	1	1	1	0	1
219	1	1	1	1	1	1	0
220	1	0	1	1	1	1	1
221	0	1	1	0	1	1	1
222	0	0	0	1	0	0	0
223	1	1	1	0	1	1	1
224	1	1	0	1	1	1	1
225	0	0	0	1	0	1	0
226	1	1	0	1	1	1	1
227	1	0	0	1	1	1	0
228	1	0	1	1	1	1	1
229	1	1	1	0	0	1	0
230	1	0	1	0	1	1	1
231	0	0	0	1	0	0	1
232	1	1	1	1	1	0	1
233	1	1	0	0	1	1	0
234	1	0	0	1	0	0	0
235	1	1	1	0	0	1	1
236	1	0	0	1	1	0	0
237	0	0	0	0	1	0	0
238	1	1	1	1	0	1	1
239	1	0	1	1	1	1	1
240	1	1	1	0	1	1	1
241	0	0	0	1	0	0	0
242	0	0	0	1	1	1	1
243	1	0	0	1	0	0	0
244	1	0	1	1	0	0	1
245	0	0	0	0	0	1	0
246	0	0	0	0	0	1	0
247	0	0	1	0	1	0	0
248	0	1	1	1	1	1	1
249	0	1	0	0	1	0	0
250	1	0	0	1	1	1	1
251	0	0	0	1	0	0	0
252	1	0	1	1	1	1	1
253	1	0	1	1	1	0	0
254	1	0	1	1	1	1	1
255	1	0	0	0	0	0	0
256	0	0	0	1	0	0	0
257	1	1	0	1	1	1	1
258	1	1	1	1	1	0	0
259	1	0	1	1	1	1	1
260	1	0	0	0	1	1	0
261	0	0	1	1	1	0	1
262	1	1	0	0	1	1	1
263	0	0	0	0	1	1	0

```

264  0    0    0    1    0    0    0
265  0    0    1    0    0    0    0
266  1    0    1    0    1    0    1
267  1    0    0    0    1    0    0
268  0    0    1    1    0    0    0
269  1    1    0    1    1    1    1
270  1    0    0    1    0    0    0
271  1    0    1    0    1    0    1
272  1    0    1    1    0    1    1
273  1    1    1    1    1    0    1
274  1    0    1    1    1    1    1
275  0    0    0    1    0    0    0
276  1    1    0    0    1    1    1
277  1    0    0    1    0    0    0
278  1    0    0    1    1    0    1
279  0    0    0    1    0    0    1
280  1    1    1    1    1    0    1
281  0    0    1    0    0    1    1
282  1    0    0    0    0    0    0
283  1    0    1    0    0    1    1
284  1    1    0    1    1    1    1
285  0    1    0    0    1    1    1
286  1    0    0    0    1    1    0
287  0    0    0    1    0    0    0
288  0    0    1    1    1    0    1
289  1    0    1    1    1    1    1
290  0    0    1    0    0    0    0
291  0    0    0    1    0    0    0
292  1    0    0    1    1    1    1
293  1    0    1    1    1    1    1
294  1    0    1    1    1    1    1
295  1    1    1    1    0    1    1
296  1    0    1    0    1    1    0
297  1    0    1    1    1    0    1
298  1    0    1    1    1    1    1
299  1    0    0    1    0    0    0
300  0    0    0    0    1    1    0

```

RUN;

```

*****
*****                               MACRO EXECUTION:                               *****
*****                               NO CHANGES TO MAKE IN THIS SECTION           *****
*****;

```

```

* Call master file;
  %INCLUDE "&macroloc.\LCDM_Mplus2.sas";

* Generate Mplus .dat file and input script;
  %CreateMplusInput(
    filesave=&filesave., filename=&filename., saslibname=&saslibname.,
    Qname=&Qname., dataname=&dataname., IDname=&IDname., itemstem=&itemstem.,
    itemlist=&itemlist., numitem=&numitem., ordervar=&ordervar.,
    maxitemorder=&maxitemorder.,
    attstem=&attstem., attcat=&attcat., numatt=&numatt., numclass=&numclass.,
    structon=&structon., structorder=&structorder., loosen=&loosen.,
    processors=&processors.);

```

```
* Running Mplus - program to call, input file, output file;  
  X CALL "C:\Program Files\Mplus\mplus.exe"  
        "&filesave.\&filename..inp"  
        "&filesave.\&filename..out";
```

```
* Calling in master program to import Mplus output;  
  %ImportMplusOutput;
```

*** The following datasets have now been created in the work library:

ReadMplus = all Mplus output text in one variable (one row per line)

Respondents = original item responses, prob of class membership, most likely class membership

ClassCounts = model-estimated class counts and proportions

Thresholds = Estimates, SE, Z-score, and p-values per item per threshold

Itemparms = LCDM Intercept (order=0), main effects (order=1), and interactions (order=2+) per item;

*** If using a structural model, these datasets are also created:

ClassMeans = Mean estimates, SE, Z-score, and p-values per class

StrucParms = Intercept (order=0), main effects (order=1), and interactions (order=2+) per attribute;

```
* Run this to output these tables of results as SAS datasets and as an excel workbook;  
  %WriteData;
```



```

* MACRO Classpattern:
* Creates an initial pattern of as many 0s and 1s in a row as needed
* Copies that pattern into separate datasets as needed
* Concatenates (then deletes) the separate datasets
* Result is one dataset per attribute with enough rows for all classes
* MACRO Mergelt:
* Merges (then deletes) datasets from eaach attribute together; */
%MACRO ClassPattern;
    %LET divisor = 2;
    %LET repeat = 1;
    %LET totclass = &numclass.;
    %DO a=1 %TO &numatt.;
        %DO r=1 %TO %EVAL(&repeat.);
            DATA att&a.&r.;
                DO c=1 TO &totclass.;
                    IF c LE (&totclass./&divisor.) THEN classatt&a.=0; ELSE
classatt&a.=1;
                OUTPUT;
            END; DROP c;
        run;
        %LET repeat = %EVAL(&repeat.+1);
    %END;
    %IF &a.=1 %THEN %LET order = 1; %ELSE %LET order=%EVAL(&order.*2);
    DATA att&a.; SET %DO loop=1 %TO &order.; Att&a.&loop. %END; ; run;
    PROC DATASETS LIB=WORK NOLIST; DELETE %DO loop=1 %TO &order.;
Att&a.&loop. %END; ;
    RUN; QUIT;
    %LET totclass = %EVAL(&totclass./2);
    %END;
%MEND ClassPattern;
%ClassPattern;
%MACRO Mergelt;
DATA classpattern; RETAIN class; MERGE
    %DO a=1 %TO &numatt.; Att&a. %END; ;
    class = _N_; RUN;
PROC DATASETS LIB=WORK NOLIST; DELETE %DO a=1 %TO &numatt.; Att&a. %END; ;
RUN; QUIT;
%MEND Mergelt;
%Mergelt;

* Creating item*attribute kernal table from classpattern matrix and Q matrix;
* Result is one row per class per item;
DATA kernel;
    DO class=1 TO &numclass.;
        DO item=1 TO &numitem.;
            class=class; item=item; OUTPUT;
        END;
    END;
RUN;
DATA kernel; MERGE kernel classpattern; BY class; RUN;
PROC SORT DATA=kernel; BY item class; RUN;
DATA kernel; MERGE kernel Qmatrix; BY item; RUN;

* Scoring by class and item to create kernel threshold values;
* Result is set of variables called "scoreatt" to be activated
if the item requires the attribute AND the class has it;

```

```

DATA kernel; SET kernel;
  ARRAY aclass(&numatt.) classatt1-classatt&numatt.;
  ARRAY aitem(&numatt.) itematt1-itematt&numatt.;
  ARRAY ascore(&numatt.) scoreatt1-scoreatt&numatt.;
  DO i=1 TO &numatt.;
    IF aclass(i)=1 AND aitem(i)=1 THEN ascore(i)=1; ELSE ascore(i)=0;
  END; DROP i;
RUN;
* Creating single variables that hold the "scoreatt" and attribute patterns;
DATA kernel; LENGTH scorepattern $10 attpattern $&numatt.; SET kernel;
  scorepattern = CAT(OF item scoreatt1-scoreatt&numatt.);
  attpattern = CAT(OF scoreatt1-scoreatt&numatt.);
RUN;
PROC SORT DATA=kernel; BY item scorepattern; RUN;
* Creates count variable for whether each "scoreatt" is unique
  to an item to create a new threshold index as needed;
DATA kernel; SET kernel; BY item scorepattern; RETAIN thresh;
  IF FIRST.Item THEN thresh=1;
  IF scorepattern=LAG1(scorepattern) THEN thresh=thresh;
  ELSE thresh=thresh+1;
RUN;
* The threshold count is one unit too high, so this FIXES it;
DATA kernel; RETAIN class item scorepattern thresh; SET kernel;
  thresh=thresh-1;
RUN;
PROC SORT DATA=kernel; BY item class; RUN;

```

```

*****
*****                               Data Manipulation:                               *****
*****                               Writing Equations out of Patterns                    *****
*****
*****
*****
*****

```

```

* Creating variables needed for parameters of the LCDM equations;
%MACRO Equations;
DATA kernel;
  * Listing all character variables eventually created;
  LENGTH NewStruc NewStrucAdd NewParm NewParmAdd
        Main OrderMain Order12way Order22way
        Order13way Order23way Order33way
        OrderMainMax Order12wayMax Order22wayMax
        Order13wayMax Order23wayMax Order33wayMax $500;
  SET kernel;
  *** STRUCTURAL MODEL PARAMETERS;
  * Dummy variable to use as placeholder;
  BeginS = .;
  * Main effects;
  %DO a=1 %TO &numatt.;
    IF classatt&a.=1 THEN G_1&a.= "G_1&a. ";
    ELSE G_1&a. = G_1&a.;
  %END;
  * Two-way interactions;
  %IF &structorder.>1 %THEN %DO;
    %DO Num1=1 %TO &numatt.;
      %DO Num2=2 %TO &numatt.;
        %LET check=%EVAL(&Num1.-&Num2.);
      ;
    ;
  ;
  ;

```

```

                                %IF &check.<0 %THEN %DO;
                                IF classatt&Num1.=1 AND classatt&Num2.=1
                                    THEN                                G_2&Num1.&Num2.=
"G_2&Num1.&Num2. ";
                                    ELSE                                G_2&Num1.&Num2.=
G_2&Num1.&Num2.;
                                %END;
                                %END;
                                %END;
                                %END;
                                * Three-way interactions;
                                %IF &structorder.>2 %THEN %DO;
                                    %DO Num1=1 %TO &numatt.;
                                        %DO Num2=2 %TO &numatt.;
                                            %DO Num3=3 %TO &numatt.;
                                                %LET check1=%EVAL(&Num1.-&Num2.);
                                                %LET check2=%EVAL(&Num2.-&Num3.);
                                                %IF &check1.<0 AND &check2.<0 %THEN %DO;
                                                    IF classatt&Num1.=1 AND classatt&Num2.=1 AND
classatt&Num3.=1
                                                        THEN                                G_3&Num1.&Num2.&Num3.=
"G_3&Num1.&Num2.&Num3. ";
                                                        ELSE                                G_3&Num1.&Num2.&Num3.=
G_3&Num1.&Num2.&Num3.;
                                                    %END;
                                                %END;
                                            %END;
                                        %END;
                                    %END;
                                %END;
                                * Four-way interactions;
                                %IF &structorder.>3 %THEN %DO;
                                    %DO Num1=1 %TO &numatt.;
                                        %DO Num2=2 %TO &numatt.;
                                            %DO Num3=3 %TO &numatt.;
                                                %DO Num4=4 %TO &numatt.;
                                                    %LET check1=%EVAL(&Num1.-&Num2.);
                                                    %LET check2=%EVAL(&Num2.-&Num3.);
                                                    %LET check3=%EVAL(&Num3.-&Num4.);
                                                    %IF &check1.<0 AND &check2.<0 AND
&check3.<0 %THEN %DO;
                                                        IF classatt&Num1.=1 AND classatt&Num2.=1
AND classatt&Num3.=1
                                                            AND classatt&Num4.=1
                                                                THEN
G_4&Num1.&Num2.&Num3.&Num4.= "G_4&Num1.&Num2.&Num3.&Num4. ";
                                                                ELSE
G_4&Num1.&Num2.&Num3.&Num4.= G_4&Num1.&Num2.&Num3.&Num4.;
                                                                %END;
                                                            %END;
                                                        %END;
                                                    %END;
                                                %END;
                                            %END;
                                        %END;
                                    %END;
                                %END;
                                %END;
                                * Five-way interactions;
                                %IF &structorder.>4 %THEN %DO;
                                    %DO Num1=1 %TO &numatt.;

```



```

Beginl=.;
* Intercepts per item (1);
  %DO i=1 %TO &numitem.;
    IF item=&i. THEN Lint = "L&i._0  ";
    ELSE Lint = Lint;
  %END;
* Main effects per item (up to # attributes);
  %DO Num1=1 %TO &numatt.;
    %DO i=1 %TO &numitem.;
      IF item=&i. AND scoreatt&Num1.=1 THEN L&i._1&Num1. =
"L&i._1&Num1.  ";
      ELSE L&i._1&Num1. = L&i._1&Num1.;
    %END;
  %END;
* Two-way interactions per item;
  %DO i=1 %TO &numitem.;
    %DO Num1=1 %TO &numatt.;
      %DO Num2=2 %TO &numatt.;
        %LET check=%EVAL(&Num1.-&Num2.);
        %IF &check.<0 AND &&itemorder&i.>1 %THEN %DO;
          IF item=&i. AND scoreatt&Num1.=1 AND
scoreatt&Num2.=1
          THEN L&i._2&Num1.&Num2.=
"L&i._2&Num1.&Num2.  ";
          ELSE L&i._2&Num1.&Num2.=
L&i._2&Num1.&Num2.;
        %END;
      %END;
    %END;
  %END;
* Three-way interactions per item;
  %DO i=1 %TO &numitem.;
    %DO Num1=1 %TO &numatt.;
      %DO Num2=2 %TO &numatt.;
        %DO Num3=3 %TO &numatt.;
          %LET check1=%EVAL(&Num1.-&Num2.);
          %LET check2=%EVAL(&Num2.-&Num3.);
          %IF &check1.<0 AND &check2.<0 AND
&&itemorder&i.>2 %THEN %DO;
            IF item=&i. AND scoreatt&Num1.=1 AND
scoreatt&Num2.=1 AND scoreatt&Num3.=1
            THEN L&i._3&Num1.&Num2.&Num3.=
"L&i._3&Num1.&Num2.&Num3.  ";
            ELSE L&i._3&Num1.&Num2.&Num3.=
L&i._3&Num1.&Num2.&Num3.;
          %END;
        %END;
      %END;
    %END;
  %END;
* Four-way interactions per item;
  %DO i=1 %TO &numitem.;
    %DO Num1=1 %TO &numatt.;
      %DO Num2=2 %TO &numatt.;
        %DO Num3=3 %TO &numatt.;
          %DO Num4=4 %TO &numatt.;

```



```

* Last variable to use as dummy place to end series;
  Last1 = .;
* Creating string variables to use in NEW equations;
  CALL CATX(" ", NewParm, OF Begin1--Last1);
  NewParm = TRANWRD(NewParm, ".", " ");
  NewParm = STRIP(NewParm);
  NewParmAdd = TRANWRD(STRIP(NewParm), " ", "+");
* Creating index for longest possible number of parms;
  Long1Parm = LENGTH(STRIP(NewParm));
RUN;
* Saving longest parm as new variable for NEW statement;
PROC SORT DATA=kernel; BY item DESCENDING Long1Parm; RUN;
DATA Save1Long; SET kernel; BY item;
  IF FIRST.item THEN NewParmMax=NewParm; ELSE DELETE; RUN;
* Merge longest parm back into data;
DATA kernel; LENGTH NewParmMax $100; MERGE Save1Long kernel; BY item; RUN;
* Clearing extra datasets;
PROC DATASETS LIB=WORK NOLIST; DELETE Save1Long; RUN; QUIT;

* Creating string variables for ordering constraints;
%LET numatt1 = %EVAL(&numatt.-1);
%LET numatt2 = %EVAL(&numatt.-2);
DATA kernel; SET kernel;
  * Main effects;
  CALL CATX(" ", Main, OF L1_11--L&numitem._1&numatt.);
  * Combining into one variable;
  OrderMain = CATT(TRANWRD(STRIP(Main), " ", ">0; " ), ">0; " );
  * Creating index for longest possible number of parms;
  LongOrderMain = LENGTH(STRIP(OrderMain));
* 1 Two-way interactions - need two sets of loops for both possible constraints;
* Creating dummy variable to begin first two-way interaction series;
BeginO12way=.;
%DO i=1 %TO &numitem.;
  %DO Num1=1 %TO &numatt.;
    %DO Num2=2 %TO &numatt.;
      %LET check1=%EVAL(&Num1.-&Num2.);
      %IF &check1.<0 AND &&itemorder&i.>1 %THEN %DO;
        IF item=&i. AND scoreatt&Num1.=1 AND
scoreatt&Num2.=1 THEN DO;
          O1L&i._2&Num1.&Num2.          =
"L&i._2&Num1.&Num2.-L&i._1&Num1.;      ";
          END;
        %END;
      %END;
    %END;
  %END;
  * Creating dummy variable to end first two-way interaction series;
  EndO12way=.;
* 2 Two-way interactions - need two sets of loops for both possible constraints;
* Creating dummy variable to begin second two-way interaction series;
BeginO22way=.;
%DO i=1 %TO &numitem.;
  %DO Num1=1 %TO &numatt.;
    %DO Num2=2 %TO &numatt.;
      %LET check1=%EVAL(&Num1.-&Num2.);
      %IF &check1.<0 AND &&itemorder&i.>1 %THEN %DO;

```

```

IF item=&i. AND scoreatt&Num1.=1 AND
scoreatt&Num2.=1 THEN DO;
O2L&i._2&Num1.&Num2. =
"L&i._2&Num1.&Num2.-L&i._1&Num2.;" ;
END;
%END;
%END;
%END;
%END;
* Creating dummy variable to end second two-way interaction series;
EndO22way=.;
%IF &maxitemorder.>1 %THEN %DO;
* Combining into one variable;
Order12way = CATS(OF BeginO12way--EndO12way);
Order12way = TRANWRD(Order12way, ".", " ");
Order22way = CATS(OF BeginO22way--EndO22way);
Order22way = TRANWRD(Order22way, ".", " ");
* Creating index for longest possible number of parms;
LongOrder12way = LENGTH(STRIP(Order12way));
LongOrder22way = LENGTH(STRIP(Order22way));
%END;
* 1 Three-way interactions - need three sets of loops for all three sets of constraints;
* Creating dummy variable to begin first three-way interaction series;
BeginO13way=.;
%DO i=1 %TO &numitem.;
%DO Num1=1 %TO &numatt.;
%DO Num2=2 %TO &numatt.;
%DO Num3=3 %TO &numatt.;
%LET check1=%EVAL(&Num1.-&Num2.);
%LET check2=%EVAL(&Num2.-&Num3.);
%IF &check1.<0 AND &check2.<0 AND
&&itemorder&i.>2 %THEN %DO;
IF item=&i. AND scoreatt&Num1.=1 AND
scoreatt&Num2.=1 AND scoreatt&Num3.=1 THEN DO;
O1L&i._3&Num1.&Num2.&Num3. = "L&i._3&Num1.&Num2.&Num3.-"
(L&i._2&Num2.&Num3.+L&i._2&Num1.&Num3.+L&i._1&Num3.);";
END;
%END;
%END;
%END;
%END;
%END;
* Creating dummy variable to end first three-way interaction series;
EndO13way=.;
* 2 Three-way interactions - need three sets of loops for all three sets of constraints;
* Creating dummy variable to begin second three-way interaction series;
BeginO23way=.;
%DO i=1 %TO &numitem.;
%DO Num1=1 %TO &numatt.;
%DO Num2=2 %TO &numatt.;
%DO Num3=3 %TO &numatt.;
%LET check1=%EVAL(&Num1.-&Num2.);
%LET check2=%EVAL(&Num2.-&Num3.);
%IF &check1.<0 AND &check2.<0 AND
&&itemorder&i.>2 %THEN %DO;

```

```

IF item=&i. AND scoreatt&Num1.=1 AND
scoreatt&Num2.=1 AND scoreatt&Num3.=1 THEN DO;

    O2L&i._3&Num1.&Num2.&Num3. = "L&i._3&Num1.&Num2.&Num3.>-
(L&i._2&Num2.&Num3.+L&i._2&Num1.&Num2.+L&i._1&Num2.);";
    END;
    %END;
    %END;
    %END;
    %END;
    * Creating dummy variable to end second three-way interaction series;
    EndO23way=.;
* 3 Three-way interactions - need three sets of loops for all three sets of constraints;
* Creating dummy variable to begin third three-way interaction series;
BeginO33way=.;
%DO i=1 %TO &numitem.;
    %DO Num1=1 %TO &numatt.;
        %DO Num2=2 %TO &numatt.;
            %DO Num3=3 %TO &numatt.;
                %LET check1=%EVAL(&Num1.-&Num2.);
                %LET check2=%EVAL(&Num2.-&Num3.);
                %IF &check1.<0 AND &check2.<0 AND
&&itemorder&i.>2 %THEN %DO;
                    IF item=&i. AND scoreatt&Num1.=1 AND
scoreatt&Num2.=1 AND scoreatt&Num3.=1 THEN DO;

                        O3L&i._3&Num1.&Num2.&Num3. = "L&i._3&Num1.&Num2.&Num3.>-
(L&i._2&Num1.&Num3.+L&i._2&Num1.&Num2.+L&i._1&Num1.);";
                        END;
                        %END;
                        %END;
                        %END;
                        %END;
                        * Creating dummy variable to end third three-way interaction series;
                        EndO33way=.;
                        %IF &maxitemorder.>2 %THEN %DO;
                            * Combining into one variable;
                            Order13way = CATS(OF BeginO13way--EndO13way);
                            Order13way = TRANWRD(Order13way, ' ', ' ');
                            Order23way = CATS(OF BeginO23way--EndO23way);
                            Order23way = TRANWRD(Order23way, ' ', ' ');
                            Order33way = CATS(OF BeginO33way--EndO33way);
                            Order33way = TRANWRD(Order33way, ' ', ' ');
                            * Creating index for longest possible number of parms;
                            LongOrder13way = LENGTH(STRIIP(Order13way));
                            LongOrder23way = LENGTH(STRIIP(Order23way));
                            LongOrder33way = LENGTH(STRIIP(Order33way));
                        %END;
                    RUN;
                    *** Saving longest parm as new variable for ordering statements;
                    * Main effects;
                    PROC SORT DATA=kernel; BY item DESCENDING LongOrderMain; RUN;
                    DATA SaveMain; SET kernel; BY item;
                    * Save max value;

```

```

        IF FIRST.item THEN OrderMainMax=STRIP(OrderMain); ELSE DELETE;
    RUN;
    * Merge longest parm back into data;
    DATA kernel; MERGE SaveMain kernel(DROP=OrderMainMax); BY item; RUN;
    * Clearing extra datasets;
    PROC DATASETS LIB=WORK NOLIST; DELETE SaveMain; RUN; QUIT;
* Two-way interactions;
%IF &maxitemorder. >1 %THEN %DO;
* Do for each set;
* Set 1;
    PROC SORT DATA=kernel; BY item DESCENDING LongOrder12way; RUN;
    DATA Save2way; SET kernel; BY item;
        * Save max value;
        IF FIRST.item THEN Order12wayMax=STRIP(Order12way); ELSE DELETE;
        * Add blanks between semi-colons;
        Order12wayMax = TRANWRD(Order12wayMax, ';' , ' ');
    RUN;
    * Merge longest parm back into data;
    DATA kernel; MERGE Save2way kernel(DROP=Order12wayMax); BY item; RUN;
* Set 2;
    PROC SORT DATA=kernel; BY item DESCENDING LongOrder22way; RUN;
    DATA Save2way; SET kernel; BY item;
        * Save max value;
        IF FIRST.item THEN Order22wayMax=STRIP(Order22way); ELSE DELETE;
        * Add blanks between semi-colons;
        Order22wayMax = TRANWRD(Order22wayMax, ';' , ' ');
    RUN;
    * Merge longest parm back into data;
    DATA kernel; MERGE Save2way kernel(DROP=Order22wayMax); BY item; RUN;
    * Clearing extra datasets;
    PROC DATASETS LIB=WORK NOLIST; DELETE Save2way; RUN; QUIT;
%END;
* Three-way interactions;
%IF &maxitemorder. >2 %THEN %DO;
* Do for each set;
* Set 1;
    PROC SORT DATA=kernel; BY item DESCENDING LongOrder13way; RUN;
    DATA Save3way; SET kernel; BY item;
        * Save max value;
        IF FIRST.item THEN Order13wayMax=STRIP(Order13way); ELSE DELETE;
        * Add blanks between semi-colons;
        Order13wayMax = TRANWRD(Order13wayMax, ';' , ' '); RUN;
    * Merge longest parm back into data;
    DATA kernel; MERGE Save3way kernel(DROP=Order13wayMax); BY item; RUN;
* Set 2;
    PROC SORT DATA=kernel; BY item DESCENDING LongOrder23way; RUN;
    DATA Save3way; SET kernel; BY item;
        * Save max value;
        IF FIRST.item THEN Order23wayMax=STRIP(Order23way); ELSE DELETE;
        * Add blanks between semi-colons;
        Order23wayMax = TRANWRD(Order23wayMax, ';' , ' '); RUN;
    * Merge longest parm back into data;
    DATA kernel; MERGE Save3way kernel(DROP=Order23wayMax); BY item; RUN;
* Set 2;
    PROC SORT DATA=kernel; BY item DESCENDING LongOrder33way; RUN;
    DATA Save3way; SET kernel; BY item;

```

```

        * Save max value;
        IF FIRST.item THEN Order33wayMax=STRIP(Order33way); ELSE DELETE;
        * Add blanks between semi-colons;
        Order33wayMax = TRANWRD(Order33wayMax, ';', ' '); RUN;
    * Merge longest parm back into data;
    DATA kernel; MERGE Save3way kernel(DROP=Order33wayMax); BY item; RUN;
* Clearing extra datasets;
    PROC DATASETS LIB=WORK NOLIST; DELETE Save3way; RUN; QUIT;
%END;
PROC SORT DATA=kernel; BY item class; RUN;
%MEND Equations;

```

```
%Equations;
```

```

*** Creating datasets for use in writing script;
* Full item by class data for writing class-based commands;
DATA itembyclass; SET kernel; RUN;
PROC SORT DATA=itembyclass; BY class item; RUN;
* Class-level data for writing structural model;
DATA classlevel; SET kernel; WHERE item=1; RUN;
PROC SORT DATA=classlevel; BY class; RUN;
* Item by threshold data for writing item-based commands;
DATA itembythresh; SET kernel; RUN;
PROC SORT NODUPKEY DATA=itembythresh; BY item thresh; RUN;
DATA itembythresh; SET itembythresh; DROP class classatt1-classatt&numatt.; RUN;

```

```

*****
*****                               Writing Mplus Script                               *****
*****
*****

```

```

* Writing Mplus script for DCM;
%MACRO WriteMplus;
DATA _NULL_;
* Datasets to draw values from;
%LET dataitembyclass=%SYSFUNC(OPEN(work.itembyclass,i));
%LET dataitembythresh=%SYSFUNC(OPEN(work.itembythresh,i));
%LET dataclasslevel=%SYSFUNC(OPEN(work.classlevel,i));

* Name of Mplus file to be written with returns and breaks;
FILE "&filesave.\&filename..inp" PRINT;
* Mplus TITLE command;
    PUT @1 "TITLE: ! Section that appears in header of output file";
    PUT @5 "DCM for &dataname. with &numatt. attributes and &structorder.-order
structural model,";
    PUT @5 "&numitem. items, and maximum &maxitemorder.-order item model,";
    PUT @5 "Saturated structural model (Mplus default)." /;
* Mplus DATA command;
    PUT @1 "DATA: ! Location of free format data file"
    / @5 "FILE = &filesave.\&dataname..dat,"
    / ;
* Mplus VARIABLE command;
    PUT @1 "VARIABLE:"
    / @5 "NAMES = &IDname. mitem1-mitem&numitem.;" @40 "! List of variables in
data file"

```

```

/      @5 "USEVARIABLE = mitem1-mitem&numitem.;"    @40 "! Variables to be
analyzed"
/      @5 "CATEGORICAL = mitem1-mitem&numitem.;"    @40 "! Binary outcomes"
/      @5 "CLASSES = c(&numclass.);" @40 "! Number of possible attribute patterns
(2^A)"
/      @5 "IDvariable = &IDname.;"    @40 "! Person ID variable to save respondent
data"
/      ;
* Mplus ANALYSIS command;
  PUT @1 "ANALYSIS:"
/      @5 "TYPE = MIXTURE;" @40 "! Estimates latent classes"
/      @5 "STARTS = 0;" @40 "! Turn off multiple random start feature"
/      @5 "PROCESSORS = &processors.;" @40 "! Number of processors available"
%IF &loosen.=1 %THEN %DO;
/      @5 "mconvergence= .1;" @40 "! Less stringent convergence criteria"
/      @5 "muconvergence= .1;"
/      @5 "convergence= .1;"
/      @5 "logcriterion= .1;"
/      @5 "rlogcriterion= .1;"
/      @5 "h1convergence= .1;"
/      @5 "mcconvergence= .1;"
/      @5 "miterations= 10000;"
/      @5 "LOGHIGH=          10;"
/      @5 "LOGLOW=         -10;"
%END;
/      ;
* Mplus MODEL command;
  PUT @1 "MODEL:"
/      ;

* Code for structural model if needed;
  %IF &structon.=1 %THEN %DO;
    PUT @1 "% " "OVERALL" "%";
    * List means per class;
    %DO c=1 %TO %EVAL(&numclass.-1);
      PUT @1 "[c#&c.] (m&c.); ! Latent variable mean for class &c.";
    %END; PUT / ;
  %END;

* List class-specific models;
  %LET counter=1;
  %DO c=1 %TO &numclass.;
    PUT @1 "% " "c#&c." "% " " ! Model for Class &c.";
    %DO i=1 %TO &numitem.;
      * Gets entire row of data;
      %LET
row=%SYSFUNC(FETCHOBS(&dataitembyclass.,&counter.));
      * Grabs specific value for threshold from that row;
      %LET tnum=%SYSFUNC(GETVARN(&dataitembyclass,
%SYSFUNC(VARNUM(&dataitembyclass,thresh)));
      PUT @5 "[mitem&i.$1] (T&i._&tnum.);" @30 "! Item &i. Thresh
&tnum.";
      %LET counter=%EVAL(&counter.+1);
    %END;
  PUT / ;

```

```

%END;

* Mplus MODEL CONSTRAINT and NEW commands;
  PUT @1 "MODEL CONSTRAINT: ! Used to define LCDM parameters"
  /      @1 "! Mplus uses P(X=0) rather than P(X=1) so multiply by -1"
  /      ;

* Code for structural model if needed;
  %IF &structon.=1 %THEN %DO;
  PUT   @1 "! STRUCTURAL MODEL";
  * List class-specific means;
    %LET c=1;
    * Gets entire row of data;
    %LET row=%SYSFUNC(FETCHOBS(&dataclasslevel.,&c.));
    * Grabs specific value for threshold from that row;
    %LET newstrucline=%SYSFUNC(GETVARC(&dataclasslevel.,

%SYSFUNC(VARNUM(&dataclasslevel.,NewStrucMax)));
    PUT @1 "NEW(G_0 &newstrucline.);";
    %DO c=1 %TO %EVAL(&numclass.-1);
    * Gets entire row of data;
    %LET row2=%SYSFUNC(FETCHOBS(&dataclasslevel.,&c.));
    * Grabs specific value for threshold from that row;
    %LET strucadd=%SYSFUNC(GETVARC(&dataclasslevel.,

%SYSFUNC(VARNUM(&dataclasslevel.,NewStrucAdd)));
    %LET strucmaxadd=%SYSFUNC(GETVARC(&dataclasslevel.,

%SYSFUNC(VARNUM(&dataclasslevel.,NewStrucMaxAdd)));
    PUT @1 "m&c.= &strucadd. - (&strucmaxadd.);";
    %END;
    PUT @1 "G_0= - (&strucmaxadd.);" /;
  %END;

* NEW commands per item;
  %LET counter=1;
  %LET counterd= %EVAL(&counter.-1);
  * Have to do it manually the first time through;
  %LET i=1;
  * Gets entire row of current data;
  %LET nowrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counter.));
  * Grabs specific value for variable from that row;
  %LET newline=%SYSFUNC(GETVARC(&dataitembythresh,

%SYSFUNC(VARNUM(&dataitembythresh.,NewParmMax)));
  PUT @1 "! Item &i: Define LCDM parameters present for item &i.";
  PUT @1 "NEW(&newline.);";
  *** Equations per item;
  * Have to do first threshold manually;
  %LET t=1;
  * Gets entire row of current data;
  %LET
nowrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counter.));
  * Grabs specific value for variable from that row;
  %LET threshline=%SYSFUNC(GETVARC(&dataitembythresh,

```

```

%SYSFUNC(VARNUM(&dataitembythresh,NewParmAdd)));
      PUT @1 "T&i._&t.=-(&threshline.);" @72 "! Item &i. Thresh &t.";
      %LET counter=%EVAL(&counter.+1); %LET counterd=
%EVAL(&counter.-1);
      * Iteratively through rest of thresholds;
      %DO t=2 %TO &numclass.;
      * Gets entire row of current data;
      %LET
nowrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counter.));
      * Grabs specific value for variable from that row;
      %LET nowitem=%SYSFUNC(GETVARN(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,item))););
      * Grabs specific value for variable from that row;
      %LET threshline=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,NewParmAdd))););
      * Gets entire row of previous data;
      %LET
pastrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counterd.));
      * Grabs specific value for variable from that row;
      %LET pastitem=%SYSFUNC(GETVARN(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,item))););
      %IF &nowitem.=&pastitem. %THEN %DO;
      PUT @1 "T&i._&t.=-(&threshline.);" @72 "! Item &i.
Thresh &t.";
      %LET counter=%EVAL(&counter.+1); %LET counterd=
%EVAL(&counter.-1);
      %END;
      %END;
      *** Ordering constraints per item;
      * Gets entire row of current data;
      %LET nowrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counterd.));
      * Grabs specific value for variable from that row;
      %LET listmain=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,OrderMainMax))););
      * Grabs specific value for variable from that row;
      %IF &&itemorder&i.>1 %THEN %DO;
      %LET list12way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order12wayMax))););
      %LET list22way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order22wayMax))););
      %END;
      %IF &&itemorder&i.>2 %THEN %DO;
      %LET list13way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order13wayMax))););
      %LET list23way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order23wayMax))););
      %LET list33way=%SYSFUNC(GETVARC(&dataitembythresh,

```

```

%SYSFUNC(VARNUM(&dataitembythresh,Order33wayMax)));
  %END;
  * Need to check to see if all terms exist, otherwise get "" in input;
  PUT @1 "! Main effect order constraints"
  / @1 "&listmain.";
  %IF &&itemorder&i.>1 AND %LENGTH(&list12way.)>0 %THEN %DO;
  PUT @1 "! Two-way interaction order constraints"
  / @1 "&list12way."
  / @1 "&list22way."; %END;
  %IF &&itemorder&i.>2 AND %LENGTH(&list13way.)>0 %THEN %DO;
  PUT @1 "! Three-way interaction order constraints"
  / @1 "&list13way."
  / @1 "&list23way."
  / @1 "&list33way."; %END;
  PUT /;

  * Iterially through rest of items, checking for new thresholds to write;
  %DO i=2 %TO &numitem.;
    * Gets entire row of current data;
    %LET nowrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counter.));
    * Grabs specific value for variable from that row;
    %LET newline=%SYSFUNC(GETVARC(&dataitembythresh,

%SYSFUNC(VARNUM(&dataitembythresh,NewParmMax)));
  PUT @1 "! Item &i: Define LCDM parameters present for item &i.";
  PUT @1 "NEW(&newline.);";
  *** Equations per item;
  * Have to do first threshold manually;
  %LET t=1;
    * Gets entire row of current data;
    %LET
nowrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counter.));
    * Grabs specific value for variable from that row;
    %LET threshline=%SYSFUNC(GETVARC(&dataitembythresh,

%SYSFUNC(VARNUM(&dataitembythresh,NewParmAdd)));
  PUT @1 "T&i._&t.=-(&threshline.);" @72 "! Item &i. Thresh &t.";
  %LET counter=%EVAL(&counter.+1); %LET counterd=
%EVAL(&counter.-1);
  * Iteratively through rest of items;
  %DO t=2 %TO &numclass.;
    * Gets entire row of current data;
    %LET
nowrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counter.));
    * Grabs specific value for variable from that row;
    %LET nowitem=%SYSFUNC(GETVARN(&dataitembythresh,

%SYSFUNC(VARNUM(&dataitembythresh,item)));
  * Grabs specific value for variable from that row;
  %LET threshline=%SYSFUNC(GETVARC(&dataitembythresh,

%SYSFUNC(VARNUM(&dataitembythresh,NewParmAdd)));
  * Gets entire row of previous data;
  %LET
pastrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counterd.));

```

```

* Grabs specific value for variable from that row;
%LET pastitem=%SYSFUNC(GETVARN(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,item)))));
    %IF &nowitem.=&pastitem. %THEN %DO;
        PUT @1 "T&i._&t.=-(&threshline.);" @72 "! Item &i.
Thresh &t.";
        %LET counter=%EVAL(&counter.+1); %LET counterd=
%EVAL(&counter.-1);
        %END;
    %END;
*** Ordering constraints per item;
* Gets entire row of current data;
%LET nowrow=%SYSFUNC(FETCHOBS(&dataitembythresh.,&counterd.));
* Grabs specific value for variable from that row;
%LET listmain=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,OrderMainMax)))));
    * Grabs specific value for variable from that row;
    %IF &&itemorder&i.>1 %THEN %DO;
        %LET list12way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order12wayMax)))));
        %LET list22way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order22wayMax)))));
        %END;
        %IF &&itemorder&i.>2 %THEN %DO;
            %LET list13way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order13wayMax)))));
            %LET list23way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order23wayMax)))));
            %LET list33way=%SYSFUNC(GETVARC(&dataitembythresh,
%SYSFUNC(VARNUM(&dataitembythresh,Order33wayMax)))));
        %END;
        * Need to check to see if all terms exist, otherwise get "" in input;
        PUT @1 "! Main effect order constraints"
        / @1 "&listmain.";
        %IF &&itemorder&i.>1 AND %LENGTH(&list12way.)>0 %THEN %DO;
            PUT @1 "! Two-way interaction order constraints"
            / @1 "&list12way."
            / @1 "&list22way."; %END;
        %IF &&itemorder&i.>2 AND %LENGTH(&list13way.)>0 %THEN %DO;
            PUT @1 "! Three-way interaction order constraints"
            / @1 "&list13way."
            / @1 "&list23way."
            / @1 "&list33way."; %END;
        PUT / ;
    %END;

* Mplus OUTPUT command;
    PUT @1 "OUTPUT:"
    / @5 "TECH10; ! Request additional model fit statistics"

```

```

/      ;

* Mplus SAVEDATA command;
  PUT @1 "SAVEDATA: ! Format, name of posterior probabilities of class membership file"
/      @5 "FORMAT = F10.5;"
/      @5 "FILE = &filesave.\respondents_&dataname..dat;"
/      @5 "SAVE = CPROBABILITIES;"
/      ;

run;

* Close datasets used in program;
%LET dataitembyclass=%SYSFUNC(CLOSE(&dataitembyclass.));
%LET dataitembythresh=%SYSFUNC(CLOSE(&dataitembythresh.));
%LET dataclasslevel=%SYSFUNC(CLOSE(&dataclasslevel.));

%MEND WriteMplus;

%WriteMplus;

/* Close log files being printed to if had activated;
   PROC PRINTTO; run; */

%MEND CreateMplusInput;

%MACRO ImportMplusOutput;

* Reading in person-level output;
DATA respondents; RETAIN &IDname.;
  INFILE "&filesave.\respondents_&dataname..dat" DLM=" " ;
  INPUT &itemlist. &IDname. cprob1-cprob&numclass. class;
RUN;

DATA ClassPattern2; SET ClassPattern; DROP class; RUN;

* This is to compute the probability of mastery for each examinee;
PROC IML;
  USE ClassPattern2;
  READ ALL VAR _ALL_ INTO pattmat;

  USE Respondents;
  READ ALL VAR {&IDname.} INTO idvec;
  nexam=NROW(idvec);

  cproblast=CONCAT("cprob",CHAR(1,1,0));
  DO i=2 to &numclass.;
    IF i < 10 THEN cproblast=cproblast||CONCAT("cprob",CHAR(i,1,0));
    ELSE IF i < 100 THEN
cproblast=cproblast||CONCAT("cprob",CHAR(i,2,0));
    ELSE
cproblast=cproblast||CONCAT("cprob",CHAR(i,3,0));
  END;

  USE respondents;
  READ ALL VAR cproblast INTO probmat;

```

```

attprob=probmatt*pattmat;

attprob=idvec||attprob;

cname="{&IDname}";

DO i=1 TO &numatt.;
    IF i < 10 THEN; cname=cname//CONCAT("prob_&attstem.",CHAR(i,1,0));
    ELSE
    cname=cname//CONCAT("prob_&attstem.",CHAR(i,2,0));
END;

CREATE attprob FROM attprob [colname=cname];
APPEND FROM attprob;

QUIT;

* Trimming up respondents file -- cutting out items and adding attribute probabilities;
PROC SORT DATA=respondents; BY &IDname.; RUN;
PROC SORT DATA=attprob; BY &IDname.; RUN;
DATA respondents (DROP=&itemlist.); MERGE respondents attprob; BY &IDname.; RUN; */

* Reading ALL output into a one-variable file;
DATA ReadMplus;
    INFILE "&filesave.\&filename..out" DSD TRUNCOVER LRECL=9000;
    INPUT line $200.;

run;

* Do if used a structural model;
%IF &structon.=1 %THEN %DO;
    * Getting class means;
    DATA ClassMeans; SET ReadMplus;
        * Selecting only class means;
        WHERE INDEX(line, "#");
        * Deleting lines from model input;
        IF INDEX(line, "!") THEN DELETE;
        * Extracting values into separate variables;
        class = INPUT(SUBSTR(line, ANYDIGIT(line)), 8.0);
        CMeanEst = INPUT(SCAN(line,2, " "), 10.3);
        CMeanSE = INPUT(SCAN(line,3, " "), 10.3);
        CMeanZ = INPUT(SCAN(line,4, " "), 10.3);
        CMeanP = INPUT(SCAN(line,5, " "), 10.3);
    DROP line;
RUN;
    * Getting NEW parameters for structural model;
    DATA StrucParms; SET ReadMplus;
        * Selecting only class means;
        WHERE INDEX(line, "G") AND INDEX(line, "_");
        * Deleting lines from model input;
        IF INDEX(line, ")") THEN DELETE;
        * Extracting values into separate variables;
        parm = SCAN(line,1, " ");
        Order = INPUT(SUBSTR(line, ANYPUNCT(line)+1, 1), 8.0);
        Atts = INPUT(SUBSTR(line, ANYPUNCT(line)+2), 8.0);
        IF Order = 0 THEN Atts = 0;

```

```

                StrucEst = INPUT(SCAN(line,2, " "), 10.3);
                StrucSE = INPUT(SCAN(line,3, " "), 10.3);
                StrucZ = INPUT(SCAN(line,4, " "), 10.3);
                StrucP = INPUT(SCAN(line,5, " "), 10.3);
        DROP line;
    RUN;
%END;

* Getting estimated model class counts and proportions;
DATA ClassCounts; SET ReadMplus;
    temp = SUBSTR(line, 1, FIND(line, " "));
    %DO i=1 %TO &numclass.;
        IF temp=&i. THEN DO;
            class = INPUT(temp, 3.0);
            estcount = INPUT(SCAN(line,2, " "), 10.3);
            estprop = INPUT(SCAN(line,3, " "), 10.3);
        END;
    %END;
    IF NMISS(estcount,estprop)>0 THEN DELETE;
    KEEP class estcount estprop;
RUN;
DATA ClassCounts; SET ClassCounts;
    IF _N_ > &numclass. THEN DELETE;
RUN;

* Merge class attribute patterns into class counts;
DATA ClassCounts; MERGE ClassCounts ClassPattern; BY class; RUN;

* Calculate reliabilities;
PROC IML;
    USE ClassCounts;
    READ ALL VAR {estprop} INTO cprob;

    cname="prob_&attstem.1";
    DO i=2 TO &numatt.;
        IF i < 10 THEN; cname=cname||CONCAT("prob_&attstem.",CHAR(i,1,0));
        ELSE
            cname=cname||CONCAT("prob_&attstem.",CHAR(i,2,0));
    END;

    USE Respondents;
    READ ALL VAR cname INTO rprob;
    nexam=NROW(rprob);

    USE ClassPattern2;
    READ ALL VAR _ALL_ INTO pattmat;

    attprob=cprob`*pattmat;

    reliability=J(1,&numatt.,0);
    DO i=1 TO nexam;
        DO j=1 to &numatt.;
            vartrue=attprob[1,j]*(1-attprob[1,j]);
            varerror=rprob[i,j]*(1-rprob[i,j]);
            reliability[1,j]=reliability[1,j]+(vartrue/(vartrue+varerror));
        END;
    END;

```

```

END;
reliability=reliability/nexam;

attnumber=J(&numatt.,1,0);
DO i=1 TO &numatt.;
    attnumber[i,1]=i;
END;
outmat=attnumber||attprob`||reliability`;

CREATE AttProbRel FROM outmat [colname = {Attribute Probability Reliability}];
APPEND FROM outmat;

QUIT;

* Needed for parsing item number out of threshold label;
    %IF &numitem.<10 %THEN %LET digit=1;
%ELSE %IF &numitem.<100 %THEN %LET digit=2;
%ELSE
    %LET digit=3;

* Getting thresholds per class per item;
DATA Thresholds; SET ReadMplus;
    * Selecting only thresholds;
        WHERE INDEX(line, "$");
    * Deleting lines with comments (from model input);
        IF INDEX(line, "!") THEN DELETE;
    * Extracting values into separate variables;
        Item = INPUT(SUBSTR(line, 6, &digit.), 8.0);
        Thresh = INPUT(SUBSTR(line, ANYPUNCT(line)+1, FIND(line, " ")), 8.0);
        ThreshEst = INPUT(SCAN(line,2, " "), 10.3);
        ThreshSE = INPUT(SCAN(line,3, " "), 10.3);
        ThreshZ = INPUT(SCAN(line,4, " "), 10.3);
        ThreshP = INPUT(SCAN(line,5, " "), 10.3);

RUN;
DATA Thresholds; RETAIN class item thresh; SET Thresholds;
    * Adding index for class;
    IF _N_=1 THEN class = 1;
    IF item > LAG1(item) THEN class=class; ELSE class=class+1;
    DROP line;

RUN;
* Merge class pattern and item attribute pattern info;
DATA Thresholds; MERGE Thresholds ItembyClass
    (KEEP= class item &itemstem. itematt1-itematt&numatt. classatt1-classatt&numatt.);
    BY class item; RUN;
DATA Thresholds; RETAIN class item &itemstem. thresh; SET Thresholds; RUN;

* Getting NEW parameters for item model;
DATA ItemParms; SET ReadMplus;
    * Selecting only class means;
    WHERE INDEX(line, "L") AND INDEX(line, "_");
    * Deleting lines from model input;
    IF INDEX(line, ")") OR INDEX(line, ">") OR INDEX(line, "FILE") THEN DELETE;
    parm = SCAN(line,1, " ");
    Item = INPUT(SUBSTR(line, ANYDIGIT(line), ANYPUNCT(line)-2), 8.0);
    Order = INPUT(SUBSTR(line, ANYPUNCT(line)+1, 1), 8.0);
    Atts = INPUT(SUBSTR(line, ANYPUNCT(line)+2), 8.0);
    IF Order = 0 THEN Atts = 0;

```

```

ItemEst = INPUT(SCAN(line,2, " "), 10.3);
ItemSE = INPUT(SCAN(line,3, " "), 10.3);
ItemZ = INPUT(SCAN(line,4, " "), 10.3);
ItemP = INPUT(SCAN(line,5, " "), 10.3);
DROP line;

RUN;
* Merge back original item labels;
DATA ItemParms; MERGE ItemParms Qmatrix; BY Item; RUN;
DATA ItemParms; RETAIN parm item &itemstem. order; SET ItemParms; RUN;

* Clearing extra datasets;
PROC DATASETS LIB=WORK NOLIST;
    DELETE Classlevel ClassPattern ClassPattern2 ItembyClass ItembyThresh kernel
        AttProb;
RUN; QUIT;

%MEND ImportMplusOutput;

%MACRO WriteData;
* Run to export summary datasets to permanent library named above;
DATA folder.&filename._Qmatrix; SET Qmatrix; RUN;
DATA folder.&filename._ReadMplus; SET ReadMplus; RUN;
DATA folder.&filename._Respondents; SET Respondents; RUN;
DATA folder.&filename._ClassCounts; SET ClassCounts; RUN;
DATA folder.&filename._Thresholds; SET Thresholds; RUN;
DATA folder.&filename._ItemParms; SET ItemParms; RUN;
DATA folder.&filename._AttProbRel; SET AttProbRel; RUN;

* Run to export to an excel workbook (one dataset per worksheet);
PROC EXPORT DATA=Qmatrix OUTFILE="&filesave.\&filename..xls" DBMS=EXCEL
REPLACE; SHEET="Qmatrix"; RUN;
PROC EXPORT DATA=ReadMplus OUTFILE="&filesave.\&filename..xls"
DBMS=EXCEL REPLACE; SHEET="ReadMplus"; RUN;
PROC EXPORT DATA=Respondents OUTFILE="&filesave.\&filename..xls"
DBMS=EXCEL REPLACE; SHEET="Respondents"; RUN;
PROC EXPORT DATA=ClassCounts OUTFILE="&filesave.\&filename..xls"
DBMS=EXCEL REPLACE; SHEET="ClassCounts"; RUN;
PROC EXPORT DATA=Thresholds OUTFILE="&filesave.\&filename..xls"
DBMS=EXCEL REPLACE; SHEET="Thresholds"; RUN;
PROC EXPORT DATA=ItemParms OUTFILE="&filesave.\&filename..xls"
DBMS=EXCEL REPLACE; SHEET="ItemParms"; RUN;
PROC EXPORT DATA=AttProbRel OUTFILE="&filesave.\&filename..xls"
DBMS=EXCEL REPLACE; SHEET="Attribute Reliability"; RUN;

* Additional output if using a structural model;
%IF &structon.=1 %THEN %DO;
* Run to export summary datasets to permanent library named above;
DATA folder.&filename._Classmeans; SET ClassMeans; RUN;
DATA folder.&filename._StrucParms; SET StrucParms; RUN;
* Run to export to an excel workbook (one dataset per worksheet);
PROC EXPORT DATA=ClassMeans OUTFILE="&filesave.\&filename..xls"
DBMS=EXCEL REPLACE; SHEET="ClassMeans"; RUN;

```

```
PROC EXPORT DATA=StrucParms OUTFILE= "&filesave.\&filename.xls"  
DBMS=EXCEL REPLACE; SHEET="StrucParms"; RUN;  
%END;
```

```
%MEND WriteData;
```

Anexo 3: Función de Probabilidad Según Modelo 3-PL

Patrón	Theta	Prob 1	Prob 2	Prob 3	Prob 4	Prob 5	Prob 6	Prob 7
1000000	-1,183	0,385	0,154	0,141	0,544	0,381	0,418	0,044
0001000	-1,159	0,385	0,156	0,145	0,544	0,388	0,421	0,047
1001000	-1,149	0,385	0,157	0,147	0,544	0,391	0,423	0,049
0000010	-0,96	0,387	0,174	0,180	0,548	0,450	0,450	0,086
1000010	-0,931	0,387	0,177	0,186	0,549	0,459	0,454	0,094
0001010	-0,922	0,388	0,177	0,188	0,549	0,462	0,456	0,097
0101000	-0,902	0,388	0,179	0,192	0,550	0,468	0,458	0,102
1001010	-0,884	0,388	0,181	0,195	0,550	0,474	0,461	0,108
0000100	-0,752	0,394	0,194	0,223	0,555	0,516	0,480	0,156
0010000	-0,749	0,394	0,194	0,224	0,555	0,517	0,481	0,157
0101010	-0,725	0,395	0,197	0,230	0,556	0,525	0,484	0,168
0001100	-0,713	0,396	0,198	0,232	0,556	0,529	0,486	0,173
0011000	-0,709	0,396	0,198	0,233	0,556	0,530	0,487	0,175
1000100	-0,659	0,400	0,203	0,245	0,559	0,546	0,494	0,199
1010000	-0,654	0,401	0,204	0,246	0,559	0,547	0,495	0,202
0000110	-0,619	0,405	0,208	0,255	0,560	0,558	0,500	0,221
0010010	-0,616	0,405	0,208	0,255	0,560	0,559	0,500	0,222
0100100	-0,608	0,406	0,209	0,257	0,561	0,562	0,502	0,227
1101010	-0,597	0,407	0,210	0,260	0,561	0,565	0,503	0,233
0001110	-0,58	0,410	0,212	0,264	0,562	0,571	0,506	0,243
1001100	-0,571	0,411	0,213	0,267	0,563	0,573	0,507	0,248
0101100	-0,569	0,411	0,213	0,267	0,563	0,574	0,507	0,250
1011000	-0,564	0,412	0,214	0,268	0,563	0,576	0,508	0,253
0110010	-0,502	0,424	0,220	0,284	0,566	0,595	0,517	0,292
0010100	-0,495	0,426	0,221	0,286	0,567	0,597	0,518	0,297
0101110	-0,468	0,432	0,224	0,293	0,568	0,605	0,522	0,315
0011100	-0,459	0,435	0,225	0,296	0,569	0,608	0,523	0,322
0010110	-0,412	0,449	0,230	0,309	0,572	0,622	0,530	0,356
0000001	-0,405	0,452	0,231	0,311	0,572	0,624	0,531	0,361
0001001	-0,371	0,465	0,235	0,320	0,574	0,635	0,536	0,386
1000110	-0,35	0,474	0,238	0,326	0,576	0,641	0,539	0,402
0100001	-0,331	0,483	0,240	0,331	0,577	0,646	0,542	0,417
0100011	-0,272	0,515	0,247	0,349	0,582	0,663	0,551	0,464
0010001	-0,266	0,519	0,248	0,350	0,582	0,665	0,552	0,469
0001101	-0,236	0,538	0,251	0,359	0,585	0,674	0,556	0,493
0000111	-0,216	0,552	0,254	0,365	0,586	0,679	0,559	0,509
0010011	-0,214	0,554	0,254	0,366	0,587	0,680	0,559	0,511
1001110	-0,192	0,570	0,257	0,372	0,588	0,686	0,562	0,528
0001111	-0,185	0,576	0,257	0,375	0,589	0,688	0,563	0,534

1101100	-0,174	0,585	0,259	0,378	0,590	0,691	0,565	0,543
0100111	-0,163	0,593	0,260	0,381	0,591	0,694	0,567	0,552
0010101	-0,158	0,598	0,261	0,383	0,592	0,695	0,567	0,556
0101111	-0,132	0,619	0,264	0,391	0,594	0,702	0,571	0,576
0011101	-0,127	0,624	0,265	0,392	0,595	0,703	0,572	0,580
1100110	-0,126	0,625	0,265	0,393	0,595	0,704	0,572	0,581
1110010	-0,124	0,626	0,265	0,393	0,595	0,704	0,572	0,582
0011111	-0,081	0,665	0,270	0,407	0,599	0,716	0,578	0,616
0110111	-0,065	0,679	0,272	0,412	0,601	0,720	0,581	0,628
1101110	-0,048	0,695	0,275	0,417	0,603	0,724	0,583	0,640
1011100	-0,037	0,705	0,276	0,420	0,604	0,727	0,585	0,649
0111111	-0,03	0,711	0,277	0,423	0,605	0,729	0,586	0,654
1010110	-0,019	0,721	0,278	0,426	0,606	0,731	0,587	0,662
1000001	-0,01	0,729	0,280	0,429	0,607	0,734	0,589	0,668
1011110	0,053	0,784	0,288	0,449	0,614	0,749	0,597	0,712
1111100	0,061	0,790	0,289	0,452	0,615	0,751	0,599	0,717
1001001	0,062	0,791	0,289	0,452	0,615	0,751	0,599	0,717
1110110	0,069	0,797	0,290	0,454	0,616	0,753	0,600	0,722
1000011	0,07	0,798	0,290	0,454	0,616	0,753	0,600	0,723
1001011	0,149	0,855	0,301	0,480	0,627	0,772	0,611	0,771
1111110	0,149	0,855	0,301	0,480	0,627	0,772	0,611	0,771
1100011	0,157	0,861	0,302	0,482	0,628	0,773	0,612	0,775
1000101	0,164	0,865	0,303	0,485	0,629	0,775	0,613	0,779
1000111	0,251	0,911	0,315	0,513	0,642	0,794	0,625	0,824
1010011	0,253	0,912	0,315	0,513	0,642	0,794	0,625	0,824
1100101	0,26	0,915	0,316	0,516	0,643	0,796	0,626	0,828
1001101	0,266	0,918	0,317	0,518	0,644	0,797	0,627	0,830
1011001	0,268	0,919	0,317	0,518	0,644	0,797	0,627	0,831
1100111	0,364	0,951	0,330	0,549	0,660	0,816	0,641	0,870
1110011	0,366	0,952	0,331	0,550	0,660	0,817	0,641	0,871
1010101	0,376	0,954	0,332	0,553	0,662	0,819	0,642	0,875
1001111	0,391	0,958	0,334	0,558	0,665	0,822	0,644	0,880
1011011	0,395	0,959	0,335	0,559	0,665	0,822	0,645	0,881
1101101	0,406	0,961	0,336	0,563	0,667	0,824	0,646	0,885
1111001	0,409	0,962	0,337	0,563	0,668	0,825	0,647	0,886
1010111	0,505	0,978	0,351	0,594	0,686	0,842	0,659	0,914
1101111	0,582	0,986	0,362	0,618	0,701	0,855	0,670	0,931
1111011	0,586	0,986	0,363	0,619	0,701	0,855	0,670	0,932
1011101	0,604	0,987	0,366	0,624	0,705	0,858	0,672	0,936
1110111	0,68	0,992	0,377	0,647	0,720	0,870	0,682	0,949
1011111	0,848	0,997	0,403	0,695	0,756	0,892	0,703	0,970

1111101	0,875	0,997	0,407	0,702	0,762	0,895	0,706	0,972
1111111	1,172	1,000	0,454	0,776	0,825	0,926	0,741	0,989

Anexo 4: Clase Latente por Patrón de Respuesta

El presente cuadro muestra la relación entre cada patrón de respuesta y la probabilidad de dominio a cada habilidad latente; en consecuencia, muestra la relación del patrón de respuesta con el perfil de clase al cual pertenece.

<i>Patrón de Respuesta</i>	<i>Probabilidad α_1</i>	<i>Probabilidad α_2</i>	<i>Probabilidad α_3</i>	<i>Perfil de Clase</i>
0000001	0,00798	0	0,16122	[000]
0000010	0,00036	0	0,00728	[000]
0000110	0,00754	0	0,01441	[000]
0000111	0,00754	0	0,01441	[000]
0001000	0,00032	0	0,00218	[000]
0001001	0,02622	0	0,17664	[000]
0001010	0,00121	0	0,00812	[000]
0001100	0,00681	0	0,00866	[000]
0001101	0,36227	0,00001	0,46078	[000]
0001110	0,02482	0	0,03157	[000]
0001111	0,59855	0,00002	0,7613	[101]
0010001	0,03475	0	0,70225	[001]
0010010	0,00409	0	0,08255	[000]
0010011	0,04444	0,00001	0,89798	[001]
0010100	0,02401	0	0,04587	[000]
0010101	0,43164	0,00002	0,82468	[101]
0010110	0,07963	0	0,15214	[000]
0011100	0,07611	0	0,09681	[000]
0011101	0,71776	0,00003	0,91294	[101]
0011111	0,76664	0,00003	0,97509	[101]
0100001	0,02357	0	0,47625	[000]
0100011	0,03822	0	0,77238	[001]
0100100	0,00953	0	0,0182	[000]
0100111	0,45605	0,00002	0,87127	[101]
0101000	0,00152	0	0,01026	[000]
0101010	0,00553	0	0,03727	[000]
0101100	0,0312	0	0,03968	[000]
0101110	0,10506	0	0,13363	[000]
0101111	0,73736	0,00003	0,93785	[101]
0110010	0,01478	0	0,29857	[000]
0110111	0,5172	0,00002	0,9881	[101]
0111111	0,782	0,00003	0,99463	[101]
1000000	0,00598	0,18943	0,01228	[000]
1000001	0,11268	0,5594	0,54877	[011]
1000010	0,02136	0,21872	0,04393	[000]
1000011	0,16799	0,7465	0,81803	[011]
1000100	0,11247	0,27485	0,05556	[000]
1000101	0,72818	0,84339	0,85196	[111]
1000110	0,80985	0,918	0,94751	[111]
1000110	0,31532	0,44945	0,15577	[000]
1000111	0,80985	0,918	0,94751	[111]
1001000	0,01972	0,20045	0,01787	[000]
1001001	0,29839	0,64508	0,64025	[011]

1001010	0,06813	0,25543	0,06173	[000]
1001011	0,40338	0,80939	0,86551	[011]
1001100	0,29792	0,42363	0,13093	[000]
1001101	0,89971	0,92254	0,93646	[111]
1001110	0,60664	0,67809	0,26662	[110]
1001111	0,93449	0,95131	0,97266	[111]
1010000	0,02986	0,27223	0,13236	[000]
1010011	0,19908	0,85842	0,98219	[011]
1010101	0,83184	0,9391	0,98605	[111]
1010110	0,65129	0,76776	0,69365	[111]
1010111	0,83984	0,94641	0,99551	[111]
1011000	0,09342	0,31806	0,18249	[000]
1011001	0,43985	0,8708	0,95622	[011]
1011011	0,45427	0,89346	0,9875	[011]
1011100	0,68493	0,74447	0,64897	[111]
1011101	0,94306	0,95847	0,99449	[111]
1011110	0,86215	0,89102	0,81689	[111]
1011111	0,94611	0,96099	0,99771	[111]
1100011	0,19396	0,83995	0,9551	[011]
1100101	0,81524	0,92377	0,96458	[111]
1100110	0,50915	0,63309	0,46609	[111]
1100111	0,83541	0,94221	0,98842	[111]
1101010	0,14138	0,37647	0,23735	[000]
1101100	0,51099	0,60027	0,41614	[000]
1101101	0,93661	0,95313	0,98586	[111]
1101110	0,77646	0,81962	0,63234	[111]
1101111	0,94443	0,95959	0,99409	[111]
1110010	0,15082	0,68467	0,72735	[011]
1110011	0,20174	0,86796	0,99618	[011]
1110110	0,78931	0,89853	0,91462	[111]
1110111	0,84204	0,9485	0,99904	[111]
1111001	0,45516	0,89523	0,99042	[011]
1111011	0,45836	0,90024	0,99733	[011]
1111100	0,87052	0,89833	0,89739	[111]
1111101	0,9463	0,96116	0,99883	[111]
1111110	0,92617	0,94438	0,95476	[111]
1111111	0,94695	0,96169	0,99951	[111]

