

**PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ**

**ESCUELA DE POSGRADO**



**MODELO G-DINA APLICADO AL DIAGNÓSTICO DE  
DESÓRDENES MENTALES**

TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAGÍSTER EN  
ESTADÍSTICA

**Presentado por:**

**Pilar Denisse Villena Guzmán**

**Asesor: Mg. Enver Gerald Tarazona Vargas**

**Miembros del jurado:**

**Dr. Víctor Giancarlo Sal y Rosas Celi**

**Dr. Luis Hilmar Valdivieso Serrano**

LIMA, PERÚ

2018

## Dedicatoria

Este trabajo está dedicado a mis padres, por su sacrificio y su amor infinito e incondicional. Y a mis hermanas, porque somos como una sola, y sin ellas no estar'ia completa.



*Uno es valiente cuando, sabiendo que ha perdido ya antes de empezar, empieza a pesar de todo y sigue hasta el final pase lo que pase. Uno vence raras veces, pero alguna vez vence.*

Harper Lee

## **Agradecimientos**

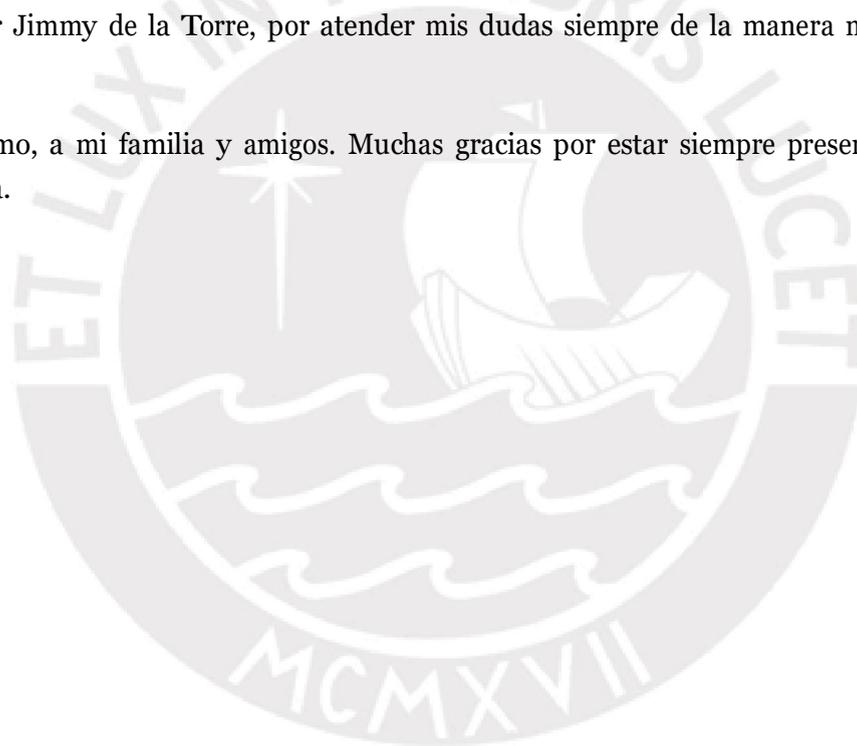
Quisiera expresar mi más profundo agradecimiento al profesor Enver Tarazona por su dedicación y apoyo a lo largo de este camino.

Al profesor Luis Valdivieso, por sus sugerencias e ideas para este trabajo.

Al doctor Johann Vega Dienstmaier por su invaluable colaboración para la realización de este trabajo.

Al profesor Jimmy de la Torre, por atender mis dudas siempre de la manera más cordial y amable.

Y por último, a mi familia y amigos. Muchas gracias por estar siempre presente de una u otra forma.



## Resumen

Actualmente, uno de los modelos de diagnóstico cognitivo (MDC) más usados es el modelo DINA. Sin embargo, este modelo presenta varias restricciones que hacen que en muchas ocasiones, no sea el que mejor se ajusta a la realidad. En ese contexto, nace una generalización del modelo DINA, denominado G-DINA (Generalized deterministic input, noisy and gate). En el presente estudio se presentan los fundamentos y propiedades del modelo G-DINA y su aplicación en un área en el que su uso todavía no es muy común: la psicología. Así, se evaluaron los resultados de una muestra de pacientes de un hospital general de Lima a los que se les aplicó el test SRQ-18 que evalúa la presencia de desórdenes mentales. Se muestra el proceso de selección del mejor modelo para cada ítem, los resultados de los parámetros obtenidos, los diagnósticos para los 10 primeros pacientes y una distribución de los perfiles de estos pacientes. Finalmente se presenta un estudio de simulación que tiene por finalidad estudiar el efecto del tamaño de muestra en la estimación de los parámetros en el contexto de la aplicación de este estudio.

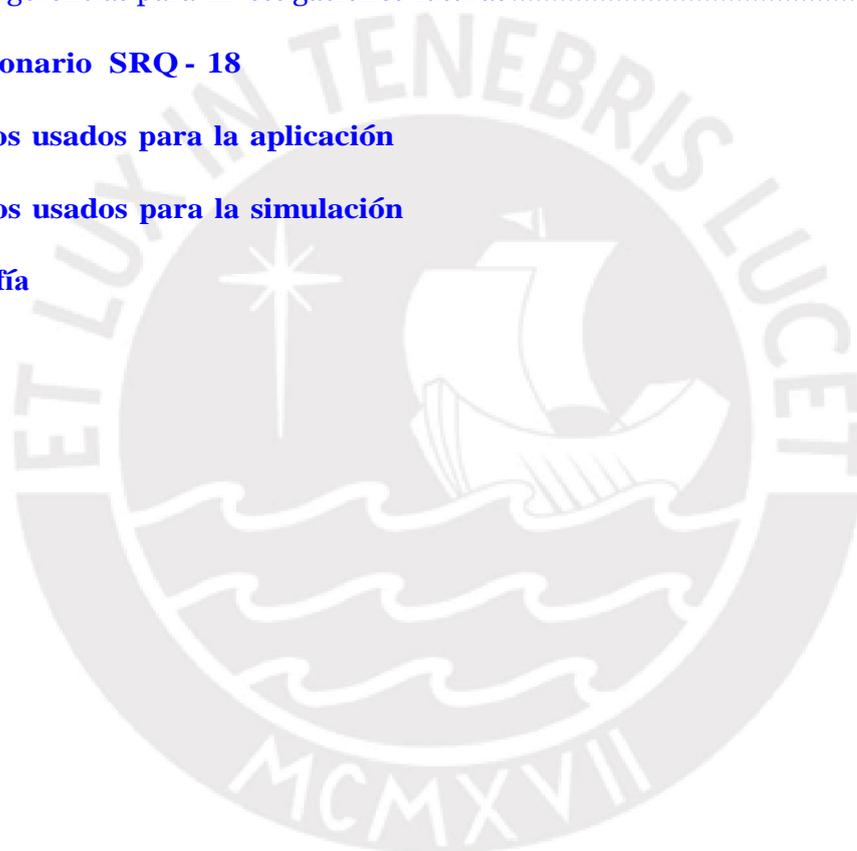
**Palabras-clave:** G-DINA, Modelos de diagnóstico Cognitivo, estimación de parámetros, Test de Wald, Disimilaridad.

# Índice general

<b>Lista de Abreviaturas</b>	<b>VII</b>
<b>Índice de figuras</b>	<b>VIII</b>
<b>Índice de cuadros</b>	<b>IX</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Consideraciones Preliminares.....	1
1.2. Objetivos.....	2
1.3. Organización del Trabajo .....	2
<b>2. El modelo G-DINA</b>	<b>4</b>
2.1. Formulación del modelo .....	4
2.2. Ejemplo.....	6
2.3. Casos especiales.....	8
2.3.1. Modelo DINA.....	8
2.3.2. Modelo DINO.....	9
2.3.3. Modelos aditivos .....	9
2.4. Función de verosimilitud .....	10
<b>3. Estimación del modelo</b>	<b>11</b>
3.1. Estimador de máxima verosimilitud marginal .....	11
3.2. Matriz de Diseño .....	15
3.3. Comparación de modelos .....	16
3.3.1. Test de Wald .....	16
3.3.2. Disimilitud entre los modelos reducidos .....	17
<b>4. Estudio de Simulación</b>	<b>18</b>
4.1. Simulación de los datos y estimación de parámetros .....	19
4.2. Resultados .....	19
4.2.1. Ítems que miden un atributo .....	19
4.2.2. Ítems que miden dos atributos .....	23
4.2.3. Ítems que miden tres atributos .....	26
<b>5. Aplicación</b>	<b>28</b>
5.1. Introducción .....	28
5.2. Instrumentos de medición .....	29

5.3. Muestra.....	29
5.4. Análisis de los datos .....	30
5.5. Selección del modelo para cada ítem .....	32
5.6. Estimación de los parámetros de los ítems .....	34
5.7. Estimación de perfiles .....	36
5.8. Diagnóstico de los pacientes .....	37
5.9. Comparación con otros modelos .....	37
5.10 .....	38
Conclusiones .....	38
<b>6. Conclusiones</b> .....	<b>39</b>
6.1. Conclusiones.....	39
6.2. Sugerencias para investigaciones futuras.....	40
<b>A. Cuestionario SRQ - 18</b> .....	<b>41</b>
<b>B. Códigos usados para la aplicación</b> .....	<b>43</b>
<b>C. Códigos usados para la simulación</b> .....	<b>45</b>
<b>Bibliografía</b> .....	<b>51</b>

**Índice general**



## Lista de Abreviaturas

G-DINA	<i>Generalized Deterministic Inputs, Noisy “And” gate.</i>
DINO	<i>Deterministic Inputs, Noisy “Or” gate.</i>
DINA	<i>Deterministic Inputs, Noisy “And” gate.</i>
MDC	Modelos de Diagnóstico Cognitivo.
CDM	<i>Cognitive Diagnosis Models.</i>
A-CDM	<i>Additive Cognitive Diagnosis Models.</i>
LLM	<i>Linear Logistic Model.</i>
R-RUM	<i>Reduced Reparametrized Unified Model.</i>
EM	<i>Expectation-Maximization.</i>
SRQ	<i>Self Reporting Questionnaire.</i>
AIC	Criterio de Información de Akaike.
BIC	Criterio de Información Bayesiano.
MMLE	Estimación de máxima verosimilitud marginal.
RMSE	<i>Root mean squared error.</i>
Minsa	Ministerio de Salud.
WHO	<i>World Health Organization</i>

# Índice de figuras

A.1. Cuestionario SRQ 18.....42



## Índice de cuadros

4.1. Comparación de parámetros de intercepto de ítems que requieren un atributo	21
4.2. Comparación de parámetros de efecto principal de ítems que requieren un atributo .....	22
4.3. Comparación de parámetros de intercepto de ítems que requieren dos atributos	23
4.4. Comparación de parámetros del 1er efecto principal de ítems que requieren dos atributos .....	23
4.5. Comparación de parámetros del 2do efecto principal de ítems que requieren dos atributos .....	24
4.6. Comparación de parámetros del efecto interacción de ítems que requieren dos atributos.....	24
4.7. Comparación del intercepto $P(00)$ : probabilidad de responder afirmativamente al ítem dado que no se posee ninguno de los atributos requeridos .....	25
4.8. Comparación de los parámetros $P(10)$ : Probabilidad de responder afirmativamente al ítem dado que se posee el primer atributo requerido.....	25
4.9. Comparación de los parámetros $P(01)$ : Probabilidad de responder afirmativamente al ítem dado que se posee el segundo atributo requerido .....	26
4.10. Comparación de los parámetros $P(11)$ : Probabilidad de responder afirmativamente al ítem dado que se posee los dos atributos requeridos .....	26
4.11. Comparación de los parámetros del ítem que evalúa 3 atributos .....	27
5.1. Descripción de los síndromes evaluados en la prueba SRQ-18 .....	30
5.2. Matriz Q.....	31
5.3. Número de ítems evaluados en cada síndrome .....	31
5.4. Valores del Estadístico de Wald para los 5 modelos evaluados.....	32
5.5. Los p-valores del test de Wald .....	33
5.6. Disimilitudes entre el modelo G-DINA y los modelos reducidos .....	33
5.7. Selección del mejor modelo para cada ítem .....	34
5.8. Parámetros estimados de los ítems que miden un desorden .....	35
5.9. Parámetros estimados de los ítems que miden dos desórdenes .....	35
5.10. Parámetros estimados del ítem que mide tres desórdenes (Ítem 3) .....	36
5.11. Perfil de desórdenes mentales y prevalencia .....	36
5.12. Probabilidades de presentar los desórdenes para 10 pacientes .....	37
5.13. Resultados del ajuste de los modelos a los datos.....	38

## Capítulo 1

# Introducción

### 1.1. Consideraciones Preliminares

Un correcto diagnóstico del estado o situación de una persona es esencial en diversos campos de investigación como la educación, psicología, salud, etc. Para este fin, existen modelos como la Teoría Clásica de los Tests o la Teoría de Respuesta al Ítem que nos permiten estimar un puntaje que represente la presencia de cierto atributo o habilidad. Existen también los modelos de diagnóstico cognitivo (MDC), que son nuestro objeto de estudio, los cuáles nos permiten ahondar un poco más en la identificación de los perfiles de los atributos (como la “*habilidad*” en el campo de la educación, y el “*desorden mental*” en el campo de la psiquiatría) de los encuestados en función a sus respuestas en el test (Sosa, 2017).

Uno de los MDC más usados es el modelo DINA (*Deterministic Input Noisy and gate*). Este modelo divide, para cada ítem  $j$  (o pregunta en una prueba o test), a las personas en dos grupos, el primero compuesto por los individuos que presenten todos los atributos requeridos para dar una respuesta correcta a este ítem, y el segundo formado por los individuos que carecen de al menos uno de los atributos requeridos. Con este modelo, se asume que todos los individuos en el mismo grupo tienen la misma probabilidad de responder al ítem correctamente. Sin embargo, esta asunción no siempre es cierta para el segundo grupo, debido a que los individuos en él tienen diferentes grados de deficiencia con respecto a los atributos requeridos. El modelo G-DINA (del acrónimo en inglés *Generalized Deterministic Input Noisy and gate*) cuestiona la suposición de igual probabilidad de responder correctamente para los individuos del segundo grupo y propone una generalización. Así, en lugar de formar solo dos grupos, el modelo G-DINA forma  $2^{K_j^*}$  grupos, donde  $K_j^*$  es el número de atributos requeridos para el ítem  $j$ . Por ejemplo cuando  $K_j^* = 2$ , en lugar de 2, se crean 4 grupos latentes, donde cada uno de ellos pueden tener diferentes probabilidades de éxito.

Respecto a su aplicación, los MDC han sido en su mayoría desarrollados y aplicados en el área de la educación; sin embargo, no existe nada inherente a su construcción que no le permita ser aplicados en otros campos como la psicología y la psiquiatría. Así en los últimos años, se han desarrollado estos modelos para evaluar patologías en el juego, desórdenes mentales, entre otros (De la Torre et al., 2015).

Por otra parte, según un informe del Instituto de Análisis y Comunicación “*Integración*”, en el Perú los desórdenes mentales son la principal causa de años de vida saludables perdidos y se estima que existen 4 millones de peruanos con alguna enfermedad mental de los cuales solo el 20% recibe tratamiento adecuado (Integración, 2016). Dicho informe indica que se

necesitan más mediciones del estado de la salud mental en el país para que de esa manera sea un problema socialmente reconocido. Es por esta razón que es importante la actualización de los modelos de diagnóstico en este campo.

## 1.2. Objetivos

El objetivo general de la tesis es estudiar las propiedades del modelo G-DINA y aplicarlo a un conjunto de datos reales. De manera específica:

- Presentar los fundamentos y propiedades del modelo G-DINA.
- Presentar el proceso de estimación de los parámetros del modelo G-DINA.
- Realizar las rutinas en el software R para la estimación de los parámetros del modelo G-DINA.
- Realizar un estudio de simulación a fin de evaluar el efecto del número de datos en el contexto de nuestra aplicación.
- Aplicar el modelo a un conjunto de datos reales que evalúa la presencia de desórdenes mentales en un grupo de pacientes de consultorio externo de psiquiatría del Hospital Nacional Cayetano Heredia (Lima, Perú).

## 1.3. Organización del Trabajo

El presente documento ha sido dividido en 6 capítulos. En el segundo capítulo se presentan los conceptos y la formulación del modelo G-DINA (Generalized Deterministic Input Noisy and gate), para luego describir los casos más usados como el modelo DINA, DINO y los modelos aditivos A-CDM, LLM y R-RUM, finalizando con la presentación de su función de verosimilitud.

En el capítulo 3 se muestra la estimación del modelo para el cual se utiliza el algoritmo de Esperanza y Maximización (EM), a fin de maximizar la función de verosimilitud marginal. Además, se describen los pasos necesarios para convertir los parámetros estimados en sus componentes, para luego presentar los criterios para la selección del mejor modelo para cada ítem.

En el capítulo 4 se presenta un estudio de simulación basándonos en el contexto de los datos reales usados para este trabajo, evaluando la precisión de la estimación bajo distintos tamaños de muestra. Se presentan los criterios de evaluación y los resultados para cada ítem.

En el capítulo 5 se presentan los resultados de aplicar el modelo G-DINA a un grupo de pacientes del hospital Cayetano Heredia a los que se les aplicó un test de diagnóstico de desórdenes mentales. Además, se presenta el proceso de selección del mejor modelo para cada uno de los ítems usando el test de Wald y las disimilaridades, para luego mostrar los parámetros estimados con su capacidad de discriminación. Se presenta también un diagnóstico de desórdenes mentales para los primeros 10 individuos y la distribución del estado de los pacientes.

En el sexto capítulo se detallan las conclusiones que se obtuvieron a lo largo de todas las etapas de esta investigación. Por último, se presentan algunas recomendaciones para investigaciones futuras.

Se incluyen 3 anexos, donde se presenta el cuestionario usado para este estudio, el SRQ-18, y los códigos en R usados en la aplicación y la simulación.



## Capítulo 2

### El modelo G-DINA

#### 2.1. Formulación del modelo

El modelo generalizado de entrada determinística con ruido y salida (G-DINA por sus siglas en inglés) parte de un contexto en el que  $N$  individuos denotados por  $i = 1, 2, \dots, N$ , son examinados mediante un test con  $J$  elementos de evaluación (que llamaremos ítems) denotados por  $j = 1, 2, \dots, J$ , cada uno de los cuales requieren la presencia de  $K_j$  atributos para poder ser correctamente respondidos, donde  $K_j \leq K$ , siendo  $K$  la totalidad de atributos evaluados por el test.

Este test es construido por expertos en el área de interés que también definen una matriz  $\mathbf{Q} = [q_{jk}]$  que nos indica los atributos requeridos por cada ítem  $j$ , donde  $k = 1, 2, \dots, K$ . Así,  $q_{jk}$  toma el valor de 1 si el atributo  $k$  es relevante para responder correctamente el ítem  $j$  y 0 en caso contrario.

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1K} \\ q_{21} & q_{22} & \dots & q_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ q_{J1} & q_{J2} & \dots & q_{JK} \end{bmatrix}$$

Por ejemplo, para un test con 10 ítems que evalúan la presencia de 4 atributos, la matriz  $\mathbf{Q}$  podría estar dada por:

$$\mathbf{Q} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Aquí la fila 1 indica que para responderse correctamente el ítem 1, es necesario que el individuo posea el segundo y cuarto atributo, mas no el primero y tercero.

De los resultados del test obtenemos una matriz observable binaria  $\mathbf{x}=[x_{ij}]$  de orden  $N \times J$  que representa las respuestas de los  $N$  entrevistados a los  $J$  ítems. De tal manera,  $x_{ij}$  toma el valor de 1 si el individuo respondió correctamente al ítem  $j$  y 0 en caso contrario.

$$\mathbf{x} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1J} \\ x_{21} & x_{22} & \dots & x_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NJ} \end{bmatrix}$$

Para cada individuo  $i$  se asume que existe un vector latente  $a_i = [a_{i1}, a_{i2}, \dots, a_{iK}]$  tal que  $a_{ik} = 1$  si el individuo  $i$  posee el atributo  $k$ , y  $a_{ik} = 0$  en caso contrario. Estos, que llamaremos perfiles latentes, son desconocidos y su estimación es uno de los objetivos de este modelo.

Para cada ítem  $j$ , se separan a los individuos en  $2^{K_j^*}$  grupos latentes ( $l$ ), donde  $K_j^* = \sum_{k=1}^j q_{jk}$  representa el número de atributos requeridos para el ítem  $j$ . Así, cada individuo pertenecerá a solo una de estas clases. La presencia o ausencia de los atributos que no se requieren para responder correctamente este ítem, no afectan la pertenencia de un individuo a una u otra clase.

Retornando a nuestro ejemplo anterior, vemos que el ítem 3 solo requiere de la presencia de 3 atributos por lo que para este ítem se forman  $2^3 = 8$  grupos latentes.

A fines de simplificar la notación y sin perder generalidad, consideraremos siempre que son los primeros  $K_j^*$  atributos los requeridos para responder correctamente el ítem  $j$ , y que correspondientemente  $a_{lj}^*$  es el vector binario de dimensión  $K_j^*$ , que contiene 1's solo si un individuo cualquiera de la clase  $l$  posee tales componentes para responder correctamente el ítem  $j$ . Por ejemplo, el ítem 10 solo requiere dos atributos para ser respondido correctamente, por lo que el vector  $a_{lj}$  toma el valor de  $a_{lj}^* = (a_{lj1}, a_{lj2})^j$ , en lugar del vector completo  $(a_{lj1}, a_{lj2}, a_{lj3}, a_{lj4})^j$ .

En modelos más sencillos como el DINA, carecer de un atributo requerido para determinado ítem, es lo mismo que carecer de todos los atributos requeridos. Sin embargo, esto podría no ser siempre cierto, ya que un individuo que posee uno de los  $K_j^*$  atributos requeridos para el ítem  $j$ , podría tener mayor probabilidad de responder correctamente que aquel que no tiene ninguno. El modelo G-DINA relaja esta hipótesis de igual probabilidad de éxito.

Por esta razón es importante establecer una relación entre los vectores  $a_{lj}^*$  y  $a_{lj}^*$ . Así, para esta investigación, diremos que  $a_{lj}^* < a_{lj}^*$  si el perfil de atributos  $a_{lj}^*$  posee menos atributos de los requeridos para el ítem  $j$  que  $a_{lj}^*$ , es decir, si  $\sum_{k=1}^{K_j^*} a_{lj}^* < \sum_{k=1}^{K_j^*} a_{lj}^*$ .

En el ejemplo anterior, para el ítem 3, que requiere la presencia de tres atributos, podemos afirmar que el vector  $a_{33}^* = (0,0,1)$  es menor que el vector  $a_{13}^* = (1,1,0)$  ya que posee una "menor" cantidad de atributos requeridos.

En adelante denotaremos a la probabilidad de que los entrevistados con perfil de atributos  $a_{lj}^*$  respondan el ítem  $j$  correctamente por:  $P(X_j = 1 | a_{lj}^*) = P(a_{lj}^*)$ , siendo  $X_j$  la variable aleatoria que denota la respuesta de estos entrevistados al ítem  $j$ . Como es natural estas probabilidades deberán de satisfacer que  $P(a_{lj}^*) \leq P(a_{lj}^*)$  cuando  $a_{lj}^* < a_{lj}^*$ . El modelo

G-DINA plantea para estas probabilidades la ecuación:

$$P(\hat{a}_{ij}) = \delta_{j0} + \sum_{k=1}^{K^*} \delta_{jk} a_{ik} + \sum_{k=1}^{K^*-1} \sum_{k'=k+1}^{K^*} \delta_{jkk'} a_{ik} a_{ik'} + \dots + \delta_{j12\dots K^*} \prod_{k=1}^{K^*} a_{ik} \quad (2.1)$$

donde  $\delta_j = [\delta_{j0}, \delta_{j1}, \dots, \delta_{jK^*}, \delta_{j12}, \dots, \delta_{j12\dots K^*}]_j$  representa un vector de parámetros a estimarse para el ítem  $j$ , siendo  $\delta_{j0}$  un parámetro de intercepto (concretamente la probabilidad de que los encuestados respondan correctamente al ítem  $j$  cuando no posean ninguno de los atributos requeridos para responder satisfactoriamente este ítem),  $\delta_{jk}$  representa el efecto principal debido al atributo  $k$  (el incremento marginal en la probabilidad de responder correctamente al ítem  $j$  como resultado de la presencia del atributo  $k$ );  $\delta_{jkk'}$  representa el efecto de segundo orden, (el incremento marginal en la probabilidad de obtener una respuesta correcta debido a la presencia de los atributos  $k$  y  $k'$ ); y así sucesivamente hasta  $\delta_{j12\dots K^*}$  que representa el cambio en la probabilidad de obtener una respuesta correcta debido a la presencia de todos los atributos requeridos.

Uno de los aspectos que resaltan la importancia del modelo G-DINA, es que además de medir la contribución marginal de poseer un atributo particular en las probabilidades de responder correctamente un ítem, puede medir también (si existiera) el efecto conjunto de dos o más atributos en ella.

### 2.2. Ejemplo

Se presenta el siguiente ejemplo a fin de ilustrar mejor los conceptos estudiados. Se aplica un test de diagnóstico a 100 pacientes de un hospital psiquiátrico de Lima. Este test tiene 10 ítems que evalúan la presencia de 3 atributos. Sea entonces la matriz Q del test:

$$Q = \begin{matrix} & \begin{matrix} 1 & 0 & 0 \\ \square & 0 & 1 & 0 & \square \\ & 0 & 0 & 1 \\ \square & 1 & 0 & 1 & \square \\ 0 & 1 & 1 \\ \square & 1 & 1 & 0 & \square \\ & 1 & 0 & 1 \\ \square & 1 & 1 & 0 \\ \square & 0 & 1 & 1 & \square \\ & 1 & 1 & 1 \end{matrix} \end{matrix}$$

Vemos entonces que los tres primeros ítems requieren la presencia de solo un atributo cada uno; los siguientes 6 requieren 2 atributos cada uno, mientras que el décimo ítem requiere la presencia de los tres atributos. Así por ejemplo, la fila cuatro de la matriz nos indica que para marcar afirmativamente la pregunta 4 se requiere del primer y tercer atributo pero no del segundo.

Cada individuo  $i$  presenta un perfil  $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \alpha_{i3})$  que indica su estado con respecto a los atributos evaluados. Al igual que él, otros individuos compartirán este perfil y por tanto pertenecerán a una misma clase. Por ejemplo, si el perfil del entrevistado  $i$  fuera  $\alpha_i = (1, 0, 0)$  significa que este solo posee el primer atributo. Este perfil no es directamente observable, por lo que tiene que ser estimado a través del modelo G-DINA.

Cada ítem tiene diferente cantidad de parámetros estimados, que depende de la cantidad de atributos que necesita. Así por ejemplo, para el ítem 2, que solo requiere la presencia de un atributo (el segundo), supongamos que tenemos los siguientes resultados:

$$P(0) = \delta_0 = 0.1420$$

$$P(1) = \delta_0 + \delta_1 = 0.1420 + 0.6408 = 0.7828$$

donde  $P(0)$  representa la probabilidad de que los individuos que no presentan el atributo 2 respondan afirmativamente al ítem 2 mientras que  $P(1)$  es la probabilidad de responder afirmativamente al ítem 2 dado que el individuo presenta el atributo 2. Los valores  $\delta_0$  y  $\delta_1$  se interpretan como los efectos de la presencia de los atributos, y aquí quien posea el atributo 2, tendrá una probabilidad 0.6408 mayor de responder correctamente a este ítem que uno que no lo tenga.

Para el ítem 4, que de acuerdo a la matriz Q requiere de la presencia del primer y tercer atributo, supongamos que tenemos los siguientes parámetros:

$$P(00) = \delta_0 = 0.1160$$

$$P(10) = \delta_0 + \delta_1 = 0.1160 + 0.2946$$

$$P(01) = \delta_0 + \delta_2 = 0.1160 + 0.3565 = 0.4725$$

$$P(11) = \delta_0 + \delta_1 + \delta_2 + \delta_{12} = 0.1160 + 0.2946 + 0.3565 + 0.2440 = 0.8951$$

donde  $P(00)$  es la probabilidad de responder correctamente el ítem 4 dado que el individuo no posee ni el primer ni el tercer atributo que son los requeridos para responder correctamente este ítem. Los valores  $P(10)$  y  $P(01)$  representan la probabilidad de responder correctamente el ítem 4, dado que se posee el primer y el tercer atributo respectivamente; mientras que  $P(11)$  representa la probabilidad de responder correctamente el ítem 4, si se posee ambos atributos. Por su parte, el valor  $\delta_0$  representa el intercepto para el ítem 4; es decir, la probabilidad de responder afirmativamente sin poseer alguno de los atributos requeridos, mientras que  $\delta_1$  y  $\delta_2$  se interpretan como los efectos principales debido a la presencia del primer y tercer atributo respectivamente, es decir el cambio en la probabilidad de responder afirmativamente como resultado de poseer el atributo 1 y 3 respectivamente. Finalmente,  $\delta_{12}$  es un efecto de interacción de primer orden que representa un cambio en la probabilidad de responder afirmativamente debido a poseer los atributos 1 y 3.

Para el ítem 10, que es el único que requiere la presencia de los 3 atributos, son 8 los

parámetros a estimar. Si suponemos que estos son:

$$\begin{aligned}
 P(000) &= \delta_0 = 0.09 \\
 P(100) &= \delta_0 + \delta_1 = 0.09 + 0.14 = 0.23 \\
 P(010) &= \delta_0 + \delta_2 = 0.09 + 0.18 = 0.32 \\
 P(001) &= \delta_0 + \delta_3 = 0.09 + 0.34 = 0.43 \\
 P(110) &= \delta_0 + \delta_1 + \delta_2 + \delta_{12} = 0.09 + 0.14 + 0.18 + 0.01 = 0.42 \\
 P(101) &= \delta_0 + \delta_1 + \delta_3 + \delta_{13} = 0.09 + 0.14 + 0.34 + 0.02 = 0.59 \\
 P(011) &= \delta_0 + \delta_2 + \delta_3 + \delta_{23} = 0.09 + 0.18 + 0.34 + 0.02 = 0.63 \\
 P(111) &= \delta_0 + \delta_1 + \delta_2 + \delta_3 + \delta_{12} + \delta_{13} + \delta_{23} + \delta_{123} \\
 &= 0.09 + 0.14 + 0.18 + 0.34 + 0.01 + 0.02 + 0.02 + 0.04 = 0.84,
 \end{aligned}$$

$P(000)$  representa la probabilidad de responder afirmativamente el ítem dado que no posee ninguno de los tres atributos. Los valores  $P(100)$ ,  $P(010)$ ,  $P(001)$  representan las probabilidades de responder el ítem afirmativamente si se posee solamente el primer, segundo y tercer atributo, respectivamente. Por otro lado, los valores  $P(110)$ ,  $P(101)$  y  $P(011)$  son las probabilidades de responder afirmativamente dado que se posee dos de los tres atributos. Por último  $P(111)$  representa la probabilidad de responder afirmativamente si se posee los tres atributos. Los valores  $\delta_0$ ,  $\delta_1$ ,  $\delta_2$ ,  $\delta_3$ ,  $\delta_{12}$ ,  $\delta_{13}$ ,  $\delta_{23}$  y  $\delta_{123}$  son los efectos debido a la presencia de los atributos y sus interacciones.

### 2.3. Casos especiales

Se dice que el modelo G-DINA es saturado si el número de parámetros a ser estimados es igual al número de grupos latentes generados por el número de atributos requeridos por el ítem. Por ejemplo si para el ítem  $j$  se requieren dos atributos, implica la formación de cuatro grupos latentes. Por lo tanto, el modelo G-DINA para el ítem  $j$  con cuatro parámetros (intercepto, dos efectos principales y un efecto interacción) se considera saturado. Los modelos reducidos son obtenidos cuando se introducen restricciones a los parámetros. Por ejemplo, si decimos que el efecto interacción es cero, entonces estamos ante un modelo reducido.

Así, estableciendo las restricciones apropiadas, se pueden obtener diversos modelos reducidos a partir del modelo G-DINA. A continuación describiremos algunos de ellos: el modelo DINA (*Deterministic inputs, noisy and gate*), el modelo DINO (*Deterministic input, noisy or gate*) y los modelos aditivos A-CDM (*Additive cognitive diagnosis models*), LLM (*Linear logistic model*) y R-RUM (*Reduced reparametrized unified model*).

#### 2.3.1. Modelo DINA

El modelo DINA está especificado por:

$$P(a_{ij}^*) = \begin{cases} g_j & , \text{si } a_{ij}^* < \mathbf{1}_{\kappa_j^*} \\ 1 - s_j & , \text{en otro caso.} \end{cases} \quad (2.2)$$

donde  $\mathbf{1}_{\kappa_j^*}$  denota a un vector de unos de longitud  $\kappa_j^*$ ,  $g_j$  es la probabilidad de responder correctamente al ítem  $j$  dado que el individuo no posee todos los atributos requeridos, y

$1 - s_j$  es la probabilidad de que un individuo con todos los atributos requeridos para el ítem, no se equivoque y marque correctamente. En el modelo DINA, salvo el vector de perfil latente  $\alpha_j^* = 1_{K_j^*}$ , los  $2^{K_j^*} - 1$  grupos restantes, tienen la misma probabilidad de responder correctamente el ítem  $j$ . El modelo DINA se obtiene consecuentemente cuando en el modelo G-DINA fijamos todos los parámetros, con excepción de  $\delta_{j0}$  y  $\delta_{j12\dots K_j^*}$ , iguales a cero. Así, en términos de los parámetros del modelo G-DINA, se cumple que  $g_j = \delta_{j0}$  y  $1 - s_j = \delta_{j0} + \delta_{j12\dots K_j^*}$ .

### 2.3.2. Modelo DINO

En el modelo DINO (del acrónimo en inglés *Deterministic Input Noisy OR gate*) el ítem  $j$  divide a los individuos en aquellos que poseen al menos uno de los atributos requeridos para responder correctamente este ítem y aquellos que no tienen ninguno. Al igual que el modelo DINA, este solo presenta dos parámetros por ítem, independientemente del número de atributos requeridos.

El modelo DINO tiene la siguiente probabilidad de respuesta:

$$P(a_{ij}^*) = \begin{cases} g_j^* & , \text{ si } a_{ij}^* < \mathbf{0}_{K_j^*} \\ 1 - s_j & , \text{ en otro caso.} \end{cases} \quad (2.3)$$

donde  $\mathbf{0}_{K_j^*}$  denota a un vector de ceros de longitud  $K_j^*$ . A diferencia del modelo DINA, el parámetro  $g_j^*$  se refiere a la probabilidad de responder correctamente el ítem  $j$  cuando no se posee ninguno de los atributos requeridos; y  $1 - s_j$  es la probabilidad de responder correctamente para los individuos que poseen al menos uno de los atributos requeridos. Este modelo puede ser obtenido del modelo G-DINA, fijándose los parámetros

$$\delta_{jk} = \delta_{j k^j k^{j-1}} = (-1)^{K_j^* + 1} \delta_{j12\dots K_j^*}$$

para  $k = 1, \dots, K_j^*$ ,  $k^j = 1, \dots, K_j^* - 1$ , y  $k^{j-1} > k^j, \dots, K_j^*$ .

### 2.3.3. Modelos aditivos

Los modelos aditivos se obtienen del modelo G-DINA cuando se fijan todos los elementos de interacción en cero. Según la función de enlace que utilicen se obtienen los modelos A-CDM, LLM y R-RUM.

El Modelo de diagnóstico Cognitivo Aditivo (A-CDM por sus siglas en inglés) utiliza la función de enlace identidad y tiene la siguiente función de respuesta:

$$P(a_{ij}) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} a_{ik} \quad (2.4)$$

Al igual que el A-CDM, el Modelo Logístico Lineal (LLM por sus siglas en inglés) también se obtiene del modelo G-DINA cuando se fijan los parámetros de interacción en cero pero usando una función de enlace logit; es decir, su función de respuesta está dada por:

$$\text{logit}[P(a_{ij})] = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} a_{ik} \quad (2.5)$$

De igual manera, el Modelo Unificado Reparametrizado Reducido (R-RUM por sus siglas en inglés) usa la función de enlace log y posee la siguiente función de respuesta:

$$\log[P(\alpha_j)] = \delta_{j0} + \sum_{k=1}^{K^*} \delta_{jk} \alpha_k \quad (2.6)$$

Estos modelos plantean que poseer un atributo  $k$  incrementa de manera marginal la probabilidad de éxito del ítem  $j$  en  $\delta_{jk}$  y su aporte es independiente del de otros atributos. Cada uno de estos tres modelos poseen en su totalidad  $K_j^* + 1$  parámetros para cada ítem  $j$ .

### 2.4. Función de verosimilitud

Usando la fórmula (2.1), la probabilidad de obtener una respuesta  $x_{ij}$  s  $\{0, 1\}$  para el individuo  $i$ , que pertenece a la clase  $l$ , está dada por:

$$P(\mathbf{X}_{ij} = x_{ij} | \alpha_i = \alpha_l, \delta) = P(\alpha_j)^{x_{ij}} (1 - P(\alpha_j))^{(1-x_{ij})} = P(\alpha_j^*)^{x_{ij}} (1 - P(\alpha_j^*))^{(1-x_{ij})} \quad (2.7)$$

donde  $\delta_j = \delta_{j0}, \delta_{j1}, \dots, \delta_{jK^*}, \delta_{j12}, \dots, \delta_{j12\dots K^*}$ ,  $\mathbf{x}_{ij}$  representa la respuesta binaria de un individuo  $i$  al ítem  $j$  y  $P(\alpha_j)$  la probabilidad de obtener una respuesta correcta para el ítem  $j$  y puede ser escrita como  $P(\alpha_j^*)$  que es la forma reducida del vector de atributos. Entonces la verosimilitud del vector de respuestas  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{iJ}]$  de la persona  $i$  dado que pertenece al perfil o grupo latente  $l$  es:

$$L(\mathbf{x}_i | \alpha_l) = \prod_{j=1}^J P(\alpha_j^*)^{x_{ij}} (1 - P(\alpha_j^*))^{(1-x_{ij})} \quad (2.8)$$

Por la ley de la probabilidad total y la ecuación 2.8 concluimos entonces que la probabilidad de que el individuo  $i$  obtenga un patrón de respuestas  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{iJ}]$ , conocidos los parámetros de los ítems  $\delta$ , y las probabilidades de pertenencias a las clases latentes  $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_L) = (p(\alpha_1), p(\alpha_2), \dots, p(\alpha_L))$  está dada por:

$$P(\mathbf{X}_i = \mathbf{x}_i | \alpha_i, \boldsymbol{\pi}, \delta) = \sum_{l=1}^L L(\mathbf{x}_i | \alpha_l) p(\alpha_l) \quad (2.9)$$

donde  $\mathbf{x}_i$  representa el patrón de respuestas del individuo  $i$ .

La función de log-verosimilitud marginal de observar una muestra de respuestas de  $N$  individuos a los  $J$  ítems, viene dada finalmente por:

$$\ell(\mathbf{x}) = \log [L(\mathbf{x})] = \log \sum_{i=1}^N \sum_{l=1}^L L(\mathbf{x}_i | \alpha_l) p(\alpha_l) \quad (2.10)$$

donde  $\mathbf{x}$  representa los datos de respuestas de todos los individuos a los  $J$  ítems.

## Capítulo 3

### Estimación del modelo

Existen diversos métodos de estimación de los parámetros del modelo G-DINA. Uno de ellos es de máxima verosimilitud que permite la estimación conjunta de los parámetros de los ítems y vectores de habilidades. Sin embargo, la estimación conjunta de estas cantidades podría llevar a algunas inconsistencias debido a la gran cantidad de parámetros involucrados. Por tal razón, como lo indica [De la Torre \(2009\)](#), se recomienda usar el método de máxima verosimilitud marginal (MMLE) y el algoritmo de Esperanza-Maximización.

#### 3.1. Estimador de máxima verosimilitud marginal

La estimación de parámetros usando el método de máxima verosimilitud marginal, puede implementarse mediante el algoritmo EM (Esperanza-Maximización).

Este método, estima primero los parámetros  $P(a_j^*)$ , que representan las probabilidades de responder correctamente el ítem  $j$  para la clase  $l$ , para luego calcular sus componentes  $\delta_j = \delta_{j0}, \delta_{j1}, \dots, \delta_{jK^*}, \delta_{j12}, \dots, \delta_{j12\dots K^*}$  mediante mínimos cuadrados.

Sea entonces  $l(\mathbf{x})$  la función de log verosimilitud marginal de los datos de respuesta, dada en (2.10), es decir,

$$l(\mathbf{x}) = \log [L(\mathbf{x})] = \log \prod_{i=1}^N \sum_{l=1}^L L(x_i|a_l) p(a_l) \quad (3.1)$$

donde recordemos que  $L(x_i|a_l)$  denota a la verosimilitud del vector de respuestas de la persona  $i$  dado que este posee un vector de atributos  $a_l$ . Entonces, para estimar el modelo G-DINA, debemos maximizar  $l(\mathbf{x})$  con respecto a  $P(a_j^*)$ . Antes de empezar la iteración del algoritmo EM, se debe escoger valores iniciales para  $\pi = (p(a_1), p(a_2), \dots, p(a_L))$  y el vector de parámetros de los ítems  $\delta_j = \delta_{j0}, \delta_{j1}, \dots, \delta_{jK^*}, \delta_{j12}, \dots, \delta_{j12\dots K^*}$ .

#### Paso Esperanza

Obtenemos la probabilidad a posteriori de que el individuo  $i$  posea un perfil latente  $l$  para el ítem  $j$  dada su respuesta, mediante el teorema de Bayes:

$$p(a_j^* | x_i) = \frac{p(a_j^*) L(x_i | a_j^*)}{\sum_{l=1}^L p(a_l) L(x_i | a_l)} \quad (3.2)$$

Se calculan dos recuentos provenientes de la probabilidad a posteriori:

- El número esperado de individuos que son clasificados para el ítem  $j$  como pertenecientes al grupo latente  $l$

$$I_{a_{lj}^*} = \sum_{i=1}^N p(a_{lj}^* | x_i) \tag{3.3}$$

- El número esperado de individuos clasificados para el ítem  $j$  como pertenecientes al grupo latente  $l$  que se espera respondan correctamente el ítem  $j$ .

$$R_{a_{lj}^*} = \sum_{i=1}^N p(a_{lj}^* | x_i) x_{ij} \tag{3.4}$$

**Paso Maximización**

Este está conformado por 2 etapas:

a) Se actualiza el vector de parámetros  $\delta_j = [\delta_{j0}, \delta_{j1}, \dots, \delta_{jK^*}, \delta_{j12}, \dots, \delta_{j12\dots K^*}]$ . Las ecuaciones para las estimaciones son obtenidas a partir de la primera derivada de la log verosimilitud marginal de los datos de respuesta con respecto a los parámetros de los ítems e igualando a cero. Esta derivada, involucra los dos conteos descritos en (3.3) y (3.4) y es obtenida de la siguiente manera:

$$\frac{\partial l(\mathbf{x})}{\partial P(a_{lj}^*)} = \sum_{i=1}^N \frac{1}{L(\mathbf{x})} \frac{\partial L(x_i)}{\partial P(a_{lj}^*)} = \sum_{i=1}^N \frac{1}{L(\mathbf{x})} \sum_{l=1}^L \frac{\partial L(x_i | a_{lj}^*)}{\partial P(a_{lj}^*)} \tag{3.5}$$

Resolviendo la derivada:

$$\frac{\partial L(x_i | a_{lj}^*)}{\partial P(a_{lj}^*)} = \frac{\partial \sum_{j=1}^{Q_j} P_j(a_l)^{x_{ij}} [1 - P_j(a_l)]^{1-x_{ij}}}{\partial P(a_{lj}^*)} \tag{3.6}$$

Cuando  $j = l$  entonces  $j$  es una constante al derivar con respecto a  $j$

$$\frac{\partial L(x_i | a^*)}{\partial P(a_{lj}^*)} = \frac{P_j(a_l)^{x_{ij}} [1 - P_j(a_l)]^{1-x_{ij}}}{P_j(a_l)^{x_{ij}} [1 - P_j(a_l)]^{1-x_{ij}}} \frac{\partial \sum P_j(a_l)^{x_{ij}} [1 - P_j(a_l)]^{1-x_{ij}}}{\partial P(a_{lj}^*)} \tag{3.7}$$

La derivada de la parte derecha de la ecuación (3.7) es igual a:

$$\begin{aligned} &= [1 - P(a)]^{1-x_{ij}} x_{ij} P(a)^{x_{ij}-1} \frac{\partial P_j(a_l)}{\partial P(a_{lj}^*)} + P(a)^{x_{ij}} (1 - x_{ij}) [1 - P(a)]^{1-x_{ij}-1} \frac{-\partial P_j(a_l)}{\partial P(a_{lj}^*)} \\ &= P_j(a_l)^{x_{ij}} [1 - P_j(a_l)]^{1-x_{ij}} \frac{\partial P_j(a_l)}{\partial P(a_{lj}^*)} \frac{x_{ij}}{P_j(a_l)} - \frac{1 - x_{ij}}{1 - P_j(a_l)} \sum_{l=1}^L \frac{\partial P_j(a_l)}{\partial P(a_{lj}^*)} \\ &= P_j(a_l)^{x_{ij}} [1 - P_j(a_l)]^{1-x_{ij}} \frac{\partial P_j(a_l)}{\partial P(a_{lj}^*)} \sum_{l=1}^L \frac{x_{ij}(1 - P_j(a_l)) - P_j(a_l)(1 - x_{ij})}{P_j(a_l)(1 - P_j(a_l))} \\ &= P_j(a_l)^{x_{ij}} [1 - P_j(a_l)]^{1-x_{ij}} \frac{\partial P_j(a_l)}{\partial P(a_{lj}^*)} \frac{x_{ij} - P_j(a_l)}{P_j(a_l)(1 - P_j(a_l))} \end{aligned} \tag{3.8}$$

Entonces, reemplazando (3.8) en la derivada de la parte derecha de (3.7), obtenemos:

$$\begin{aligned} \frac{\partial L(\mathbf{x}_i|a_{ij})}{\partial P(a^{*ij})} &= \frac{\partial P(a)}{\partial P(a^{*ij})} \frac{x_{ij} - P(a)}{P(a)(1 - P(a))} \\ &= L(x_i|a^{*ij}) \frac{\partial P(a)}{\partial P(a^{*ij})} \frac{x_{ij} - P(a)}{P(a)(1 - P(a))} \end{aligned} \quad (3.9)$$

Reemplazamos los resultados de la ecuación (3.9) en la ecuación (3.5)

$$\frac{\partial l(\mathbf{x})}{\partial P(a_{ij})} = \sum_{i=1}^L L(\mathbf{x}_i) \sum_{l=1}^L p(a_{ij}^*) L(\mathbf{x}_i|a_{ij}^*) \frac{\partial P(a_l)}{\partial P(a_{ij}^*)} \frac{x_{ij} - P(a_l)}{P(a_l)(1 - P(a_l))}$$

Intercambiamos el orden de las sumatorias:

$$\begin{aligned} \frac{\partial l(\mathbf{x})}{\partial P(a_{ij}^*)} &= \sum_{l=1}^L \frac{\partial P(a_l)}{\partial P(a_{ij}^*)} \frac{1}{P(a_l)(1 - P(a_l))} \sum_{i=1}^L p(a_{ij}^*) L(\mathbf{x}_i|a_{ij}^*) [x_{ij} - P(a_l)] \\ &= \sum_{l=1}^L \frac{\partial P(a_l)}{\partial P(a_{ij}^*)} \frac{1}{P(a_l)(1 - P(a_l))} \sum_{i=1}^L p(a_{ij}^* | \mathbf{x}_i) [x_{ij} - P(a_l)] \\ &= \sum_{l=1}^L \frac{\partial P(a_l)}{\partial P(a_{ij}^*)} \frac{1}{P(a_l)(1 - P(a_l))} \sum_{i=1}^L p(a_{ij}^* | \mathbf{x}_i) x_{ij} - P(a_l) \sum_{i=1}^L p(a_{ij}^* | \mathbf{x}_i) \\ &= \sum_{l=1}^L \frac{\partial P(a_l)}{\partial P(a_{ij}^*)} \frac{1}{P(a_l)(1 - P(a_l))} \sum_{i=1}^L R_{a_{ij}^*} - P(a_l) I_{a_{ij}^*} \end{aligned} \quad (3.10)$$

Para el ítem  $j$ , la expresión (3.10) puede ser escrita como:

$$\begin{aligned} \frac{\partial l(\mathbf{x})}{\partial P(a_{ij}^*)} &= \sum_{l=1}^L \frac{\partial P(a_l)}{\partial P(a_{ij}^*)} \frac{1}{P(a_l)(1 - P(a_l))} \sum_{i=1}^L R_{a_{ij}^*} - P(a_l) I_{a_{ij}^*} \\ &= \frac{\partial P(a_{ij}^*)}{\partial P(a_{ij}^*)} \frac{1}{P(a_{ij}^*)(1 - P(a_{ij}^*))} \sum_{i=1}^L R_{a_{ij}^*} - P(a_{ij}^*) I_{a_{ij}^*} \\ &= \frac{\partial P(a_{ij}^*)}{\partial P(a_{ij}^*)} \frac{1}{P(a_{ij}^*)(1 - P(a_{ij}^*))} \sum_{i=1}^L R_{a_{ij}^*} - P(a_{ij}^*) I_{a_{ij}^*} \end{aligned}$$

Así, maximizar  $l(\mathbf{x})$  con respecto a  $P(a_{ij}^*)$  se reduce a resolver la ecuación para  $P(a_{ij}^*)$

$$\frac{1}{P(a_{ij}^*)(1 - P(a_{ij}^*))} \sum_{i=1}^L R_{a_{ij}^*} - P(a_{ij}^*) I_{a_{ij}^*} = 0 \quad (3.11)$$

Por lo tanto, el estimador de máxima verosimilitud marginal de  $\hat{P}(a_{ij}^*)$  está dada por:

$$\hat{P}(a_{ij}^*) = \frac{R_{a_{ij}^*}}{I_{a_{ij}^*}} \quad (3.12)$$

donde  $p(a_{ij}^* | \mathbf{x}_i)$  representa la probabilidad a posteriori de que el individuo  $i$  esté en el grupo  $a_{ij}^*$ ,  $I_{a_{ij}^*} = \sum_{i=1}^L p(a_{ij}^* | \mathbf{x}_i)$  es el número de individuos que se espera pertenezcan al grupo

latente  $a_{ij}^*$  y  $R_{a_{ij}^*} = \sum_{i=1}^L p(a_{ij}^* | \mathbf{x}_i) x_{ij}$  es el número de individuos en el grupo latente  $a_{ij}^*$  que

se espera respondan correctamente el ítem  $j$ .

b) Se actualiza la distribución de clases  $P(a_l)$  y se procede a calcular el número esperado de individuos en la clase de atributos  $a_l$ :

$$n_l = \sum_{i=1}^N P(a_l^* | x_i; \delta)$$

actualizando la distribución de las clases de atributos,  $P(a_l)$ , mediante:

$$P(a_l) = \frac{n_l}{N}, \quad l = 1, \dots, L$$

y las probabilidades de que el individuo presente el atributo  $k$  se definen como:

$$P(a_k) = \sum_{l=1}^L a_{lk} \cdot P(a_l | \gamma), \quad k = 1, \dots, K$$

En resumen, el primer paso del algoritmo empieza con valores iniciales para  $P(a_l^*)$ . El segundo paso consiste en calcular los valores para  $R_{a_l^*}$  y  $I_{a_l^*}$  basados en los presentes valores de  $P(a_l^*)$ . El tercer paso encuentra los valores de  $P(a_l^*)$  usando (3.11) y por último, se repiten los pasos 2 y 3 hasta conseguir la convergencia.

c) A partir de  $P(a_l | x_i)$  se puede calcular los perfiles de atributos de los individuos mediante tres métodos:

- El enfoque de clasificación máxima a priori (MAP) que considera como perfil de atributos del individuo  $i$  a la clase de atributos que presenta el mayor valor de  $P(a_l | x_i)$ :

$$\hat{\alpha}_{iMAP} = \arg_{a_l} \max \{ P(a_l | x_i) \}, \quad i = 1, \dots, N$$

- El que clasifica al individuo  $i$  en base a la estimación por máxima verosimilitud (MLE), que se obtiene maximizando:

$$\hat{\alpha}_{iMLE} = \arg_{a_l} \max \{ P(x_i | a_l) \}, \quad i = 1, \dots, N$$

- El enfoque de clasificación basado en probabilidades a posteriori (EAP), donde el perfil estimado de cada individuo es de la forma:  $\hat{\alpha}_i = \{ P(a_k = 1 | x_i) \}$ , donde el elemento  $k$  del vector  $\hat{\alpha}_i$  es la probabilidad a posteriori de que el atributo  $k$  esté presente en el individuo  $i$ . Estas probabilidades pueden ser calculadas:

$$P(a_k = 1 | x_i) = \sum_{\forall l: a_{lk}=1} P(a_l | x_i) \quad (3.13)$$

donde  $P(a_l | x_i)$  es la probabilidad a posteriori de que el individuo  $i$  tenga el perfil de atributos  $a_l$  y la sumatoria se realiza sobre todos los perfiles donde el  $k$ -ésimo elemento de  $a_l$  es 1. Así, las probabilidades que estén cerca a 1 indican una fuerte evidencia de la presencia del atributo. Para una mejor interpretación estas probabilidades pueden ser convertidas a cero

o uno basadas en ciertas reglas. Se puede establecer por ejemplo que si la probabilidad es mayor a 0.5 entonces el atributo está presente.

Para calcular los errores estándar necesitamos usar la matriz de información la cual se obtiene a través de la segunda derivada de la log-verosimilitud marginalizada con respecto a los parámetros  $P(a_{lj}^*)$  y  $P(a_{lj}^*)$

Así tenemos que:

$$\frac{\partial^2 \ln(x)}{\partial P(a_{lj}^*) \partial P(a_{lj}^*)} = - \sum_{i=1}^n \frac{p(a_{ij}^* | x_i) \frac{x_i - P(a_{lj}^*)}{P(a_{lj}^*) [1 - P(a_{lj}^*)]}}{p(a_{ij}^* | x_i) \frac{x_i - P(a_{lj}^*)}{P(a_{lj}^*) [1 - P(a_{lj}^*)]}} \quad (3.14)$$

Usando la expresión (3.12) y los valores observados  $X$  obtenemos la matriz de información para los parámetros del ítem  $j$ ,  $I(P_j^*)$ , donde  $P_j^* = P(a_{lj}^*)$  y el error estándar de  $P(a_{lj}^*)$  se calcula como la raíz cuadrada de la  $l$ -ésima diagonal de  $I^{-1}(P_j^*)$ .

### 3.2. Matriz de Diseño

Para poder convertir los parámetros estimados del modelo G-DINA, en el vector  $\delta_j = \delta_{j0}, \delta_{j1}, \dots, \delta_{jK^*}, \delta_{j12}, \dots, \delta_{j12\dots K^*}$ , se hace uso de la matriz de diseño  $M_j$ . Esta matriz es de dimensión  $2^{K^*} \times P$ , donde  $P$  es el número de parámetros del modelo de interés. Así, para un modelo saturado,  $P = 2^{K^*}$ , es decir el número de columnas es igual al número de filas.

Entonces, para definir la matriz  $M_j$ , señalaremos primero a  $A_j = \{a_{lk}\}$  una matriz  $2^{K^*} \times K^*$  de las posibles combinaciones de los atributos requeridos para el ítem  $j$ , y que se define de la siguiente manera:

$$A_j = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \square & 0 & & \square \\ \square & 1 & 0 & \dots & 0 & \square \\ \square & 0 & & & & \square \\ \square & & & \ddots & & \square \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

Para los modelos saturados, la fila  $l$  de la matriz de diseño saturada  $M_j^{(s)}$  se genera a partir de la fila  $l$  de la matriz  $A_j$ . Su primer elemento es "1", seguido por el vector  $a_{lk}$ , para  $k = 1, \dots, K^*$ ; luego por  $a_{lk} \times a_{jk}$ , para  $k = 1, \dots, K^*-1$  y  $kl = k+1, \dots, K^*$  y así sucesivamente. El último elemento de este vector es  $a^k$

Definimos ahora  $P_j = P(a_{lj}^*)$ , un vector de longitud  $2^{K^*}$ , donde  $P(a_{lj}^*)$  es la probabilidad de responder correctamente el ítem  $j$  de la clase  $l$ . Para obtener los valores estimados de  $\delta_j = \delta_{j0}, \delta_{j1}, \dots, \delta_{jK^*}, \delta_{j12}, \dots, \delta_{j12\dots K^*}$  dado  $P_j$  hacemos uso del estimador de mínimos cuadrados. Esto está motivado por la estructura del modelo de regresión lineal múltiple de  $P(a_{lj}^*)$ . Este estimador viene por tanto dado por:

$$\hat{\delta}_j = (M_j^S M_j^S)^{-1} M_j^S P_j \quad (3.15)$$

Tomemos de ejemplo un caso en el que  $K^* = 3$ , entonces nuestra matriz  $A_j$  está definida por

$$A_j = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Entonces nuestra matriz de diseño está dada por:

$$M^{(S)}_j [8 \times 8] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Los parámetros de los modelos reducidos pueden ser estimados de  $\hat{\mathcal{P}}_j^*$ , especificando las matrices de diseño apropiadas

### 3.3. Comparación de modelos

Cuando no se afecta evidentemente el ajuste de un modelo al introducirse restricciones, se prefiere un modelo reducido al saturado por muchas razones. En primer lugar, los modelos reducidos tienen interpretaciones más directas y requieren menor tamaño de muestra para su estimación. En segundo lugar, por el principio de parsimonia que indica que los modelos más sencillos deben ser preferidos ante los más complejos cuando no hay mucha diferencia entre ellos. Y por último, los modelos reducidos pueden brindar mejores tasas de clasificación que los modelos saturadas, particularmente cuando el tamaño de muestra es pequeño (Wenchao et al., 2016). A continuación, se describirá el test de Wald y el índice de disimilitud para escoger el modelo adecuado para cada ítem.

#### 3.3.1. Test de Wald

De la Torre (2011) propone el uso del test de Wald como un método de determinar si alguno de los modelos reducidos puede ser usado en lugar del modelo saturado G-DINA sin afectar los resultados del ajuste. Este test es un procedimiento a nivel de ítem, que es desarrollado un ítem a la vez siempre que  $K_j^* > 1$ . Para desarrollar el test es necesario establecer una matriz de restricciones  $R$  de dimensión  $(2^{K_j^*} - p) \times 2^{K_j^*}$ , donde  $p$  es el número de parámetros que se estiman para el modelo reducido. Por ejemplo, cuando  $K_j^* = 3$ , el modelo DINA y el modelo A-CDM presentan un valor de  $p$  igual a 2 y 4 parámetros respectivamente.

Por consiguiente, las matrices de restricciones  $R$ , de dimensiones  $6 \times 8$  y  $4 \times 8$ , para estos modelos reducidos, son:

$$\begin{aligned}
 R^{(1)}_{[6 \times 8]} &= \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix} \\
 R^{(2)}_{[4 \times 8]} &= \begin{bmatrix} 1 & -1 & -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 \end{bmatrix}
 \end{aligned}$$

Por lo tanto, las matrices de restricciones  $R^{(1)}$  para el modelo DINA, y  $R^{(2)}$  para el A-CDM, presentan las siguientes restricciones:

$$\begin{aligned}
 \delta_1 = \delta_2 = \delta_3 = \delta_{12} = \delta_{13} = \delta_{23} = 0, \\
 \delta_{12} = \delta_{13} = \delta_{23} = \delta_{123} = 0
 \end{aligned}$$

respectivamente. Notamos por ejemplo, que para el A-CDM los efectos de interacción están fijados en cero.

Luego de definir la matriz de restricciones, el estadístico de Wald,  $W$ , es calculado de la siguiente manera:

$$W = [R \times P_j] \left( R \times Var(P_j) \times R^j \right)^{-1} [R \times P_j] \tag{3.16}$$

donde  $P_j = P(\alpha_{ij})$  y  $Var(P_j)$  es la inversa de la matriz de información. Bajo la hipótesis nula  $R \times P_j = 0$ ,  $W$  presenta una distribución chi-cuadrado con  $2K^j - p$  grados de libertad.

### 3.3.2. Disimilitud entre los modelos reducidos

El índice de disimilitud ha sido desarrollado recientemente por (Wenchao et al., 2016) y es definido como una función de las diferencias de las probabilidades de éxito en todos los posibles grupos latentes entre dos modelos. Así, las disimilitudes entre dos MDCs,  $M_1$  y  $M_2$  para el ítem  $j$ , presentan la siguiente forma:

$$DS(M_1, M_2) = \min_{\varphi} \sum_l | P_{M_1}(\alpha^*) - \tilde{P}_{M_2}(\alpha^*) |_{ij} \tag{3.17}$$

donde  $P_{M_1}(\alpha_{ij}^*)$  y  $\tilde{P}_{M_2}(\alpha_{ij}^*)$  son las probabilidades de éxito para  $\alpha_{ij}^*$  basados en los modelos  $M_1$  y  $M_2$  respectivamente y  $\varphi$  es el vector de parámetros de  $M_2$ . Los valores de  $P_{M_1}(\alpha_{ij}^*)$  son fijos, mientras que los valores de  $\tilde{P}_{M_2}(\alpha_{ij}^*)$  son estimados buscando minimizar la suma del valor absoluto de las diferencias en cada elemento de los vectores  $P_{M_1}(\alpha_{ij}^*)$  y  $\tilde{P}_{M_2}(\alpha_{ij}^*)$ . Cuando  $DS(M_1, M_2) = 0$ , decimos que  $M_2$  puede aproximar a  $M_1$  perfectamente. Así, valores pequeños para  $DS(M_1, M_2)$  indican que  $M_2$  puede aproximar bien a  $M_1$ , mientras que valores grandes indican que no. Este indicador no es simétrico, es decir que  $DS(M_1, M_2)$  no es necesariamente igual que  $DS(M_2, M_1)$ .

## Capítulo 4

### Estudio de Simulación

Este capítulo presenta los estudios de simulación desarrollados, que han sido adaptados del contexto del presente trabajo; es decir del diagnóstico de los desórdenes mentales. Se presentarán los escenarios evaluados y los resultados obtenidos.

Basándonos en el contexto del test de evaluación que se usará en la aplicación, el Cuestionario de Auto-reporte o *Self Reporting* (SRQ 18), se usarán los siguientes parámetros fijos:

- Número de atributos: 4.
- Número de ítems: Se usará la misma cantidad de ítems del test SRQ-18: 22.
- Número de réplicas: Estudios como el de ([Harwellet al., 1996](#)), recomiendan un mínimo de 25 réplicas para obtener resultados confiables en estudios de simulación para la TRI. Además este número debe aumentar a medida que la complejidad de los efectos se incrementa. Por su parte ([Wenchaoet al., 2016](#)), en su ejercicio de simulación para estudiar la estimación de los parámetros usando distintos modelos de diagnóstico cognitivo, usan un número de réplicas de 100 casos por cada condición estudiada. Por lo tanto, en este estudio se usará como número de réplicas 150.
- Matriz Q: matriz de orden  $22 \times 4$  del Cuestionario de Auto-reporte o Self Reporting (SRQ 18) que se usará en la aplicación.

Se simularán matrices de respuestas dicotómicas para evaluar la precisión del recupero de los parámetros bajo diferentes tamaños de muestra. Así, se examinará el desempeño del modelo bajo distintos tamaños de muestra en el contexto de la aplicación. Por tal razón, los tamaños evaluados deberán ser relativamente pequeños, ya que en un contexto real no es muy común trabajar con gran cantidad de observaciones. Los valores evaluados son de 100, 200, 300, 400, 500 y 1000.

Para los estudios de simulación de modelos de diagnóstico cognitivo es recomendable trabajar con datos para los cuáles conocemos los verdaderos modelos y parámetros para cada ítem ([George, 2014](#)). Así, para la generación de los datos se usarán los parámetros obtenidos para todos los ítems al aplicar los modelos a nuestra muestra de estudio.

El objetivo de este estudio será:

- Evaluar la precisión con la cual se recuperan los parámetros bajo diferentes tamaños de muestra mediante el algoritmo EM.

#### 4.1. Simulación de los datos y estimación de parámetros

La generación de los datos se realiza a través de la función simGDINA del paquete GDINA en R. Así, estableciendo los valores mencionados de los parámetros se generan 50 bases de datos (número de réplicas) para cada uno de los 6 escenarios (los 6 tamaños de muestra evaluados).

Con los datos generados, se procede a ajustar el modelo seleccionado para cada ítem (aquellos determinados en la aplicación) a través de la función GDINA del paquete estadístico del mismo nombre. Estos resultados pueden ser comparados con los parámetros verdaderos, es decir aquellos con los que se empezó la simulación (que fueron obtenidos de los datos reales). Para efectuar esta comparación, seguiremos los criterios de evaluación definidos en la siguiente sección.

Con el fin de evaluar la estimación de los parámetros obtenidos, se usará la raíz del error cuadrático medio (ECM) y el error estándar.

El sesgo está definido como la sumatoria de las diferencias entre el valor estimado y el verdadero entre el número de réplicas.

$$Sesgo = \sum_{r=1}^R \hat{\pi}_r - \pi \quad (4.1)$$

donde  $\pi$  representa el valor verdadero de los parámetros mientras que  $\hat{\pi}$  es el valor estimado de estos parámetros. El valor de R es el número de réplicas a usarse, que en este caso es 150.

El Error Cuadrático Medio (RMSE por su siglas en inglés) es definido como la raíz cuadrada de la sumatoria de las diferencias al cuadrado entre los parámetros estimados y los verdaderos entre el número de réplicas.

$$RMSE = \sqrt{\frac{\sum_{r=1}^R (\hat{\pi}_r - \pi)^2}{R}} \quad (4.2)$$

donde  $\pi$  representa el valor verdadero de los parámetros mientras que  $\hat{\pi}$  es el valor estimado de estos parámetros. El valor de R es el número de réplicas a usarse, que en este caso es 150. Este indicador es usado para evaluar la varianza total del error de estimación entre los valores observados y los reales. Así, resultados con valor cero indican que no hay discrepancia entre el modelo y los valores verdaderos. Por otro lado, algunos autores señalan que valores mayores a 0.1 indican un pobre ajuste de los datos (George, 2014).

#### 4.2. Resultados

Los resultados obtenidos serán mostrados en tres secciones, donde se agruparán a los ítems de acuerdo al número de atributos evaluados. La presentación se realiza de esa manera ya que la cantidad de parámetros estimados y su interpretación depende del número de atributos evaluados y el modelo aplicado.

##### 4.2.1. Ítems que miden un atributo

En el test SRQ -18 existen 12 ítems que miden solo un atributo y a los que se les aplicó el modelo DINA, que es un caso especial del modelo G-DINA y que fue estudiado en los capítulos

previos. Por esta razón, para estos ítems solo se estiman dos parámetros: el parámetro de intercepto  $\delta_0$  y el parámetro del efecto principal  $\delta_1$ . El ítem 18, que mide dos atributos, se considerará dentro de este grupo debido a que al ser ajustado con el modelo DINO, solo requiere la estimación de dos parámetros independientemente del número de atributos que evalúe.

Por lo tanto, ambos parámetros tienen interpretación distinta dependiendo del modelo aplicado. Así, el parámetro de intercepto para el ítem 18, para el cual se usó el modelo DINO, se interpreta como la probabilidad de responder correctamente el ítem dado que el individuo no posee ninguno de los atributos requeridos para ese ítem. Para el resto de ítems, a los que se les aplicó el modelo DINA, se interpreta como la probabilidad de responder correctamente el ítem dado que el individuo carece de al menos uno de los atributos requeridos.

Por su parte, para el ítem 18, el parámetro del efecto principal se interpreta como el efecto de poseer alguno de los atributos requeridos en la probabilidad de responder afirmativamente este ítem. Para el resto de ítems, este valor se interpreta como el efecto de poseer el atributo requerido, en la probabilidad de responder afirmativamente el ítem.



Cuadro 4.1: Comparación de parámetros de intercepto de ítems que requieren un atributo

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
4	RMSE	0.0785	0.0530	0.0405	0.0347	0.0321	0.0215
	Error	-0.0122	0.0006	-0.0046	-0.0040	-0.0045	-0.0024
6	RMSE	0.0963	0.0596	0.0430	0.0457	0.0351	0.0295
	Error	-0.0075	-0.0022	-0.0080	-0.0076	-0.0031	-0.0032
9	RMSE	0.0628	0.0521	0.0342	0.0350	0.0319	0.0203
	Error	-0.0008	0.0002	-0.0033	-0.0029	-0.0021	-0.0022
11	RMSE	0.0431	0.0332	0.0290	0.0228	0.0230	0.0169
	Error	-0.0002	0.0010	-0.0014	0.0001	-0.0037	-0.0038
14	RMSE	0.0556	0.0393	0.0314	0.0285	0.0240	0.0171
	Error	-0.0041	0.0020	0.0018	0.0042	0.0016	-0.0010
15	RMSE	0.0435	0.0297	0.0252	0.0212	0.0185	0.0123
	Error	-0.0025	-0.0016	-0.0007	-0.0004	-0.0005	0.0008
16	RMSE	0.0715	0.0518	0.0442	0.0314	0.0309	0.0214
	Error	-0.0094	0.0036	0.0000	0.0048	-0.0004	-0.0001
17	RMSE	0.0545	0.0330	0.0295	0.0280	0.0214	0.0167
	Error	-0.0017	-0.0003	-0.0002	0.0022	0.0020	-0.0016
18	RMSE	0.1169	0.0765	0.0625	0.0491	0.0423	0.0315
	Error	0.0020	-0.0037	-0.0075	-0.0107	-0.0044	0.0019
19	RMSE	0.1405	0.1265	0.1144	0.0944	0.0992	0.0549
	Error	0.0115	-0.0161	-0.0080	-0.0110	-0.0192	0.0005
20	RMSE	0.1141	0.0660	0.0574	0.0394	0.0384	0.0263
	Error	-0.0210	-0.0017	-0.0069	0.0024	0.0061	-0.0056
21	RMSE	0.1316	0.1080	0.0980	0.0770	0.0770	0.0472
	Error	-0.0008	-0.0131	-0.0164	-0.0038	-0.0055	-0.0002
22	RMSE	0.0667	0.0527	0.0398	0.0374	0.0339	0.0228
	Error	0.0076	0.0033	0.0014	-0.0031	-0.0028	0.0022

Cuadro 4.2: Comparación de parámetros de efecto principal de ítems que requieren un atributo

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
4	RMSE	0.1184	0.0829	0.0631	0.0506	0.0521	0.0356
	Sesgo	0.0145	0.0040	0.0100	0.0080	0.0091	0.0097
6	RMSE	0.1045	0.0677	0.0487	0.0496	0.0376	0.0291
	Sesgo	0.0060	0.0063	0.0101	0.0111	0.0043	0.0035
9	RMSE	0.0767	0.0650	0.0433	0.0403	0.0383	0.0249
	Sesgo	0.0029	-0.0013	0.0030	0.0010	0.0010	0.0031
11	RMSE	0.0645	0.0525	0.0455	0.0384	0.0334	0.0233
	Sesgo	-0.0045	-0.0030	-0.0027	0.0004	0.0021	0.0062
14	RMSE	0.0893	0.0688	0.0541	0.0431	0.0426	0.0295
	Sesgo	0.0098	-0.0048	-0.0017	-0.0091	-0.0013	0.0011
15	RMSE	0.0732	0.0531	0.0424	0.0404	0.0349	0.0234
	Sesgo	-0.0042	0.0001	-0.0009	0.0026	0.0010	-0.0006
16	RMSE	0.0868	0.0618	0.0523	0.0362	0.0360	0.0252
	Sesgo	0.0077	-0.0017	-0.0004	-0.0051	-0.0008	-0.0003
17	RMSE	0.0856	0.0561	0.0509	0.0488	0.0386	0.0271
	Sesgo	0.0007	-0.0007	-0.0020	0.0019	0.0018	0.0013
18	RMSE	0.1358	0.0905	0.0727	0.0576	0.0522	0.0384
	Sesgo	-0.0033	0.0054	0.0067	0.0133	0.0033	-0.0001
19	RMSE	0.2037	0.1666	0.1468	0.1014	0.1102	0.0656
	Sesgo	0.0080	0.0313	0.0030	0.0206	0.0343	0.0058
20	RMSE	0.1900	0.1063	0.0812	0.0701	0.0625	0.0421
	Sesgo	0.0439	0.0005	0.0094	-0.0021	-0.0068	0.0086
21	RMSE	0.2210	0.1537	0.1469	0.0951	0.0941	0.0635
	Sesgo	0.0423	0.0384	0.0349	0.0225	0.0093	0.0103
22	RMSE	0.1796	0.1120	0.0721	0.0611	0.0521	0.0368
	Sesgo	0.0094	0.0060	0.0002	0.0093	0.0066	0.0022

Los resultados para las estimaciones de ambos parámetros se muestran en los Cuadros 4.1 y 4.2. Notamos que los resultados para el parámetro de intercepto son mejores que los del efecto principal, al presentar menor valor de RMSE en todos los ítems y escenarios evaluados. En ambos parámetros, se obtienen resultados con buen desempeño para 10 de los 13 ítems a partir de los 200 casos (RMSE < 0.1). El desempeño mejora a medida que aumenta el tamaño de muestra usado, tomando el mejor valor cuando este valor es de 1000. Por último, notamos que no es recomendable usar un tamaño de muestra menor a los 100 casos, ya que en ese caso, solo se presentan resultados mínimamente confiables para 6 de los 13 ítems evaluados, de acuerdo a los valores del RMSE.

### 4.2.2. Ítems que miden dos atributos

De acuerdo a la matriz Q, el test usado presenta 9 ítems que evalúan dos atributos, de los cuáles el ítem 18 ya fue presentado en el rubro anterior.

Del grupo restante, cuatro ítems fueron ajustados usando modelos G-DINA con función de enlace identidad, mientras que para el otro grupo se usaron los modelos con función de enlace log y logit (RRUM y LLM respectivamente). El Cuadro 4.3 muestra los resultados para el primer grupo de ítems, donde se comparan los estimados de los parámetros componentes delta  $\delta = \{\delta_0, \delta_1, \delta_2, \delta_{12}\}$ , es decir el intercepto y los efectos debido a la presencia de los atributos. Notamos entonces que para el intercepto se obtienen resultados m´inimamente confiables a partir de 200 casos de acuerdo al RMSE, los cuáles van mejorando a medida que se aumenta el tamaño de muestra.

Cuadro 4.3: Comparación de parámetros de intercepto de ítems que requieren dos atributos

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
1	RMSE	0.1669	0.0757	0.0505	0.0482	0.0420	0.0354
	Sesgo	0.0195	-0.0139	0.0061	0.0043	-0.0041	-0.0021
7	RMSE	0.1220	0.1056	0.0754	0.0557	0.0497	0.0363
	Sesgo	-0.0419	-0.0511	-0.0193	-0.0121	-0.0103	-0.0034
8	RMSE	0.1054	0.0528	0.0466	0.0399	0.0379	0.0226
	Sesgo	-0.0016	-0.0126	0.0059	-0.0077	-0.0088	0.0016
10	RMSE	0.0389	0.0276	0.0233	0.0180	0.0179	0.0116
	Sesgo	0.0039	0.0063	0.0008	0.0007	-0.0026	0.0017

De igual manera, los cuadros 4.4 y 4.5 muestran los resultados para los parámetros de los efectos de la presencia de los atributos en la probabilidad de éxito. En estos casos notamos que los mejores resultados se obtienen a partir de un tamaño de muestra de 300.

Cuadro 4.4: Comparación de parámetros del 1er efecto principal de ítems que requieren dos atributos

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
1	RMSE	0.2707	0.1237	0.0927	0.0885	0.0823	0.0602
	Sesgo	-0.0424	-0.0078	-0.0023	-0.0010	0.0020	0.0001
7	RMSE	0.1495	0.0876	0.0659	0.0555	0.0572	0.0327
	Sesgo	-0.0265	-0.0186	-0.0072	-0.0139	-0.0089	-0.0173
8	RMSE	0.1046	0.0725	0.0541	0.0426	0.0349	0.0258
	Sesgo	-0.0129	0.0077	-0.0142	0.0082	0.0072	-0.0037
10	RMSE	0.0744	0.0643	0.0509	0.0417	0.0427	0.0239
	Sesgo	0.0015	-0.0109	0.0166	0.0022	0.0041	0.0033

Cuadro 4.5: Comparación de parámetros del 2do efecto principal de ítems que requieren dos atributos

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
1	RMSE	0.2426	0.1139	0.0824	0.0677	0.0699	0.0442
	Sesgo	-0.0171	0.0214	0.0035	-0.0095	0.0184	-0.0003
7	RMSE	0.1477	0.1243	0.0814	0.0708	0.0674	0.0469
	Sesgo	0.0826	0.0715	0.0393	0.0343	0.0218	0.0152
8	RMSE	0.1196	0.0846	0.0600	0.0506	0.0536	0.0296
	Sesgo	-0.0027	0.0115	0.0051	0.0002	-0.0020	-0.0023
10	RMSE	0.0913	0.0607	0.0470	0.0396	0.0330	0.0224
	Sesgo	-0.0028	0.0011	-0.0035	0.0012	0.0004	-0.0018

El ítem 1 es el único al que se le aplica un modelo G-DINA saturado, por lo que solo éste presenta el efecto interacción  $\delta_{12}$ , es decir el incremento en la probabilidad de éxito dado que se poseen simultáneamente los dos atributos. A diferencia de los resultados anteriores, el Cuadro 4.6 señala que para este parámetro se obtiene estimaciones confiables a partir de un tamaño de muestra de 500. Este resultado concuerda con lo señalado por diversos autores, que indican que modelos saturados requieren mayor tamaño de muestra.

Cuadro 4.6: Comparación de parámetros del efecto interacción de ítems que requieren dos atributos

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
1	RMSE	0.3781	0.1705	0.1327	0.1140	0.1036	0.0759
	Sesgo	0.0529	-0.0116	-0.0051	0.0186	-0.0113	0.0069

Para el grupo de ítems cuyos parámetros fueron estimados con los modelos LLM y RRUM se obtienen parámetros componentes cuyos valores no se encuentran en la métrica de los parámetros  $\delta$  por lo que no sería correcto evaluarlos de la misma manera. Por esta razón, para este grupo de ítems, no se evaluarán los parámetros componentes, sino los que se expresan en términos de probabilidad. Así, los parámetros examinados son:  $P(00)$ , la probabilidad de responder afirmativamente cuando no se posee ninguno de los atributos requeridos;  $P(10)$  (o  $P(01)$ ) representa la probabilidad de responder afirmativamente cuando se posee el primer (o segundo) atributo requerido; y  $P(11)$  representa la probabilidad de responder el ítem afirmativamente cuando se posee los dos atributos.

El cuadro 4.7 muestra los resultados para el primer parámetro.

Cuadro 4.7: Comparación del intercepto  $P(00)$ : probabilidad de responder afirmativamente al ítem dado que no se posee ninguno de los atributos requeridos

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
2	RMSE	0.1161	0.0751	0.0557	0.0384	0.0328	0.0150
	Sesgo	0.0783	0.0403	0.0291	0.0185	0.0159	0.0071
5	RMSE	0.0648	0.0504	0.0375	0.0304	0.0264	0.0186
	Sesgo	0.0026	-0.0036	-0.0017	-0.0018	-0.0022	0.0005
12	RMSE	0.0702	0.0496	0.0372	0.0327	0.0342	0.0190
	Sesgo	-0.0038	-0.0020	0.0011	0.0020	0.0035	-0.0007
13	RMSE	0.0783	0.0473	0.0419	0.0329	0.0279	0.0234
	Sesgo	-0.0027	0.0046	-0.0035	0.0001	-0.0008	-0.0010

Para el intercepto se obtienen valores de sesgos bajos, y que van disminuyendo a medida que aumenta el tamaño de muestra. En todos los ítems se consiguen resultados confiables a partir de un tamaño de muestra de 200 de acuerdo al valor del RMSE.

Los cuadros 4.8 y 4.9, muestran los resultados de las probabilidades cuando se posee al menos uno de los atributos requeridos

Cuadro 4.8: Comparación de los parámetros  $P(10)$ : Probabilidad de responder afirmativamente al ítem dado que se posee el primer atributo requerido

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
2	RMSE	0.1487	0.0929	0.0716	0.0465	0.0403	0.0194
	Sesgo	0.1028	0.0513	0.0370	0.0230	0.0196	0.0092
5	RMSE	0.1554	0.0931	0.0644	0.0578	0.0513	0.0356
	Sesgo	0.0077	-0.0060	-0.0076	-0.0067	-0.0082	-0.0028
12	RMSE	0.1050	0.0600	0.0518	0.0427	0.0408	0.0269
	Sesgo	-0.0146	0.0027	0.0014	0.0060	0.0005	0.0000
13	RMSE	0.0936	0.0635	0.0519	0.0479	0.0437	0.0280
	Sesgo	0.0048	0.0050	0.0003	-0.0023	-0.0020	0.0033

Cuadro 4.9: Comparación de los parámetros P(01): Probabilidad de responder afirmativamente al ítem dado que se posee el segundo atributo requerido

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
2	RMSE	0.2257	0.1930	0.1290	0.1238	0.0988	0.0804
	Sesgo	-0.0826	-0.0709	-0.0335	-0.0346	-0.0154	-0.0122
5	RMSE	0.1084	0.0576	0.0415	0.0365	0.0334	0.0221
	Sesgo	0.0107	-0.0037	0.0029	0.0012	0.0024	0.0044
12	RMSE	0.0772	0.0529	0.0383	0.0352	0.0374	0.0209
	Sesgo	0.0050	-0.0039	0.0019	-0.0009	0.0050	-0.0007
13	RMSE	0.0998	0.0673	0.0493	0.0440	0.0413	0.0302
	Sesgo	-0.0006	-0.0011	-0.0054	-0.0032	0.0031	0.0002

Notamos que para estos dos parámetros tanto los sesgos, como los valores del RMSE, son mayores que los del parámetro de intercepto. Sin embargo, se obtienen valores suficientemente precisos a partir de un número de muestra de 200 en tres de los 4 ítems. Para el ítem 2 se requiere al menos de 500 casos. Al igual que el parámetro anterior, se obtienen los mejores resultados cuando el tamaño es de 1000.

Finalmente, el cuadro 4.10 presenta los resultados para el parámetro  $P(11)$

Cuadro 4.10: Comparación de los parámetros P(11): Probabilidad de responder afirmativamente al ítem dado que se posee los dos atributos requeridos

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
2	RMSE	0.1411	0.0943	0.0871	0.0760	0.0567	0.0475
	Sesgo	-0.0831	-0.0565	-0.0507	-0.0417	-0.0325	-0.0269
5	RMSE	0.1194	0.0720	0.0588	0.0538	0.0461	0.0322
	Sesgo	0.0021	-0.0005	0.0040	0.0002	0.0015	0.0043
12	RMSE	0.0844	0.0562	0.0418	0.0393	0.0292	0.0237
	Sesgo	0.0042	-0.0017	0.0031	-0.0014	0.0026	0.0000
13	RMSE	0.0809	0.0464	0.0392	0.0320	0.0297	0.0246
	Sesgo	0.0032	-0.0025	0.0010	-0.0033	0.0015	0.0032

Para este parámetro, notamos que se obtienen sesgos pequeños en todos los ítems, y una estimación precisa para los ítems a partir de un tamaño de muestra de 200 ( $RMSE < 0.1$ ).

#### 4.2.3. Ítems que miden tres atributos

Solo el ítem 3 evalúa la presencia de los tres atributos. Para este ítem se ajustó el modelo aditivo RRUM por lo que presenta 4 ( $K_j^* + 1$ ) parámetros componentes a estimar: el intercepto y tres efectos principales. Al igual que en el caso anterior y por la misma razón, para este ítem también se analizarán los estimados de los 8 ( $2^{K_j^*}$ ) parámetros principales. Así, los parámetros estimados son:  $P(000)$ , la probabilidad de responder afirmativamente el ítem cuando no se posee ninguno de los tres atributos requeridos;  $P(100)$ ,  $P(010)$  y  $P(001)$

representan la probabilidad de responder afirmativamente cuando se posee solo el primer, segundo o tercer atributo requerido, respectivamente;  $P(110)$ ,  $P(101)$ ,  $P(011)$  representan las probabilidades de responder afirmativamente cuando se poseen dos de los tres atributos requeridos y finalmente  $P(111)$  es la probabilidad de responder afirmativamente cuando se poseen los tres atributos requeridos.

El cuadro 4.11 muestra los resultados para los 8 parámetros del ítem 3 donde observamos que a partir de un tamaño de muestra de 200 se obtienen resultados precisos para 7 de estos valores. A partir de 300 casos obtenemos resultados precisos para todos los valores.

Cuadro 4.11: Comparación de los parámetros del ítem que evalúa 3 atributos

Ítem	Indicador	Tamaño de muestra					
		N=100	N=200	N=300	N=400	N=500	N=1000
P000	RMSE	0.0584	0.0485	0.0435	0.0433	0.0453	0.0366
	Sesgo	-0.0460	-0.0363	-0.0368	-0.0354	-0.0268	-0.0274
P100	RMSE	0.0632	0.0556	0.0514	0.0511	0.0504	0.0462
	Sesgo	0.0066	0.0140	0.0127	0.0137	0.0210	0.0222
P010	RMSE	0.1751	0.1162	0.0916	0.0720	0.0648	0.0487
	Sesgo	-0.0484	-0.0380	-0.0379	-0.0172	-0.0302	-0.0171
P001	RMSE	0.1077	0.0822	0.0740	0.0777	0.0872	0.0674
	Sesgo	0.0116	0.0256	0.0252	0.0207	0.0445	0.0350
P110	RMSE	0.1694	0.1079	0.0830	0.0669	0.0609	0.0401
	Sesgo	-0.0205	-0.0243	-0.0272	-0.0076	-0.0241	-0.0061
P101	RMSE	0.1130	0.0973	0.0892	0.0915	0.0968	0.0843
	Sesgo	0.0270	0.0352	0.0330	0.0276	0.0492	0.0427
P011	RMSE	0.1549	0.0882	0.0659	0.0557	0.0377	0.0544
	Sesgo	-0.0711	-0.0361	-0.0264	-0.0254	-0.0162	-0.0235
P111	RMSE	0.0442	0.0296	0.0185	0.0191	0.0126	0.0087
	Sesgo	-0.0133	-0.0102	-0.0060	-0.0079	-0.0048	-0.0036

En resumen, un tamaño de muestra de 100 casos, no es recomendable para un estudio con estas características. A partir de 200 casos se obtiene resultados precisos para al menos el 80 % de los ítems; mientras que si se usan 300 casos, este porcentaje se incrementa a 90 %. Se obtienen los mejores resultados para el parámetro de intercepto.

## Capítulo 5

### Aplicación

En el presente trabajo se aplica el modelo G-DINA al campo de la psiquiatría, usándola como herramienta para obtener un diagnóstico de desórdenes mentales. Para este fin se emplea información de los consultorios de psiquiatría del Hospital Cayetano Heredia (Lima, Perú). En este capítulo se describen los desórdenes mentales medidos, la importancia de un correcto diagnóstico y las investigaciones previas realizadas. Se presentan además los resultados obtenidos usando el modelo G-DINA mostrando los parámetros obtenidos de cada uno de los ítems. Se muestra también la interpretación de estos resultados y la distribución de perfiles existentes en la muestra estudiada. Por último se presentan las conclusiones de la aplicación del modelo.

#### 5.1. Introducción

Es importante señalar que si bien los modelos de diagnóstico cognitivo, a los cuáles pertenece el modelo G-DINA, se usan mayormente en el área educativa, no existe nada inherente a ellos que limite su uso en otras áreas (De la Torre et al., 2015). Sin embargo, mucha de la terminología empleada está orientada al campo de la educación. Para evitar cualquier tipo de confusión y buscando una generalización de la técnica, se sigue la pauta de algunos autores. Así, para el presente capítulo, se reemplaza el término “*Modelos de Diagnóstico Cognitivo*” por el de “*Modelos de Diagnóstico de Clasificación*”; y se utilizará la palabra “*desorden*” en lugar de la palabra “*atributo*” usada en los capítulos previos.

En los últimos años, el uso de estos modelos se ha incrementado tanto en el área de la educación como en el de la psicología y psiquiatría. Así por ejemplo, en estos últimos, se ha usado los CDM como herramienta de diagnóstico de desórdenes psicológicos relacionados con el juego (Templin y Henson, 2006) y de desórdenes mentales tales como ansiedad, depresión, desórdenes del pensamiento y trastorno somatomorfo (De la Torre et al., 2015). Este último estudio fue realizado con datos obtenidos de pacientes de Bélgica, usando el instrumento de medición “*MCMI – III Inventario Clínico Multiaxial de MILLON*”, que evalúa trastornos de la personalidad y otros síndromes clínicos.

El presente trabajo busca también realizar un diagnóstico de estos desórdenes mentales en un hospital del Perú, ya que, tal como ocurre a nivel mundial, estos afectan también a un porcentaje alto de la población peruana (Integración, 2016). Actualmente, el Ministerio de Salud usa el cuestionario SRQ-18 como prueba de tamizaje de estos desórdenes mentales, permitiendo identificar grupos de síntomas presentes y si el paciente alcanza cierto puntaje, deberá ser evaluado por el médico del establecimiento quien confirmará o no la presencia

del desorden (Minsa, 2015). Por lo tanto, esta prueba se presenta como una herramienta de descarte, que permite de manera rápida pero no tan específica, conocer el estado en que se encuentran los pacientes. Es por esta razón que se propone usar los resultados de este cuestionario, que se emplea en los hospitales a nivel nacional, pero aplicando el modelo G-DINA como herramienta de diagnóstico. De esta manera, se busca aprovechar los recursos ya existentes y disponibles del Ministerio de Salud para crear un perfil de los pacientes, que tiene como ventaja que ha sido elaborado con un modelo que tiene en cuenta la comorbilidad de los síndromes.

## 5.2. Instrumentos de medición

El instrumento de medición usado en el presente trabajo es el Cuestionario de Auto-reporte o Self Reporting Questionnaire (SRQ) (Ver anexo A). Este test puede ser autoplicado o llenado por el personal autorizado del hospital y permite indagar sobre situaciones específicas que la persona haya experimentado en los últimos 30 días. Ha sido diseñado por la Organización Mundial de la Salud (OMS por sus siglas en inglés) con la finalidad de ampliar la detección y atención de los problemas de salud mental en los servicios de atención primaria, especialmente en los países en desarrollo (Romero et al., 2016). Dada su importancia y relevancia a nivel mundial, ha sido validado en numerosas ocasiones en diversos estudios realizados entre 1978 y 1993 (WHO, 1994).

El SRQ-18 contiene 28 preguntas dicotómicas que se refieren a síntomas depresivos, ansiosos, trastorno psicótico y problemas con el consumo del alcohol. Las primeras 18 preguntas se refieren a trastornos depresivos o ansiosos, donde 9 o más respuestas positivas determinan que el entrevistado tiene una alta probabilidad de presentar una enfermedad mental. Las preguntas 19 a 22 son indicativas de un trastorno psicótico. Una sola respuesta positiva entre estas cuatro determina un “caso”. Una respuesta afirmativa a la pregunta 23 indica una alta probabilidad de sufrir un trastorno convulsivo. Las últimas preguntas se refieren a problemas relacionados al consumo de alcohol

## 5.3. Muestra

Los datos empleados en esta investigación consisten en la información de 192 pacientes con trastornos mentales que acudieron a los consultorios de psiquiatría del Hospital Cayetano Heredia (Lima, Perú), a quienes se les aplicó el cuestionario SRQ-18. Estos datos fueron recolectados en el marco de la investigación “Estructura de la sintomatología psiquiátrica según el SRQ-18 en pacientes ambulatorios con trastornos mentales evaluados en un hospital general de Lima” (Romero et al., 2016).

Los pacientes evaluados tenían edades entre los 18 a 65 años (58 % mujeres) y no presentaban problemas físicos o mentales que les impidieran comprender y responder correctamente. Las pruebas fueron administradas entre agosto y diciembre del 2015. A fin de evaluar solo los síndromes mencionados, se excluyen las preguntas 23 a 28 del test, que indican problemas con el consumo del alcohol. Se excluyeron también 28 registros que presentaban alguna respuesta incompleta. Para mayor detalle de la muestra usada, revisar el estudio de (Romero et al., 2016).

#### 5.4. Análisis de los datos

Para poder aplicar el modelo G-DINA a los datos y obtener los parámetros señalados, necesitamos definir cuáles son los desórdenes o síndromes que el paciente podría presentar. La siguiente tabla presenta una breve descripción de los desórdenes evaluados:

Cuadro 5.1: Descripción de los síndromes evaluados en la prueba SRQ-18

Parámetro	Síndrome	Descripción
Uno	Depresión (D)	Trastorno mental que se caracteriza por una profunda tristeza, decaimiento anímico, pérdida de interés por todo y disminución de las funciones psíquicas.
Dos	Ansiedad (A)	Estado mental que se caracteriza por una gran inquietud, intensa excitación y una extrema inseguridad.
Tres	Psicosis (P)	Estado mental grave que se caracteriza por una alteración global de la personalidad acompañada de un trastorno grave del sentido de la realidad.
Cuatro	Síntomas físicos (SF)	Estado relacionado con síntomas físicos sin explicación médica como dormir mal, pobre apetito, problemas digestivos, etc.

A continuación, la matriz Q es definida por una psicóloga experta. Dado que nuestro test presente 22 ítems que miden 4 desórdenes, la matriz Q tiene una dimensión de  $22 \times 4$ . En el siguiente cuadro se presenta la matriz Q empleada.

Cuadro 5.2: Matriz Q

Ítem	D	A	P	SF
1	0	1	0	1
2	1	0	0	1
3	1	1	0	1
4	0	1	0	0
5	0	1	0	1
6	0	1	0	0
7	0	1	0	1
8	1	1	0	0
9	1	0	0	0
10	1	1	0	0
11	1	0	0	0
12	1	1	0	0
13	1	1	0	0
14	1	0	0	0
15	1	0	0	0
16	1	0	0	0
17	1	0	0	0
18	1	1	0	0
19	0	0	1	0
20	0	0	1	0
21	0	0	1	0
22	0	0	1	0

A partir de la matriz Q construimos el Cuadro 5.3 que resume el número de desórdenes medido en cada conjunto de síntomas. Podemos observar que existen 12 ítems que miden solo un desorden, 9 ítems que evalúan 2 desórdenes y un ítem que mide 3 desórdenes. Notamos también que para diagnosticar la presencia de depresión y ansiedad existe mayor cantidad de ítems (13 y 11 respectivamente) que para la psicosis y los síntomas físicos (4 y 5 respectivamente).

Cuadro 5.3: Número de ítems evaluados en cada síndrome

Número de desórdenes	Desorden				Número de ítems
	D	A	P	SF	
Uno	6	2	4	0	12
Dos	6	8	0	4	9
Tres	1	1	0	1	1

Es necesario dejar constancia que la base de datos recibida no incluye la identificación

individual de los participantes que respondieron el cuestionario. El protocolo fue revisado por el Comité Institucional de Ética de la Universidad Peruana Cayetano Heredia.

### 5.5. Selección del modelo para cada ítem

Si bien el modelo G-DINA puede proporcionar un mejor ajuste de los datos, en ocasiones otros modelos cognitivos específicos son preferibles por distintas razones. En primer lugar, los modelos reducidos requieren menor cantidad de datos y permiten interpretaciones más directas. En segundo lugar, por el principio de parsimonia se prefieren modelos más sencillos si la diferencia con el mejor modelo no es significativamente grande.

Para la selección del mejor modelo, existe el enfoque a nivel del test y a nivel de ítem. En el primer caso se usan distintas medidas como el Criterio de información de Akaike (AIC) y el Criterio de Información Bayesiano (BIC), que serán detalladas más adelante. A nivel de ítem, (Wenchao et al., 2016) usan el test de Wald y el concepto de disimilitud para la elección del mejor modelo para cada ítem.

El proceso empieza ajustando el modelo G-DINA para los datos y calculando, para cada ítem, el estadístico de Wald, los p-valores y las disimilitudes entre los modelos reducidos y el modelo G-DINA. Para el presente trabajo, los modelos reducidos evaluados serán: DINA, DINO, A-CDM, LLM Y R-RUM que fueron descritos en los capítulos previos.

Los cuadros 5.4 y 5.5 muestran los valores del estadístico de Wald y sus p-valores para los 5 modelos estudiados. Así, para cada ítem, se descartan los modelos cuyos p-valores sean menores a 0.05, que en este caso son en su mayoría los modelos DINA y DINO. Por el contrario, se prefieren los modelos cuyos estadísticos de Wald sean mayores.

Cuadro 5.4: Valores del Estadístico de Wald para los 5 modelos evaluados

Ítem	Estadístico de Wald				
	DINA	DINO	A-CDM	LLM	R-RRUM
1	10.58	1.33	0.38	0.71	1.37
2	3.15	4.66	0.23	0.00	0.00
3	16.30	5.03	4.67	0.00	6.61
5	6.51	2.49	0.62	0.00	0.00
7	6.32	6.15	0.19	0.20	1.05
8	2.38	4.92	0.08	0.05	0.05
10	7.24	10.75	0.02	1.58	2.61
12	5.69	20.34	1.50	1.75	0.82
13	8.05	8.51	0.00	0.00	0.37
18	6.50	0.18	1.28	1.17	2.35

Cuadro 5.5: Los p-valores del test de Wald

Ítem	p-values				
	DINA	DINO	A-CDM	LLM	R-RRUM
1	0.01	0.51	0.54	0.40	0.24
2	0.21	0.10	0.63	0.99	0.99
3	0.01	0.54	0.32	1.00	0.16
5	0.04	0.29	0.43	0.99	0.99
7	0.04	0.05	0.66	0.65	0.31
8	0.30	0.09	0.77	0.82	0.83
10	0.03	0.00	0.90	0.21	0.11
12	0.06	0.00	0.22	0.19	0.37
13	0.02	0.01	0.98	0.98	0.54
18	0.04	0.91	0.26	0.28	0.13

Adicionalmente, el cuadro 5.6 muestra las disimilitudes, que pueden interpretarse como una medida que cuantifica qué tan diferente es el modelo G-DINA y los otros 5 modelos reducidos. Valores pequeños de disimilitudes nos indican que el modelo reducido puede aproximar bien el modelo G-DINA, mientras que valores grandes sugieren que no (Wenchao et al., 2016). Los valores mostrados en el cuadro 5.6 indican que los modelos DINA y DINO son los que más difieren del modelo G-DINA excepto en el ítem 18, el cual tiene como mejor opción al modelo DINO. La interpretación de estos resultados, coinciden con los del test de Wald y analizando ambos valores se procede a seleccionar el mejor modelo para cada ítem.

Cuadro 5.6: Disimilitudes entre el modelo G-DINA y los modelos reducidos

Ítem	Disimilaridades				
	DINA	DINO	A-CDM	LLM	R-RRUM
1	0.41	0.25	0.16	0.17	0.19
2	0.50	0.70	0.20	0.00	0.00
3	2.64	2.17	2.20	1.80	2.20
5	0.47	0.63	0.34	0.27	0.19
7	0.51	0.39	0.12	0.11	0.15
8	0.32	0.40	0.09	0.05	0.04
10	0.44	0.40	0.03	0.11	0.13
12	0.45	0.64	0.28	0.28	0.15
13	0.50	0.50	0.01	0.00	0.08
18	0.41	0.10	0.32	0.22	0.32

El cuadro 5.7 muestra la selección del mejor modelo para los ítems que miden más de un atributo. De estos 10 ítems, uno será modelado usando el modelo G-DINA mientras que 8 de ellos usan modelos aditivos (A-CDM, R-RUM o LLM). Solo uno de ellos usa el modelo DINO.

Cuadro 5.7: Selección del mejor modelo para cada ítem

Ítem	Modelo seleccionado
1	G-DINA
2	R-RUM
3	R-RUM
5	LLM
7	A-CDM
8	A-CDM
10	A-CDM
12	R-RUM
13	LLM
18	DINO

### 5.6. Estimación de los parámetros de los ítems

Tanto la selección del modelo para cada ítem, como la estimación de los parámetros estudiados, se realizó con el paquete G-DINA (Wenchao et al., 2017) del software estadístico R.

Luego de la selección del mejor modelo para cada ítem, se procede con la estimación de la combinación de modelos con los métodos estudiados.

Los cuadros del (5.8) al (5.10) muestran los parámetros estimados para los ítems que miden 1, 2 y 3 desórdenes respectivamente. Los ítems que miden un solo desorden presentan solo dos parámetros: la probabilidad de que los pacientes que no presentan el desorden, marquen afirmativamente la pregunta y la probabilidad de responder afirmativamente a la pregunta dado que el paciente presenta el desorden medido por el ítem. Así, para el ítem 4 que evalúa síntomas de ansiedad, la probabilidad de que un paciente que no presenta ansiedad responda afirmativamente la pregunta es 0.187; mientras que si el paciente presentara ansiedad esta probabilidad sube a 0.72.

Según De la Torre et al. (2015), se usa además la diferencia entre estos dos parámetros que puede ser visto como un indicador del poder de discriminación del ítem. En su artículo menciona que, aunque no existe un valor establecido para considerar un ítem como discriminatorio, usa el valor de 0.4, que es el valor que se usa en este trabajo. Esto quiere decir que los pacientes que presentan un desorden en particular tienen al menos 40 % más de probabilidad de responder afirmativamente a un ítem que aquellos que lo poseen.

Cuadro 5.8: Parámetros estimados de los ítems que miden un desorden

Desorden	Ítem	Combinación de desórdenes		Disc
		0 (Ausente)	1 (Presente)	
A	4	0.187	0.728	0.541
	6	0.380	0.960	0.580
D	9	0.327	0.889	0.563
	11	0.103	0.856	0.753
	14	0.150	0.631	0.482
	15	0.087	0.783	0.697
	16	0.432	0.909	0.478
P	17	0.138	0.580	0.441
	19	0.199	0.806	0.607
	20	0.577	0.656	0.078
	21	0.172	0.722	0.550
	22	0.068	0.335	0.267

Notamos que los ítems 20 y 22, que evalúan síntomas de psicosis, no son considerados como discriminatorios al presentar valores de 0.078 y 0.27 respectivamente, ambos menores que 0.4. Los otros dos síntomas, que evalúan síntomas de psicosis, presentan valores discriminatorios relativamente altos. Con respecto a los ítems que evalúan los síntomas depresivos, es el ítem 11 el que presenta mayor índice discriminatorio (0.76). Los otros ítems, 9, 14, 15, 16 y 17 presentan valores también altos de discriminación entre 0.43 y 0.69. En este grupo de ítems, solo dos evalúan síntomas ansiosos, ambos con valores altos de discriminación. En resumen, este grupo de ítems que evalúa solo un atributo es considerado como discriminante.

Cuadro 5.9: Parámetros estimados de los ítems que miden dos desórdenes

Desorden	Ítem	Combinación de desórdenes				Disc
		00	10	01	11	
A-SF	1	0.158	0.533	0.187	0.615	0.457
A-SF	2	0.000	0.001	0.812	1.000	0.999
A-SF	5	0.124	0.489	0.164	0.571	0.447
A-SF	7	0.252	0.366	0.683	0.797	0.546
A-SF	8	0.211	0.634	0.319	0.742	0.531
A-SF	10	0.026	0.438	0.187	0.599	0.573
A-D	12	0.315	0.798	0.346	0.876	0.561
A-D	13	0.170	0.667	0.323	0.824	0.809
A-D	18	0.277	0.662	0.662	0.662	0.386

El cuadro 5.9 muestra los parámetros estimados para el grupo de ítems que miden dos atributos. Para este grupo de ítems se calcula también el índice de discriminación y basados en este valor vemos que todas las preguntas, con excepción del ítem 18, se pueden considerar discriminatorias ya que presentan un valor mayor a 0.4. Se calcula además otros indicadores

como la diferencia de  $P_j(10)$  y  $P_j(01)$  que nos ayudan a determinar si alguno de los pacientes que presentan alguno de los dos desórdenes tiene más probabilidad de presentar el síntoma. Por ejemplo, analizando el ítem 13, notamos que los pacientes con depresión tienen 0.3 más de probabilidad que aquellos pacientes con solo ansiedad, de responder afirmativamente.

El ítem 1 es el único que mide los 3 desórdenes estudiados, y sus resultados se muestran en el cuadro 5.10. Notamos que los pacientes que presentan los tres desórdenes tienen 60 % más de probabilidad de responder afirmativamente que aquellos que no presentan ninguno.

Cuadro 5.10: Parámetros estimados del ítem que mide tres desórdenes (Ítem 3)

Ninguno	D	A	SF	D-A	D-SF	A-SF	D-A-SF
0.112	0.112	0.547	0.205	0.547	0.205	0.999	0.999

### 5.7. Estimación de perfiles

El cuadro 5.11 muestra la distribución de los pacientes en los distintos grupos de perfiles existentes. Notamos que el 39 % de los pacientes evaluados, pertenecen al grupo (0000) que es aquel que no presenta ningún síntoma. Además, aproximadamente un 20 % pertenece al perfil de los pacientes que presentan ansiedad, depresión y síntomas físicos. Existe otro grupo importante que pacientes que presentan ansiedad y síntomas físicos.

Cuadro 5.11: Perfil de desórdenes mentales y prevalencia

Perfil de desórdenes	Prevalencia( %)
0000	38.5 %
1000	1.8 %
0100	1.5 %
0010	1.3 %
0001	0.8 %
1100	2.1 %
1010	0.0 %
1001	2.2 %
0110	14.6 %
0101	3.6 %
0011	0.0 %
1110	20.1 %
1101	0.7 %
1011	1.5 %
0111	0.0 %
1111	11.2 %

### 5.8. Diagnóstico de los pacientes

Uno de los principales objetivos del modelo G-DINA es dar un diagnóstico al entrevistado. Así, este modelo nos permite obtener la probabilidad de cada individuo de pertenecer a cada uno de los 16 perfiles existentes. El cuadro 5.12 muestra estas probabilidades agrupadas en los cuatro desórdenes evaluados. Estableciendo como punto de corte 0.5, podemos decir que el primer individuo se encuentra en el perfil (D=1, A=1, SF=1, P=0). Es decir presenta síntomas de Depresión, ansiedad, y síntomas físicos.

Cuadro 5.12: Probabilidades de presentar los desórdenes para 10 pacientes

Paciente	Probabilidades de los desórdenes			
	D	A	P	SF
1	0.993	0.994	0.984	0.001
2	0.001	0.954	0.953	0.009
3	0.977	0.827	0.418	0.999
4	0.001	0.039	0.778	0.001
5	0.001	0.998	0.991	0.001
6	0.999	0.999	0.999	0.996
7	0.001	0.003	0.004	0.001
8	0.155	0.993	0.145	0.999
9	0.000	0.026	0.012	0.002
10	0.975	0.999	0.992	0.001

### 5.9. Comparación con otros modelos

El cuadro 5.13 presenta un resumen que compara el modelo saturado G-DINA con los modelos reducidos mencionados en este estudio: DINA, DINO, A-CDM, LLM, R-RUM y por último el modelo Mixto, que resulta de la selección del mejor modelo para cada ítem. Para evaluar el modelo que mejor ajuste a los datos observamos los indicadores Criterio de información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y la log-verosimilitud. Si bien el modelo saturado G-DINA el que presenta mejores valores según el loglik, muestra también los ajustes más débiles según el BIC. Según el AIC, es el modelo mixto el que presenta mejor desempeño, seguido por los modelos aditivos. Analizando el BIC, es también el modelo mixto el que presenta mejores resultados, seguidos por los modelos aditivos.

Cuadro 5.13: Resultados del ajuste de los modelos a los datos

Modelo	Número de parámetros	LogLik	Deviance	AIC	BIC	Chisq	df	p-value
G-DINA	83	-2,207.0	4414.0	4580.1	4846.0			
DINA	59	-2,259.9	4519.7	4637.7	4826.7	105.7	24	<0.001
DINO	59	-2,256.1	4512.1	4630.1	4819.1	98.1	24	<0.001
A-CDM	70	-2,221.9	4443.8	4583.8	4808.1	29.8	13	0.005
LLM	70	-2,221.7	4443.4	4583.4	4807.7	29.4	13	0.0051
R-RUM	70	-2,224.5	4448.9	4588.9	4813.2	34.9	13	0.001
MIXTO	70	-2,217.5	4435.0	4575.0	4799.2	20.9	13	0.07

### 5.10. Conclusiones

En el presente capítulo aplicamos el modelo G-DINA a los resultados del test SRQ-18 aplicado a 182 pacientes de psiquiatría del hospital Cayetano Heredia. Se obtuvo un diagnóstico para cada uno de ellos y una distribución de perfiles en la población. Se encontró que aproximadamente un 90 % de los ítems presentan valores altos de discriminación. Además, los ítems que miden dos desórdenes, presentan una alta diferencia en la probabilidad de marcar correctamente un ítem cuando se presenta dos desórdenes. Sin embargo, en el único ítem que mide tres desórdenes, los resultados de los efectos de las interacciones no son precisos, por lo que en ese caso sería preferible usar el modelo G-DINA pero definiendo estos efectos como cero (modelo Aditivo). La aplicación cumple con proporcionar un diagnóstico para cada uno de los pacientes y la probabilidad de presentar cada uno de los síndromes evaluados.

## Capítulo 6

### Conclusiones

En el presente trabajo se desarrolló el modelo G-DINA y se aplicó a un conjunto de datos de pacientes con trastornos mentales evaluados en consultorio externo de psiquiatría del Hospital Cayetano Heredia (Lima, Perú). Además se realizó un estudio de simulación para evaluar el efecto del tamaño de muestra en la estimación de los parámetros. Durante este proceso se llegaron a las siguientes conclusiones y recomendaciones para investigaciones futuras.

#### 6.1. Conclusiones

Respecto a la aplicación se observan las siguientes conclusiones:

- Se aplicó satisfactoriamente el modelo G-DINA a los datos, obteniendo resultados diagnósticos para cada uno de los pacientes. Además se logró obtener una distribución de los perfiles de la muestra estudiada, los que indican que un 38 % de ésta no presenta ninguno de los desórdenes estudiados, mientras que un 11 % presenta los 4 desórdenes estudiados simultáneamente. Se debe resaltar que existe un 46 % de pacientes que presentan ansiedad y síntomas físicos lo cual indica que ambos desórdenes tienen muchos síntomas en común. De igual manera, hay un 34 % que presenta los desórdenes de ansiedad y depresión al mismo tiempo, resultado que era esperado dado que estudios clínicos señalan que existe una alta comorbilidad entre ambas enfermedades.
- Aproximadamente un 90 % de los ítems presenta un alto índice de discriminación. Los ítems relacionados con el desorden Psicosis, son los que presentan menor poder para distinguir quiénes tienen el desorden y quiénes no.
- Se estimó un modelo final que ajusta un modelo distinto para cada ítem. De esta manera, se obtienen mejores resultados de desempeño y una mejor interpretación de los resultados, dado que presenta menor cantidad de parámetros que si se hubiera aplicado el modelo G-DINA a todos los ítems.

Respecto a la simulación se concluye lo siguiente

- Muestras con un tamaño de 100 casos, obtienen resultados pocos precisos en la mitad de los ítems. A partir de 200 casos, se obtienen resultados con buen desempeño para un 80 % de los ítems. Este porcentaje se incrementa a 90 % cuando la muestra es de tamaño 300. Se obtienen mejores resultados a medida que se incrementa el tamaño, por lo tanto en este caso, muestras con 1000 individuos presentan estimación más precisa.

## 6.2. Sugerencias para investigaciones futuras

- La aplicación de este estudio nació a raíz de la identificación de una oportunidad existente relacionada con los resultados del test SRQ-18 aplicado para un estudio previo. Actualmente el tema de la salud mental es abordado, entre otras formas, a través de este test, es decir, el MINSA ya evalúa a determinados pacientes con esta prueba. Por esta razón se sugiere, que dado que existen recursos disponibles en el estado, se realicen futuras investigaciones en los centros de salud estatales, y comparar los resultados obtenidos con los diagnósticos del personal calificado.
- [De la Torre et al. \(2015\)](#), recalcó la necesidad de tener solo los mejores métodos de diagnóstico en un campo tan importante como el de la salud mental aún teniendo en cuenta la dificultad de que no existiese un software para el cálculo de los parámetros. Esta dificultad que señalaba Jimmy de la Torre, actualmente ya no existe, debido al desarrollo del paquete estadístico GDINA, en el software de uso libre R, disponible a partir de abril del 2017. Por lo tanto, se recomienda el uso de este paquete para el desarrollo del modelo.
- Se recomienda validar la matriz Q usada, ya que los resultados son muy sensibles a estos valores. [De la Torre et al. \(2015\)](#) recalca la importancia de usar una matriz Q especificada con precisión para poder obtener resultados confiables. Actualmente el paquete GDINA, permite generar una matriz Q de acuerdo a los datos, pero esto debe ser evaluado cuidadosamente con el profesional experto en el área de estudio, en este caso, la psiquiatría.
- Al realizar la revisión de investigaciones que hacen uso del G-DINA en el área de la psiquiatría, no se encontró ninguna que fuera realizada a nivel nacional. Sin embargo, sí existen estudios donde se emplean diversos test, como el “Inventario Clínico Multiaxial de Millon”, para el diagnóstico de desórdenes mentales. Por lo tanto, existe la oportunidad para el uso del modelo G-DINA y se recomienda a los investigadores emplear el modelo G-DINA como complemento a sus respectivos trabajos.



## Apéndice A

### Cuestionario SRQ - 18

MINISTERIO DE SALUD E.S: \_\_\_\_\_  
 REGION DE SALUD TACNA HCL: \_\_\_\_\_

**CUESTIONARIO DE SINTOMAS (SRQ)**

FECHA: \_\_\_\_\_ EDAD: \_\_\_\_\_ SEXO: \_\_\_\_\_

NOMBRES Y APELLIDOS: \_\_\_\_\_  
 COLEGIO/SECCION/GRADO/TURNO: \_\_\_\_\_

Los datos se obtienen de los últimos 30 días

SINTOMAS	SI	NO
1. ¿Tiene dolores frecuentes de cabeza?		
2. ¿Tiene mal apetito?		
3. ¿Duerme mal?		
4. ¿Se asusta con facilidad?		
5. ¿Sufre de temblor de manos?		
6. ¿Se siente nervioso, tenso?		
7. ¿Sufre de mala digestión?		
8. ¿Es incapaz de pensar con claridad?		
9. ¿Se siente triste?		
10. ¿Llora usted con mucha frecuencia?		
11. ¿Tiene dificultad en disfrutar de sus actividades diarias?		
12. ¿Tiene dificultad para tomar decisiones?		
13. ¿Tiene dificultad en hacer su trabajo? ¿Sufre usted con su trabajo?		
14. ¿Es incapaz de desempeñar un papel útil en su vida?		
15. ¿Ha perdido interés en las cosas?		
16. ¿Se siente aburrido?		
17. ¿Ha tenido la idea de acabar con su vida?		
18. ¿Se siente cansado todo el tiempo?		
19. ¿Siente usted que alguien ha tratado de herirlo en alguna forma?		
20. ¿Es usted una persona mucho más importante que lo que piensan los demás?		
21. ¿Ha notado interferencias o algo raro en sus pensamientos?		
22. ¿Oye voces sin saber de donde vienen o que otras personas no pueden oír?		
23. ¿Ha tenido convulsiones, ataques o caídas al suelo, con movimientos de brazos y piernas; con mordeduras de lengua o pérdida del conocimiento?		
24. ¿Alguna vez le ha parecido a su familia, sus amigos, su médico o a su sacerdote que usted ha estado bebiendo demasiado licor?		
25. ¿Alguna vez a querido dejar de beber, pero no ha podido?		
26. ¿Ha tenido alguna vez dificultades en el trabajo (o estudio) a causa de la bebida, como beber en el trabajo o en el colegio, o faltar a ellos?		
27. ¿Ha estado en riñas o lo han detenido estando borracho?		
28. ¿Le a parecido alguna vez que usted bebía demasiado?		

**PUNTUACION:** Se obtiene sumando el número de respuestas positivas

Las 18 primeras preguntas se refieren a trastornos de leve o moderada intensidad como los depresivos, angustia y otros, 9 ó mas respuestas positivas en este grupo determinan que el entrevistado tiene una alta probabilidad de sufrir enfermedad mental, y por lo tanto se le considera un "caso"

Las preguntas 24 a 29 son indicativas de un trastorno psicótico (problemas relacionados con el alcohol); una sola respuesta positiva entre las 4 determina un "caso"

La respuesta positiva a la pregunta 23 indica alta probabilidad de sufrir un trastorno convulsivo

Para determinar el síndrome al que podría corresponder, se verifica si las respuestas positivas caen dentro de uno de los siguientes grupos:

Depresión los números 2-3-9-10-11-12-13-14-15-16-17-18

Angustia los números 1-3-4-5-6-7-8-18

Síndrome convulsivo epilepsia el número 23

Psicosis los números 19-20-21-22

Alcoholismo los números 24-25-26-27-28

Nombre y apellido del entrevistador

Figura A.1: Cuestionario SRQ 18

## Apéndice B

### Códigos usados para la aplicación

```
##Librerías usadas##

install.packages("readxl")
install.packages("GDINA")
library(readxl)
library(xlsx)
library(GDINA)

#####
# Leyendo los datos
#####

SRQ_sinnombres_V1 <-
read_excel("E:/MAESTRIA_PUCP/III_CICLO/SEMINARIO_DE_TESIS/DATOS/SRQ_sinnombres_V1.xlsx")

#####
#Definiendo la matriz Q
#####

Q2<- matrix(c(0,1,1,0,0,0,0,1,1,1,1,1,1,1,1,1,0,0,0,0,
              1,0,1,1,1,1,1,1,0,1,0,1,1,0,0,0,0,1,0,0,0,0,
              0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,
              1,1,1,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),nrow=22)

#####
# Calculando el Test de Wald, disimilaridades
#####

MOD_GDINA = GDINA(SRQ_sinnombres_V1[10:31], Q2, model = "GDINA",sequential = FALSE,
att.dist = "saturated",randomseed = 123456)
wmod1r <- modelcomp(MOD_GDINA,DS=TRUE)

#####
#Extrayendo las disimilaridades y los p-values del test de Wald
#####
```

```

extract(wmodlr,"DS")
extract(wmodlr,"pvalues")

#####
#Definiendo los modelos seleccionados para cada ítem
#####

modelos=c("GDINA","RRUM","RRUM","GDINA","LLM","GDINA","ACDM","ACDM","GDINA",
          "ACDM","GDINA","RRUM","LLM","GDINA","GDINA","GDINA","GDINA","DINO",
          "GDINA","GDINA","GDINA","GDINA")

#####
#Ajustando el modelo final con los modelos seleccionados para cada ítem
#####

MOD_MIXTO = GDINA(SRQ_sinnombres_V1[10:31], Q2, model = modelos, sequential = FALSE,
att.dist = "saturated", randomseed = 123456)

#####
#Extrayendo los resultados para la aplicación #####
#####

write.xlsx(MOD_MIXTO$catprob.parm,"E:/MAESTRIA/TESIS/MIXTO.xlsx")
write.xlsx(extract(MOD_MIXTO,what="posterior.prob"),"E:/MAESTRIA/TESIS/Probpost.xlsx")
write.xlsx(extract(MOD_MIXTO,what="logposterior.i"),"E:/MAESTRIA/TESIS/ProbInd.xlsx")

#####
#Comparando los modelos con ANOVA
#####

MOD_DINA = GDINA(SRQ_sinnombres_V1[10:31], Q2, model = "DINA", randomseed = 123456)
MOD_DINO = GDINA(SRQ_sinnombres_V1[10:31], Q2, model = "DINO", randomseed = 123456)
MOD_ACDM = GDINA(SRQ_sinnombres_V1[10:31], Q2, model = "ACDM", randomseed = 123456)
MOD_LLM = GDINA(SRQ_sinnombres_V1[10:31], Q2, model = "LLM", randomseed = 123456)
MOD_RRUM = GDINA(SRQ_sinnombres_V1[10:31], Q2, model = "RRUM", randomseed = 123456)
comp = anova(MOD_DINA,MOD_GDINA,MOD_DINO,MOD_ACDM, MOD_LLM, MOD_RRUM,MOD_MIXTO)

#####
# Extrayendo los resultados para el ANOVA
#####

write.xlsx(comp$IC, "E:/MAESTRIA_PUCP/III_CICLO/SEMINARIO_DE_TESIS/DATOS/anova.xlsx")
write.xlsx(comp$LR, "E:/MAESTRIA_PUCP/III_CICLO/SEMINARIO_DE_TESIS/DATOS/lr.xlsx")

```

## Apéndice C

### Códigos usados para la simulación

```
dir <- "E:/MAESTRIA_PUCP/III_CICLO/SEMINARIO_DE_TESIS/RESULTADOS"

SRQ_sinnombres_V1 <-
read_excel("E:/MAESTRIA_PUCP/III_CICLO/SEMINARIO_DE_TESIS/DATOS/SRQ_sinnombres_V1.xlsx")
Q2<- matrix(c(0,1,1,0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,0,0,0,0,
              1,0,1,1,1,1,1,1,0,1,0,1,1,0,0,0,0,0,1,0,0,0,0,
              0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,
              1,1,1,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),nrow=22)
J<-nrow(Q2)
nrep = 150
modelos=c("GDINA", "RRUM", "RRUM", "GDINA", "LLM", "GDINA", "ACDM", "ACDM", "GDINA",
          "ACDM", "GDINA", "RRUM", "LLM", "GDINA", "GDINA", "GDINA", "GDINA", "DINO",
          "GDINA", "GDINA", "GDINA", "GDINA")

#####
#Para calcular y extraer los parámetros de la aplicación
#####

Map = GDINA(SRQ_sinnombres_V1[10:31],Q2,model=modelos,randomseed=123456)
delta.list2 = Map$delta.parm
catprob.list2 = Map$catprob.parm

a1=c(Map$delta.parm$`Item 4`[1],Map$delta.parm$`Item 6`[1],Map$delta.parm$`Item 9`[1],
Map$delta.parm$`Item 11`[1],Map$delta.parm$`Item 14`[1],Map$delta.parm$`Item 15`[1],
Map$delta.parm$`Item 16`[1],Map$delta.parm$`Item 17`[1],Map$delta.parm$`Item 18`[1],
Map$delta.parm$`Item 19`[1],Map$delta.parm$`Item 20`[1],Map$delta.parm$`Item 21`[1],
Map$delta.parm$`Item 22`[1])
a2=c(Map$delta.parm$`Item 4`[2],Map$delta.parm$`Item 6`[2],Map$delta.parm$`Item 9`[2],
Map$delta.parm$`Item 11`[2],Map$delta.parm$`Item 14`[2],Map$delta.parm$`Item 15`[2],
Map$delta.parm$`Item 16`[2],Map$delta.parm$`Item 17`[2],Map$delta.parm$`Item 18`[2],
Map$delta.parm$`Item 19`[2],Map$delta.parm$`Item 20`[2],Map$delta.parm$`Item 21`[2],
Map$delta.parm$`Item 22`[2])
gs <- matrix(c(a1,a2),ncol=2)
b1=c(Map$delta.parm$`Item 7`[1],Map$delta.parm$`Item 8`[1],Map$delta.parm$`Item 10`[1])
b2=c(Map$delta.parm$`Item 7`[2],Map$delta.parm$`Item 8`[2],Map$delta.parm$`Item 10`[2])
b3=c(Map$delta.parm$`Item 7`[3],Map$delta.parm$`Item 8`[3],Map$delta.parm$`Item 10`[3])
```

```

bb = matrix(c(b1,b2,b3),ncol=3)
c1=c(Map$delta.parm$`Item 1`[1])
c2=c(Map$delta.parm$`Item 1`[2])
c3=c(Map$delta.parm$`Item 1`[3])
c4=c(Map$delta.parm$`Item 1`[4])
cc = matrix(c(c1,c2,c3,c4),ncol=4)
d1=c(Map$catprob.parm$`Item 2`[1],Map$catprob.parm$`Item 5`[1],
Map$catprob.parm$`Item 12`[1],Map$catprob.parm$`Item 13`[1])
d2=c(Map$catprob.parm$`Item 2`[2],Map$catprob.parm$`Item 5`[2],
Map$catprob.parm$`Item 12`[2],Map$catprob.parm$`Item 13`[2])
d3=c(Map$catprob.parm$`Item 2`[3],Map$catprob.parm$`Item 5`[3],
Map$catprob.parm$`Item 12`[3],Map$catprob.parm$`Item 13`[3])
d4=c(Map$catprob.parm$`Item 2`[4],Map$catprob.parm$`Item 5`[4],
Map$catprob.parm$`Item 12`[4],Map$catprob.parm$`Item 13`[4])
dd = matrix(c(d1,d2,d3,d4),ncol=4)
e1=c(Map$catprob.parm$`Item 3`[1])
e2=c(Map$catprob.parm$`Item 3`[2])
e3=c(Map$catprob.parm$`Item 3`[3])
e4=c(Map$catprob.parm$`Item 3`[4])
e5=c(Map$catprob.parm$`Item 3`[5])
e6=c(Map$catprob.parm$`Item 3`[6])
e7=c(Map$catprob.parm$`Item 3`[7])
e8=c(Map$catprob.parm$`Item 3`[8])
ee = matrix(c(e1,e2,e3,e4,e5,e6,e7,e8),ncol=8)

#####
#Creando los objetos para guardar los resultados
#####

##para los ítems con 2 parámetros delta a estimar (identity link function)

resdelt0_1<- matrix(0,13,nrep);rmse_delt0_1<-matrix(0,13);
error0delt0_1<-matrix(0,13);resdelt1_1<- matrix(0,13,nrep);
rmse_delt1_1<-matrix(0,13);error0delt1_1<-matrix(0,13);
Int_1at <- matrix(0,13,nrep);Int_1at_ME <- matrix(0,13,nrep);

###Para los items con 3 parámetros delta a estimar (identity link function)

resdelt0_2<- matrix(0,3,nrep);rmse_delt0_2<-matrix(0,3);
error0delt0_2<-matrix(0,3);resdelt1_2<- matrix(0,3,nrep);
rmse_delt1_2<-matrix(0,3);error0delt1_2<-matrix(0,3);
resdelt2_2<- matrix(0,3,nrep);rmse_delt2_2<-matrix(0,3);
error0delt2_2<-matrix(0,3);Int_2at <- matrix(0,3,nrep)
Int_2at_ME<-matrix(0,3,nrep);Int_2at_ME2<-matrix(0,3,nrep)

###Para los items con 4 parámetros delta a estimar (identity link function)

resdelt0_3<- matrix(0,1,nrep);rmse_delt0_3<-matrix(0,1);

```

```

error_delt0_3<-matrix(0,1);res_delt1_3<- matrix(0,1,nrep);
rmse_delt1_3<-matrix(0,1);error_delt1_3<-matrix(0,1);
res_delt2_3<- matrix(0,1,nrep);rmse_delt2_3<-matrix(0,1);
error_delt2_3<-matrix(0,1);res_delt3_3<- matrix(0,1,nrep);
rmse_delt3_3<-matrix(0,1);error_delt3_3<-matrix(0,1);
Int_3at <- matrix(0,1,nrep);Int_3at_ME <- matrix(0,1,nrep);
Int_3at_ME2<-matrix(0,1,nrep);Int_3at_ME3<-matrix(0,1,nrep);

##Para los ítems con 4 parámetros principales a estimar (log o logit function)

resP00_2<- matrix(0,4,nrep);rmse_P00_2<-matrix(0,4);
errorP00_2<-matrix(0,4);resP10_2<- matrix(0,4,nrep);
rmse_P10_2<-matrix(0,4);errorP10_2<-matrix(0,4);
resP01_2<- matrix(0,4,nrep);rmse_P01_2<-matrix(0,4);
errorP01_2<-matrix(0,4);resP11_2<- matrix(0,4,nrep);
rmse_P11_2<-matrix(0,4);errorP11_2<-matrix(0,4)
Int_2at_P00 <- matrix(0,4,nrep);Int_2at_P10 <- matrix(0,4,nrep);
Int_2at_P01<-matrix(0,4,nrep);Int_2at_P11<-matrix(0,4,nrep)

##Para los ítems con 8 parámetros principales a estimar(log o logit function)

resP000_3<- matrix(0,1,nrep);rmse_P000_3<-matrix(0,1);
errorP000_3<-matrix(0,1);resP100_3<- matrix(0,1,nrep);
rmse_P100_3<-matrix(0,1);errorP100_3<-matrix(0,1);
resP010_3<- matrix(0,1,nrep);rmse_P010_3<-matrix(0,1);
errorP010_3<-matrix(0,1);resP001_3<- matrix(0,1,nrep);
rmse_P001_3<-matrix(0,1);errorP001_3<-matrix(0,1)
resP110_3<- matrix(0,1,nrep);rmse_P110_3<-matrix(0,1)
errorP110_3<-matrix(0,1);resP101_3<- matrix(0,1,nrep);
rmse_P101_3<-matrix(0,1);errorP101_3<-matrix(0,1);
resP011_3<- matrix(0,1,nrep);rmse_P011_3<-matrix(0,1);
errorP011_3<-matrix(0,1);resP111_3<- matrix(0,1,nrep);
rmse_P111_3<-matrix(0,1);errorP111_3<-matrix(0,1);
Int_3at_P000<- matrix(0,1,nrep);Int_3at_P100 <- matrix(0,1,nrep);
Int_3at_P010<-matrix(0,1,nrep);Int_3at_P001<-matrix(0,1,nrep);
Int_3at_P110<-matrix(0,1,nrep);Int_3at_P101<-matrix(0,1,nrep);
Int_3at_P011<-matrix(0,1,nrep);Int_3at_P111<-matrix(0,1,nrep);

###Escenario
Lista <- list(c("100"), c("200"), c("300"),c("400"),c("500"),c("1000"))

#####
###Simulación para ítems que requieren dos parámetros delta#####
#####
for (i in 1:length(Lista)) {
  ptm <- proc.time()
  N <- eval(parse(text=Lista[[i]][1]))

```

```

#Simulación
for (p in 1:nrep)
{sim <- simGDINA(N=N,Q2,delta.parm = delta.list2, model = modelos)
datos = sim$dat
modelo1 = GDINA(datos, Q2, model = modelos,randomseed = 123456)
Int_1at[,p]=c(modelo1$delta.parm$`Item 4`[1],modelo1$delta.parm$`Item 6`[1],
modelo1$delta.parm$`Item 9`[1],modelo1$delta.parm$`Item11`[1],
modelo1$delta.parm$`Item 14`[1],modelo1$delta.parm$`Item 15`[1],
modelo1$delta.parm$`Item 16`[1],modelo1$delta.parm$`Item 17`[1],
modelo1$delta.parm$`Item 18`[1],modelo1$delta.parm$`Item 19`[1],
modelo1$delta.parm$`Item 20`[1],modelo1$delta.parm$`Item 21`[1],
modelo1$delta.parm$`Item 22`[1])
##ME: Main effect
Int_1at_ME[,p]=c(modelo1$delta.parm$`Item 4`[2],modelo1$delta.parm$`Item 6`[2],
modelo1$delta.parm$`Item 9`[2],modelo1$delta.parm$`Item11`[2],
modelo1$delta.parm$`Item 14`[2],modelo1$delta.parm$`Item 15`[2],
modelo1$delta.parm$`Item 16`[2],modelo1$delta.parm$`Item 17`[2],
modelo1$delta.parm$`Item 18`[2],modelo1$delta.parm$`Item 19`[2],
modelo1$delta.parm$`Item 20`[2],modelo1$delta.parm$`Item 21`[2],
modelo1$delta.parm$`Item 22`[2])
}
#Resultados de los escenarios
for(h in 1:13){
  for(cont in 1:nrep){
    resdelt0_1[h,cont] <- (Int_1at[h,cont] - gs[h,1])^2
    rmse_delt0_1[h] <- sqrt(sum(resdelt0_1[h,])/nrep)
    errordelt0_1[h] <- mean(Int_1at[h,] - gs[h,1])
    resdelt1_1[h,cont] <- (Int_1at_ME[h,cont] - gs[h,2])^2
    rmse_delt1_1[h] <- sqrt(sum(resdelt1_1[h,])/nrep)
    errordelt1_1[h] <-mean(Int_1at_ME[h,] -gs[h,2])
  }
}
write.csv(rmse_delt0_1, file=paste(dir,"/deltaintercepto",i,".csv",sep = ""))
write.csv(errordelt0_1, file=paste(dir,"/error_deltaint_1it",i,".csv",sep = ""))
write.csv(rmse_delt1_1, file=paste(dir,"/deltamainef",i,".csv",sep = ""))
write.csv(errordelt1_1, file=paste(dir,"/errormainef",i,".csv",sep = ""))
}

#####
###Simulación para ítems que requieren tres parámetros delta
#####
for (i in 1:length(Lista)) {
  ptm <-proc.time()
  N <- eval(parse(text=Lista[[i]][1]))

```

```

#Simulación
for (p in 1:nrep)
{sim <- simGDINA(N=N,Q2,delta.parm = delta.list2, model = modelos)
datos = sim$dat
modelo1 = GDINA(datos, Q2, model = modelos,randomseed = 123456)
Int_2at[,p]=c(modelo1$delta.parm$`Item 7`[1],modelo1$delta.parm$`Item 8`[1],
modelo1$delta.parm$`Item 10`[1])
Int_2at_ME[,p]=c(modelo1$delta.parm$`Item 7`[2],modelo1$delta.parm$`Item 8`[2],
modelo1$delta.parm$`Item 10`[2])
Int_2at_ME2[,p]=c(modelo1$delta.parm$`Item 7`[3],modelo1$delta.parm$`Item 8`[3],
modelo1$delta.parm$`Item 10`[3])
}
# Resultados de los escenarios
for(h in 1:4){
  for(cont in 1:nrep){
    resdelt0_2[h,cont] <- (Int_2at[h,cont] - bb[h,1])^2
    rmse_delt0_2[h] <- sqrt(sum(resdelt0_2[h,])/nrep)
    errordelt0_2[h] <-mean(Int_2at[h,] - bb[h,1])

    resdelt1_2[h,cont] <- (Int_2at_ME[h,cont] - bb[h,2])^2
    rmse_delt1_2[h] <- sqrt(sum(resdelt1_2[h,])/nrep)
    errordelt1_2[h] <-mean(Int_2at_ME[h,] - bb[h,2])

    resdelt2_2[h,cont] <- (Int_2at_ME2[h,cont] - bb[h,3])^2
    rmse_delt2_2[h] <- sqrt(sum(resdelt2_2[h,])/nrep)
    errordelt2_2[h] <-mean(Int_2at_ME2[h,] - bb[h,3])

  }
}
write.csv(rmse_delt0_2, file=paste(dir,"/deltainterc_2atrb",i,".csv",sep = ""))
write.csv(rmse_delt1_2, file=paste(dir,"/deltamainef_2atrb",i,".csv",sep = ""))
write.csv(rmse_delt2_2, file=paste(dir,"/deltamainef2_2atrb",i,".csv",sep = ""))
write.csv(errordelt0_2, file=paste(dir,"/error_deltaint_2at",i,".csv",sep = ""))
write.csv(errordelt1_2, file=paste(dir,"/error_deltaMainef_2at",i,".csv",sep = ""))
write.csv(errordelt2_2, file=paste(dir,"/error_deltaMainef2_2at",i,".csv",sep = ""))
}

```



## Bibliografía

- De la Torre, J. (2009). Dina model and parameter estimation: A didactic., *Journal of Educational and Behavioral Statistics* **34**(1): 115–130. <http://journals.sagepub.com/doi/abs/10.3102/1076998607309474>.
- De la Torre, J. (2011). The generalized dina model framework, *Psycometrika* **76**: 179–199. <https://link.springer.com/article/10.1007/s11336-011-9207-7>.
- De la Torre, J., Andries van der Ark, L. y Rossi, G. (2015). Analysis of clinical data from cognitive diagnosis modeling framework, *Measurement and Evaluation in Counseling and Development* pp.1–16. <https://doi.org/10.1177/0748175615569110>.
- De la Torre, J. y Young-Sun, L. (2013). Evaluating the wald test for item-level comparison of saturated and reduced models in cognitive diagnosis, *Journal of Educational Measurement* **50**: 355–373. <http://journals.sagepub.com/doi/abs/10.1177/0146621615621717?journalCode=apma>.
- Galeshi, R. (2012). *Cognitive Diagnostic Model, a Simulated-Based Study: Understanding Compensatory Reparameterized Unified Model (CRUM)*, PhD thesis, Faculty of the Virginia Polytechnic Institute and State University.
- George, A. (2014). *Investigating CDMs: Blending theory with practicality*.
- George, A., Robitzsch, A., Kiefer, T., Gro, J. y Ünlü (2016). The r package cdm for cognitive diagnosis models, *Journal of Statistical Software* **74**: 1–24. <http://journals.sagepub.com/doi/abs/10.1177/0146621615621717?journalCode=apma>.
- Harwell, M., Stone, C., Hsu, T. y Kirisci, L. (1996). Monte carlo studies in item response theory, *Applied Psychological Measurement* **20**: 101–125. <https://conservancy.umn.edu/bitstream/handle/11299/119086/1/v20n2p101.pdf>.
- Integración (2016). *INFORME 2016 - 6*. <http://www.integracion.pe/80-peruanos-problemas-salud-mental-no-atendidos/>.
- Minsa (2015). *Salud mental: Estadísticas, estrategias y desafíos relacionados con la etapa de vida joven*. [ftp://ftp2.minsa.gob.pe/descargas/mas/adolescente/DAIS/02\\_depresion\\_en\\_jovenes.pdf](ftp://ftp2.minsa.gob.pe/descargas/mas/adolescente/DAIS/02_depresion_en_jovenes.pdf).
- Ravand, H. y Robitzsch, A. (2015). Cognitive diagnostic modeling using r, *Practical Assessment, Research Evaluation* **20**(11): 1–12. <http://pareonline.net/getvn.asp?v=20&n=11>.
- Romero, A., Sánchez, F., Lozano, A., Ruíz, P. y Vega, J. (2016). Estructura de la sintomatología psiquiátrica según el srq-18 en pacientes ambulatorios con trastornos mentales evaluados en un hospital general de lima., *Revista de neuro-psiquiatría* **79**(3): 142–151. <http://journals.sagepub.com/doi/abs/10.1177/0146621615621717?journalCode=apma>.
- Sosa, Y. K. (2017). *Modelo dina aplicado a la evaluación de matemática en estudiantes de segundo grado de secundaria*, Master's thesis, Pontificia Universidad Católica del Perú.
- Templin, J. y Henson, R. (2006). Measurement of psychological disorders using cognitive diagnosis models, *Psychological Methods* **11**(3): 287–305. [https://www.researchgate.net/publication/6835883\\_Measurement\\_of\\_psychological\\_disorders\\_using\\_cognitive\\_diagnosis\\_models\\_Psychological\\_Methods\\_113\\_287-305](https://www.researchgate.net/publication/6835883_Measurement_of_psychological_disorders_using_cognitive_diagnosis_models_Psychological_Methods_113_287-305).

- Wenchao, M., De la Torre, J. y Sorrel, M. (2017). The generalized dina model framework. R package version 1.4.2 <https://cran.r-project.org/web/packages/GDINA/index.html>.
- Wenchao, M., Iaconangelo, C. y De la Torre, J. (2016). Model similarity, model selection and attribute classification, *Applied Psychological Measurement* **40**: 200–217. <http://journals.sagepub.com/doi/abs/10.1177/0146621615621717?journalCode=apma>.
- WHO(1994). *A user's guide to the Self Reporting Questionnaire*. [http://apps.who.int/iris/bitstream/handle/10665/61113/WHO\\_MNH\\_PSF\\_94.8.pdf?sequence=1&isAllowed=y](http://apps.who.int/iris/bitstream/handle/10665/61113/WHO_MNH_PSF_94.8.pdf?sequence=1&isAllowed=y).

