

PONTIFICIA UNIVERSIDAD  
CATÓLICA DEL PERÚ

Escuela de Posgrado



MODELOS DE TRANSICIÓN DE CLASES  
LATENTES

Tesis para optar el grado académico de Maestro en Estadística que  
presenta:

Ronald Eduardo Minchola Alza

Asesor:

Luis Hilmar Valdivieso Serrano

Lima, 2024

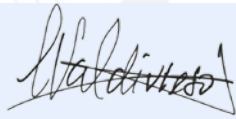
## Informe de Similitud

Yo, Luis Hilmar Valdivieso Serrano, docente de la Escuela de Posgrado de la Pontificia Universidad Católica del Perú, asesor(a) de la tesis titulada “Modelos de Transición de Clases Latentes”, de el autor Ronald Eduardo Minchola Alza, dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 10%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 15/08/2024.
- He revisado con detalle dicho reporte y la Tesis o Trabajo de investigación, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha:

Lima, 12 de Setiembre de 2024.

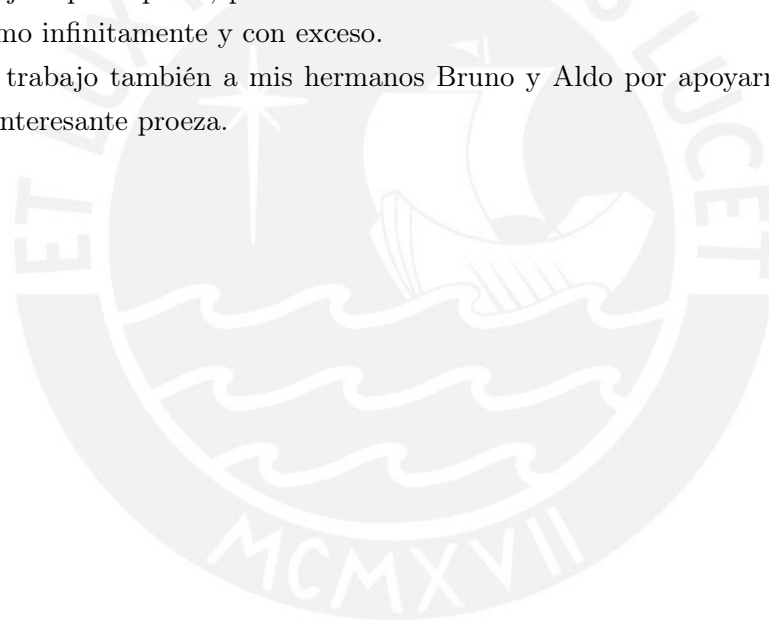
Apellidos y nombres del asesor: <u>Valdivieso Serrano, Luis Hilmar</u>	
DNI: 07958730	Firma
ORCID: <a href="https://orcid.org/0000-0002-8975-7557">https://orcid.org/0000-0002-8975-7557</a>	

# Dedicatoria

A Dios, a Hilda Mary Alza Silva y José Ronier Minchola Zavaleta con ese puro amor infinito que se tiene por los padres, por haber inculcado en mi muchas enseñanzas para ser buena persona y un excelente profesional. Por haberme dado lo mejor en todo aspecto, mil gracias siempre.

A Sinthia por ser esposa y compañera por su comprensión y paciencia, a mis hijos Sergio, Emilio y Valery por ser fuente de inspiración cada día y en cada paso que doy. Algún día leerán este trabajo espero que sí, pues infiero su curiosidad estadística siendo ellos todavía pequeños. Los amo infinitamente y con exceso.

Dedicar este trabajo también a mis hermanos Bruno y Aldo por apoyarme moralmente en esta difícil e interesante proeza.



# Agradecimientos

Quiero agradecer a todos mis profesores desde la profesora Dorita en mi primer grado de estudios en el colegio Manuel Pardo-Chiclayo, al buhonero que me enseñaba fracciones, MCM y MCD cuando todavía era un jovencito allá en la lejana década de los ochentas en la desaparecida tienda de Comercial Rominza, pasando por todos mis profesores de las academias y del CEP Claretiano-Trujillo recordando al gran Osiel Linares con sus tizas de colores y sus Diagramas de Venn, sería extenso detallar a todos los que han contribuido a mi formación pero de hecho los recuerdo muchísimo

Agradecer infinitamente a todos los profesores de la Maestría en Estadística muy en especial al Dr. Luis Valdivieso Serrano por su calidad de persona y ser humano, además de su sabiduría y paciencia hacia mi persona.

Agradecer también con mucho corazón a mi tío Efraim y mi dulce tía Josefina Chepita por haberme apoyado, cobijándome del frío limeño mil gracias me han dado mucho cariño y enseñanzas diariamente.

A Manuel Merino y Walter Chanavá por la gran ayuda brindada en varios aspectos del presente trabajo a pesar del momento vivido (pandemia) me inculcaron a seguir adelante deberás que Dios los puso en mi camino.

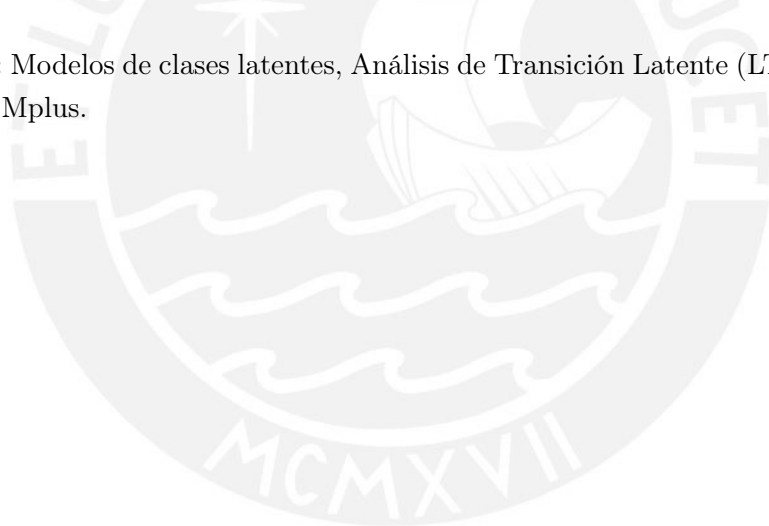
Con mucho cariño un agradecimiento sincero a los colegas del Departamento Académico de Estadística de la Universidad Nacional de Piura (UNP) Walter Alva, Carlos Gonzales, Lemin Abanto, José Fiestas, Cosme Correa quienes cubrieron mi carga académica durante mis estudios de la Maestría en Estadística de la PUCP.

Fuerza Carlos A. Mannucci de la ciudad de Trujillo!!!!!!!

# Resumen

A pesar de que los modelos de clases latentes han recibido gran atención en la evaluación social y otros campos, estos modelos son estáticos o de corte transversal y hay poco trabajo investigativo en modelos longitudinales. Este trabajo estudia el análisis de transición de clases latentes (LTA), observando el cambio en la clasificación de clases a través del tiempo. Se desarrolla el modelo teórico para dos períodos y se extiende luego este para más de dos períodos, estudiándose sus parámetros y estimación a través del algoritmo de Esperanza-Maximización (EM). Se detallan también criterios para la selección de modelos definiendo teóricamente la razón de su uso. Finalmente, se realiza un estudio de aplicación en Mplus y en R sobre una base de datos de acoso escolar(bullying).

**Palabras clave:** Modelos de clases latentes, Análisis de Transición Latente (LTA), Algoritmo EM, Invarianza, Mplus.



# Abstract

Although latent class models have received great attention in social assessment and other fields, these models are static or cross-sectional and there is little research work on longitudinal models. This work studies latent class transition analysis (LTA), observing the change in class classification over time. The theoretical model is developed for two periods and then extended for more than two periods, studying its parameters and estimation through the Expectation-Maximization (EM) algorithm. Criteria for selecting models are also detailed, theoretically defining the reason for their use. Finally, an application study is carried out in Mplus and R on a bullying database.



# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Consideraciones preliminares y justificación . . . . .	1
1.2. Objetivos . . . . .	2
1.3. Organización del Trabajo . . . . .	3
<b>2. Modelos de clases latentes</b>	<b>4</b>
2.1. Modelos de clases latentes . . . . .	4
2.2. Cadenas de Markov . . . . .	6
<b>3. Modelos de transición de clases latentes</b>	<b>10</b>
3.1. Introducción . . . . .	10
3.1.1. La prevalencia de los estados latentes . . . . .	13
3.1.2. Las probabilidades de transición . . . . .	13
3.1.3. Las probabilidades de respuesta al ítem . . . . .	14
3.1.4. Restricción de parámetros en el análisis de transición de clases latentes	14
<b>4. Estimación del modelo</b>	<b>17</b>
4.1. Función de verosimilitud . . . . .	18
4.2. Algoritmo EM . . . . .	19
4.3. Criterios para la selección del modelo . . . . .	24
<b>5. Aplicación a un estudio de acoso escolar</b>	<b>27</b>
5.1. Metodología del recojo de la información . . . . .	29
5.2. Metodología de análisis . . . . .	29
5.3. Preocupación por el acoso escolar . . . . .	29
5.4. Resultados . . . . .	30
5.5. Acoso escolar y su relación con otras variables . . . . .	37
<b>6. Conclusiones y sugerencias</b>	<b>43</b>
<b>Bibliografía</b>	<b>45</b>
<b>A. Anexos</b>	<b>48</b>
A.1. Código Mplus para la estimación del modelo LTA . . . . .	48
A.2. Código R para biplots del análisis de correspondencias múltiple . . . . .	60

# Lista de Abreviaturas

EM = Algoritmo de Esperanza Maximización

LCA = Análisis de clases latentes

LTA = Análisis de transición de clases latentes

SAS = Statistical analysis system

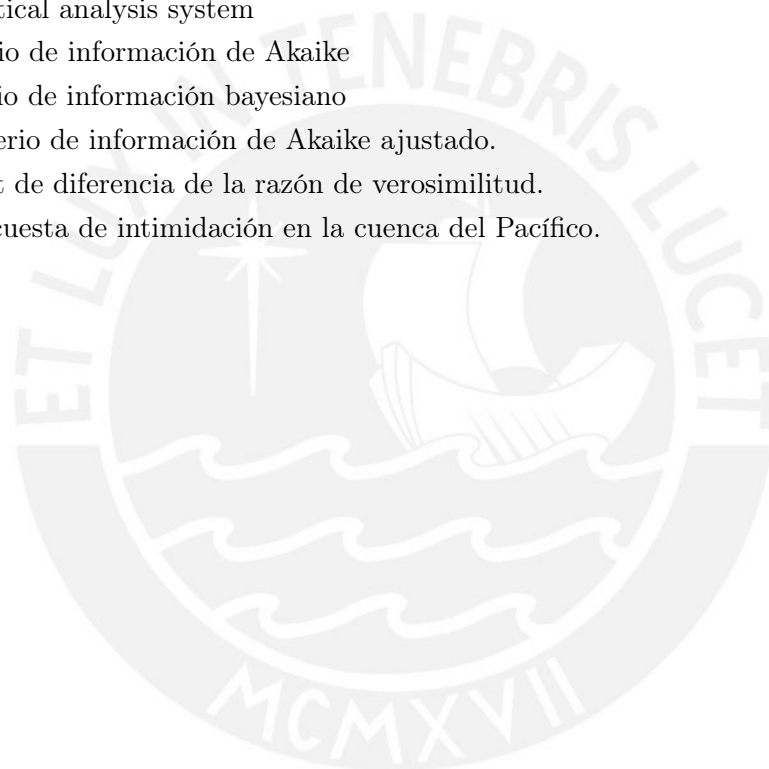
AIC = Criterio de información de Akaike

BIC = Criterio de información bayesiano

CAIC = Criterio de información de Akaike ajustado.

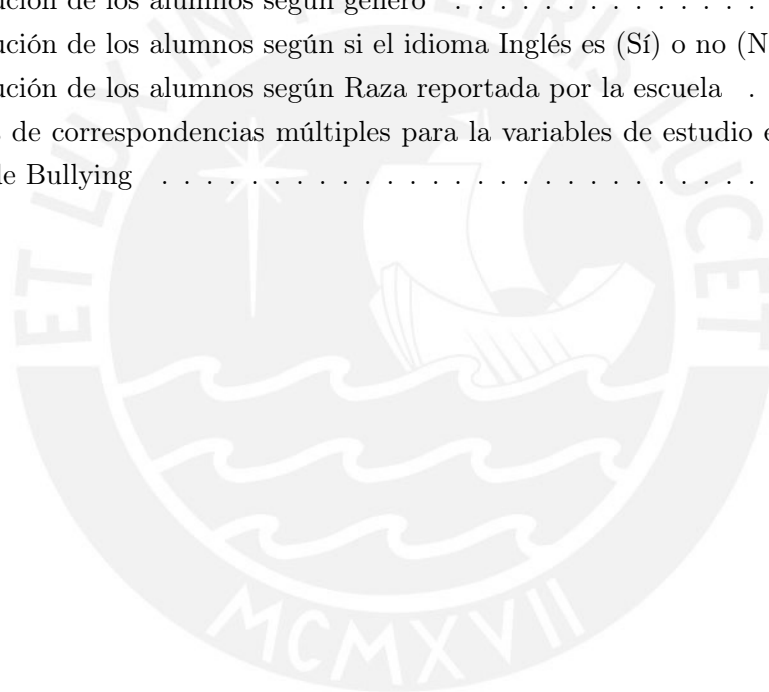
LRDT = Test de diferencia de la razón de verosimilitud.

PRBm = Encuesta de intimidación en la cuenca del Pacífico.



# Índice de figuras

5.1. Porcentajes de Bullying según tiempo y respuesta . . . . .	30
5.2. Perfiles latentes para el modelo de 3 soluciones . . . . .	34
5.3. Perfiles latentes para el modelo de 4 soluciones . . . . .	34
5.4. Distribución de alumnos en los diferentes grados para el estudio de Bullying .	37
5.5. Distribución de los alumnos según género . . . . .	38
5.6. Distribución de los alumnos según si el idioma Inglés es (Sí) o no (No) nativo	38
5.7. Distribución de los alumnos según Raza reportada por la escuela . . . . .	39
5.8. Análisis de correspondencias múltiples para la variables de estudio en el problema de Bullying . . . . .	42



# Índice de tablas

5.1. Respuestas marginales para los 6 ítems en la encuesta sobre el acoso escolar PRBm . . . . .	30
5.2. Resultados del análisis de clases latentes en cada periodo de tiempo. . . . .	31
5.3. Resultados de la medida de invarianza longitudinal. . . . .	32
5.4. Estimaciones de las prevalencias en los estados latentes en cada punto en el tiempo . . . . .	35
5.5. Estimación de la matriz de transición del tiempo 1 al tiempo 2 . . . . .	36
5.6. Estimación de la matriz de transición del tiempo 2 al tiempo 3 . . . . .	36
5.7. Resultados de la invarianza en las probabilidades de transición . . . . .	37
5.8. Distribución porcentual de las clases latentes según periodos de estudio . . . .	39
5.9. Prueba de asociación Chi cuadrado entre la clasificación latente y el grado de estudios . . . . .	40
5.10. Prueba de asociación Chi cuadrado entre la clasificación latente y el género .	40
5.11. Prueba de asociación Chi cuadrado entre la clasificación latente y el idioma .	40
5.12. Prueba de asociación Chi cuadrado entre la clasificación latente y la raza . . .	41

# Capítulo 1

## Introducción

### 1.1. Consideraciones preliminares y justificación

El estudio de datos longitudinales desempeña una función crucial en el progreso de la ciencia. En estos estudios el investigador busca estudiar factores de riesgo para encontrar oportunidades óptimas de intervención. Dentro de estos métodos destaca el análisis de transición de clases latentes (LTA). Este es un modelo latente de cadenas de Markov, donde el estado de membresía en cada punto del tiempo es no observable, pero medible a partir de un conjunto de variables (o ítems) manifiestas (Collins & Lanza, 2010). En el análisis de transición de clases latentes (LTA), el modelo de medición en cada punto del tiempo es un modelo de clases latentes, donde la relación entre las variables manifiestas son explicadas por una variable latente categórica subyacente y el desarrollo longitudinal de esta es resumida en sus probabilidades de transición. Usualmente cuando queremos analizar datos longitudinales las preguntas de investigación están en relación al cambio a través del tiempo. El análisis de transición de clases latentes (LTA) es una aplicación longitudinal del modelo del análisis de clases latentes, que tiene por objetivo identificar si existe un cambio entre las clases latentes a través del tiempo (Lanza et al., 2003). Mientras que el análisis de clases latentes es aplicado a un conjunto de variables recolectadas en un tiempo específico e identifica las clases latentes en base a los patrones de respuesta, el análisis de transición de clases latentes es aplicado a mediciones repetidas en el tiempo para estudiar el movimiento entre las clases latentes identificadas.

El LTA fue desarrollado inicialmente para estudiar el cambio secuencial por etapas de un tipo de variables latentes llamadas variables latentes dinámicas (Collins et al., 1993). Las variables latentes dinámicas incluyen características tales como actitudes y patrones de personalidad que cambian a través del tiempo. La técnica fue posteriormente utilizada para el estudio sobre las intervenciones como, por ejemplo, el abandono del consumo de tabaco, la disminución de malas conductas (comportamiento inapropiado) y el crecimiento académico (Boscardin et al., 2008); (Compton et al., 2008); (Trentacosta et al., 2011) entre otros. Las aplicaciones respecto al LTA aparecen desde principios de los 90s con publicaciones en varias áreas del conocimiento como las ciencias sociales, psicología, educación y salud pública.

En esta investigación, detallaremos el modelo LTA y describiremos su uso con datos concernientes al abuso escolar a partir de una muestra de 1180 estudiantes del quinto al noveno

grado durante tres semestres académicos. Esta aplicación está basada en el artículo de Ryoo et al. (2018).

El LTA presenta varias razones para su uso: como representar variables latentes multidimensionales, modelar y predecir el cambio a través del tiempo, el cual es en cierto sentido discreto y permitir conocer si algunos estados latentes pueden tener prevalencias muy bajas en etapas iniciales, pero a medida que los individuos hacen la transición en el tiempo, su prevalencia aumenta (Lanza et al., 2003).

Este modelo, permitirá entonces a un investigador responder directamente a un conjunto de preguntas como: ¿existe algún cambio entre las clases latentes a través del tiempo? y, si la respuesta es positiva, explicar ¿cómo caracterizar dicho cambio a través de las probabilidades?

Así, como en el análisis de clases latentes, en este nuevo modelo se estiman también las probabilidades de respuesta al ítem, pero además la prevalencia en las clases latentes que se entiende como el número de casos de un evento en una población en un momento dado y la incidencia de las transiciones entre las clases latentes junto con una medida del error.

Tres grupos de parámetros son estimados en un modelo LTA. El propósito de este estudio es examinar exhaustivamente el modelo LTA y la implementación de la estimación de sus parámetros desde un punto de vista clásico. Se ilustrará ello con una aplicación basada en datos reales y se usará para esta implementación el software estadístico Mplus.

## 1.2. Objetivos

El objetivo principal de la investigación es estudiar el modelo de transición de clases latentes para datos longitudinales, así como su implementación y aplicación. De manera específica:

- Revisar la literatura acerca de los modelos de clases latentes y de transición de clases latentes.
- Presentar los fundamentos y propiedades de los modelos de análisis de clases latentes(LCA), los modelos de transición de clases latentes(LTA) y su implementación computacional.
- Estudiar el proceso de estimación de parámetros de los modelos de transición de clases latentes bajo el enfoque clásico o frecuentista y el uso del paquete Mplus.
- Realizar un estudio de aplicación a efectos de analizar la performance del proceso de estimación del modelo LTA.

### 1.3. Organización del Trabajo

En el Capítulo 2 se estudian los modelos de clases latentes y las cadenas de Markov en tiempo discreto, que son las bases para entender mejor el modelo de transición de clases latentes. En el Capítulo 3 se detalla el modelo de transición de clases latentes y los parámetros involucrados. En el capítulo 4 se estudia el proceso de estimación de los parámetros del modelo vía el algoritmo EM y su implementación en el software Mplus. En el Capítulo 5 se presenta una aplicación sobre el acoso escolar (bullying). En el Capítulo 6 se presentan algunas conclusiones, recomendaciones y sugerencias para futuras investigaciones que se podrían derivar de este trabajo. Se incluirá finalmente en Anexos los códigos en Mplus y R-studio para la estimación del modelo en estudio.



## Capítulo 2

# Modelos de clases latentes

En diversas disciplinas como las ciencias sociales, ciencias de la salud y especialmente en las ciencias del comportamiento, existe un gran interés en aproximarse a constructos teóricos que, aunque no pueden ser observados directamente, se considera que influyen en variables que sí pueden medirse. Esto nos permite acercarnos a estos constructos y obtener un mejor entendimiento de las variables observadas.

Estos constructos son frecuentemente denominados variables latentes en la literatura estadística, y su modelo estadístico describe la distribución conjunta de las variables que pueden observarse, también conocidas como variables manifiestas, y las variables latentes.

### 2.1. Modelos de clases latentes

El modelo de clases latentes constituye un tipo específico dentro de los modelos de variables latentes, en el cual tanto las variables observables (o manifiestas) como las no observables (o latentes) son categóricas. La finalidad es agrupar a los individuos según las variables que pueden observarse. La naturaleza categórica de la variable latente se establece a partir de una base teórica o evidencia previa, o simplemente por razones prácticas (Lazarsfeld & Henry, 1968).

Las ventajas del análisis de clases latentes son diversas: simplifica la complejidad de conjuntos de datos al agruparlos en clases, permite la estimación de varias probabilidades y facilita el análisis de datos categóricos sin necesidad de transformaciones adicionales.

El modelo estándar parte de dos supuestos: dentro de cada clase, todos los individuos tienen las mismas probabilidades de responder a las variables observables, y se cumple la

independencia condicional, lo que significa que las respuestas entre individuos son independientes dado que pertenecen a la misma clase.

Para ilustrar, se examinará un modelo de clases latentes con variables observables binarias, donde la variable latente  $L$  es unidimensional y se organiza en  $C$  categorías o clases, denotadas por un índice específico  $\ell = 1, 2, \dots, C$ . Sea  $\pi_{\ell j}$  la probabilidad de obtener una respuesta positiva a la variable manifiesta  $X_j$ , donde  $j$  es la  $j$ -ésima respuesta al ítem, dado que el individuo pertenece a la clase  $\ell$ ; es decir,  $\pi_{\ell j} = P(X_j = 1|L = \ell)$  y sea  $\eta_\ell = P(L = \ell)$ , la probabilidad a priori de que un individuo sea parte de la clase  $\ell$ . Entonces, la función de probabilidad conjunta del vector de respuestas observadas o patrones de respuesta  $\mathbf{X} = (X_1, X_2, \dots, X_p)$  de un individuo viene dada por

$$f(\mathbf{x}) = P(\mathbf{X} = \mathbf{x}) = \sum_{\ell=1}^C P(\mathbf{X} = \mathbf{x}|L = \ell)P(L = \ell) = \sum_{\ell=1}^C g(\mathbf{x}|\ell)\eta_\ell, \quad (2.1)$$

donde:

$$g(\mathbf{x}|\ell) = P(\mathbf{X} = \mathbf{x}|L = \ell) = \prod_{j=1}^p \pi_{\ell j}^{x_j} (1 - \pi_{\ell j})^{1-x_j} \quad (2.2)$$

y  $x_j$  vale 1 (si la respuesta al ítem  $j$  es correcta o positiva) o 0 (si es incorrecta o negativa).

De esta manera, cada individuo es descrito por dos dimensiones: el vector de variables manifiestas  $\mathbf{X}$  y el indicador de pertenencia a la clase. La probabilidad a posteriori de que un individuo con un patrón de respuestas  $\mathbf{x} = (x_1, x_2, \dots, x_p)$  pertenezca a la clase  $\ell$  es por tanto:

$$h(\ell|\mathbf{x}) = P(L = \ell|\mathbf{X} = \mathbf{x}) = \frac{P(\mathbf{X} = \mathbf{x}|L = \ell)P(L = \ell)}{P(\mathbf{X} = \mathbf{x})} = \frac{\eta_\ell \prod_{j=1}^p \pi_{\ell j}^{x_j} (1 - \pi_{\ell j})^{1-x_j}}{\sum_{\ell=1}^C \eta_\ell \prod_{j=1}^p \pi_{\ell j}^{x_j} (1 - \pi_{\ell j})^{1-x_j}}. \quad (2.3)$$

Esta función se emplea posteriormente como criterio para asignar a un individuo a la clase que tiene la probabilidad más alta de que pertenezca, dado el patrón de respuestas que presenta. El modelo necesita de los parámetros  $\eta_\ell$  y  $\pi_{\ell j}$ , además, se debe calcular la adecuación del modelo. También es crucial identificar y interpretar las clases latentes subyacentes de manera que se relacionen significativamente con los datos.

## 2.2. Cadenas de Markov

Extenderemos ahora el modelo de clases latentes a datos de caracter longitudinal para averiguar, entre otras, cosas como cambian con los datos la pertenencia de un individuo a las clases. Para ello requeriremos del concepto de cadenas de Markov que pasamos a revisar brevemente (Parzen, 1979).

Consideremos un conjunto  $E$  finito o numerable. Sea  $X_0, X_1, X_2, \dots$  una sucesión de variables aleatorias discretas las cuales toman valores en  $E$  y están definidas sobre cierto espacio probabilístico. El conjunto  $E$  se denominará espacio de estados, y sus elementos estados. Las variables aleatorias  $X_0, X_1, \dots, X_n$  se dicen independientes si y sólo si, se cumple que:

$$P(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \prod_{k=1}^n P(X_k = x_k), \quad (2.4)$$

para todo estado  $x_0, x_1, \dots, x_n$ . En este caso los eventos  $A = \{X_0 = x_0, \dots, X_n = x_n\}$  y  $B = \{X_{n+1} = x_{n+1}, \dots, X_{n+m} = x_{n+m}\}$  son independientes para cualesquier  $n = 0, 1, \dots$  y  $m = 1, 2, \dots$  y cualquier sucesión de estados  $x_0, \dots, x_{n+m}$ . La dependencia markoviana consiste en que la probabilidad del evento B depende sólo del valor que toma la variable aleatoria  $X_n$  y no de los valores que toman las variables aleatorias  $X_0, \dots, X_{n-1}$ . Si el índice de la sucesión representa el tiempo discreto y  $n$  es el instante presente en una cadena de Markov, podemos decir que la probabilidad de que un suceso ocurra en los instantes futuros  $n+1, \dots, n+m$ , depende solamente del estado en que se encuentra la sucesión en el instante presente  $n$  y no de los estados en que se encontró en los instantes pasados  $0, 1, \dots, n-1$ . Formalmente, diremos que la sucesión  $\{X_n\}$  es una cadena de Markov con espacio de estados discreto si:

$$P(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0) = P(X_{n+1} = x_{n+1} | X_n = x_n), \quad (2.5)$$

para todo  $n = 1, 2, \dots$  y cualquier sucesión de estados  $x_0, \dots, x_{n+1}$  en  $E$ , siempre que  $P(X_n = x_n, \dots, X_0 = x_0) > 0$ . A la identidad (2.5) se le llama la propiedad markoviana. En adelante, para no recargar notaciones, asumiremos, sin pérdida de generalidad, que el espacio de estados  $E$  es un subconjunto de  $\mathbb{N}$ . En tal sentido, diremos que una cadena de Markov es homogénea en el tiempo si para todo par de estados  $i, j$ , la probabilidad condicional  $P(X_{n+1} = j | X_n = i)$  no depende de  $n$ ; es decir,

$$P(X_1 = j | X_0 = i) = P(X_2 = j | X_1 = i) = \dots = P(X_{n+1} = j | X_n = i). \quad (2.6)$$

En general, utilizamos en adelante la expresión cadena de Márkov para referirnos a una cadena de Markov homogénea en el tiempo. La caracterización de la cadena en este caso se definirá a través de las probabilidades de transición

$$p_{ij} = P(X_1 = j | X_0 = i) \quad (2.7)$$

y de las probabilidades iniciales

$$\pi_i^0 = P(X_0 = i)$$

A la matriz  $P = [p_{ij}]$ , con  $i \in E, j \in E$  (posiblemente infinita) se le denomina la matriz de transición, y al vector  $\pi^0 = (\pi_i^0)$ , la distribución inicial de la cadena de Markov. Es sencillo, ver que la matriz de transición satisface las siguientes propiedades: (Parzen, 1979)

- i)  $p_{ij} \geq 0$ , para todo par de estados  $i, j$  en  $E$ .
- ii)  $\sum_{j \in E} p_{ij} = 1$  para todo estado  $i$  en  $E$ .

Por su parte, la distribución inicial satisface las propiedades:

- i)  $\pi_i^0 \geq 0$ , para todo estado  $i$  en  $E$ .
- ii)  $\sum_{i \in E} \pi_i^0 = 1$ .

La definición de una cadena de Markov abarca varios casos, incluyendo secuencias de variables aleatorias independientes e idénticamente distribuidas, así como secuencias formadas por sumas parciales de variables aleatorias también independientes e idénticamente distribuidas, cuando estas toman valores enteros.

Ahora consideremos las probabilidades de transición de orden  $n$  de una cadena de Markov.  $\{X_n\}$  con espacios de estados  $E$ , matriz de transición  $P$  y distribución inicial  $\pi^0$ . Dados dos estados  $i \in E, j \in E$ , estas vienen dadas, respectivamente, por

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i) \quad \text{y} \quad \pi_i^{(n)} = P(X_n = i), \quad (2.8)$$

a la matriz  $P^{(n)} = [p_{ij}^{(n)}]$ , se le llama la matriz de transición de orden  $n$  y al vector  $\pi^{(n)} = (\pi_i^{(n)})$ , con  $i \in E$ , la distribución de probabilidad de la cadena de Markov en el periodo  $n$ . Observemos que  $\pi^{(0)}$  es la distribución inicial de la cadena de Markov y  $P^{(1)} = P$  su matriz de transición. Las probabilidades de transición de orden  $n$  satisfacen la ecuación de Chapman-Kolgomorov:

$$p_{ij}^{(m+n)} = \sum_{k \in E} p_{ik}^{(m)} p_{kj}^{(n)}, \quad (2.9)$$

para todo par de índices  $m, n$  y todo par de estados  $i, j$ . Esto es, en notación matricial, (2.9) tiene la forma

$$P^{(m+n)} = P^{(m)} \times P^{(n)}, \quad (2.10)$$

donde  $\times$  denota el producto de matrices. En efecto, aplicando la fórmula de la probabilidad total y la propiedad de Markov (2.5) se obtiene:

$$\begin{aligned} p_{ij}^{(m+n)} &= P(X_{m+n} = j | X_0 = i) = \sum_{k \in E} P(X_{m+n} = j, X_m = k | X_0 = i) = \\ &= \sum_{k \in E} P(X_{m+n} = j | X_m = k) P(X_m = k | X_0 = i) = \sum_{k \in E} p_{ik}^{(m)} p_{kj}^{(n)}. \end{aligned} \quad (2.11)$$

Análogamente, se prueba que para la distribución de probabilidad en el instante  $n$  tiene lugar la identidad

$$\pi_j^{(m+n)} = \sum_{k \in E} \pi_k^{(m)} p_{kj}^{(n)}, \quad (2.12)$$

para todo par de índices  $m, n$  y todo estado  $j$ . En notación matricial esto se escribe como:

$$\pi^{(m+n)} = \pi^{(m)} \times P^{(n)}$$

En particular, de (2.10) resulta que  $P^{(n)} = P \times P^{(n-1)}$  y  $P^{(n)} = P^n$ . En conclusión, la matriz de transición de orden  $n$  es la potencia  $n$ -ésima de la matriz de transición  $P$ , y es correcto interpretar el superíndice  $n$  en la notación  $P^n$  como la potencia  $n$ -ésima de la matriz  $P$ . La distribución de probabilidad en el instante  $n$  se obtiene mediante la fórmula

$$\pi^n = \pi^0 \times P^n, \quad (2.13)$$

que también se escribe en su  $j$ -ésima componente como:

$$\pi_j^{(n)} = \sum_{k \in E} \pi_k^0 p_{kj}^{(n)}, \quad (2.14)$$

Calculemos ahora, para una elección de índices  $n_1, \dots, n_k \in \mathbb{N}$  arbitraria, la distribución del vector aleatorio  $(X_{n_1}, \dots, X_{n_k})$ , que llamamos la distribución finito-dimensional de la cadena de Markov. Es claro que para este cálculo es suficiente conocer probabilidades de la forma

$$P(X_0 \in A_0, \dots, X_n \in A_n), \quad n = 0, 1, \dots$$

donde  $A_0, \dots, A_n$  son subconjuntos arbitrarios de  $E$ . A su vez, para calcular estas últimas

probabilidades, es suficiente, dada una sucesión de estados  $i_0, \dots, i_n$ , conocer las probabilidades

$$\begin{aligned}
 P(X_n = i_n, \dots, X_0 = i_0) &= \\
 &= P(X_n = i_n | X_{n-1} = i_{n-1}) \dots P(X_1 = i_1 | X_0 = i_0) P(X_0 = i_0) \\
 &= p_{i_{n-1}i_n} \dots p_{i_0i_1} \pi_{i_0}.
 \end{aligned} \tag{2.15}$$

Este último cálculo muestra que las distribuciones finito dimensionales de una cadena de Markov pueden determinarse si se conoce la matriz de transición  $P$  y su distribución inicial  $\pi^0$ . Más aún, es posible demostrar que la probabilidad de un evento que depende de una cantidad arbitraria, no necesariamente finita, de variables aleatorias de la cadena de Markov, se puede hallar a partir de  $P$  y de  $\pi^0$  y que, reciprocamente, dados un conjunto  $E$ , finito o numerable, una matriz  $P = [p_{ij}]$ , con  $i \in E$  y  $j \in E$  que cumple las propiedades i) y ii) y un vector  $\pi = (\pi_i)_{i \in E}$  que satisface las propiedades i) y ii), existe un espacio de probabilidades y una cadena de Markov homogénea con espacio de estados  $E$  y variables aleatorias definidas en este espacio de probabilidades, que tiene a  $P$  como matriz de transición y a  $\pi^0$  como distribución inicial. De aquí se obtiene la siguiente conclusión: Las propiedades probabilísticas de una cadena de Markov, dependen únicamente de la matriz transición y de la distribución inicial de la cadena de Markov. (Parzen, 1979)

## Capítulo 3

# Modelos de transición de clases latentes

### 3.1. Introducción

El modelo de transición de clases latentes (LTA), llamado también de mezcla de clases latentes de Markov, es un modelo de variable latente similar al de clases latentes (LCA). Sin embargo, mientras que el LCA usa datos de corte transversal, el LTA es usado con datos tipo panel o longitudinal para investigar si algún cambio ha ocurrido dentro de las clases latentes a través del tiempo. Cuando se dispone de datos longitudinales el modelo de transición de clases latentes permite direccionar un conjunto de preguntas: ¿Existe algún cambio entre las clases latentes a través del tiempo? Si es así ¿cómo se puede caracterizar este cambio? Si un individuo se encuentra en una clase latente particular en el tiempo  $t$ , ¿cuál es la probabilidad de que este individuo permanezca en la misma clase latente en el tiempo  $t + 1$ , y cual la probabilidad de que el individuo esté en una clase latente diferente? El LTA es una forma de ajustar modelos que abordan este conjunto de preguntas adicionalmente al conjunto de preguntas abordadas. Es por eso que esta investigación propone, no solo estudiar las clases latentes de algunos fenómenos sociales, sino también observar como los individuos transitan por las diferentes clases latentes a través del tiempo. El LTA, al igual que el LCA, estima probabilidades de respuesta al ítem. Por lo tanto, se estiman también la prevalencia de las clases latentes y la incidencia de las transiciones entre las clases latentes mientras se ajusta un error de medición. El modelo de transición de clases latentes con solo dos puntos en el tiempo se plantea como:

$$P(\mathbf{Y} = \mathbf{y}) = \sum_{\ell_1=1}^C \sum_{\ell_2=1}^C \delta_{\ell_1} \tau_{\ell_2|\ell_1} \prod_{t=1}^2 \prod_{m=1}^M \prod_{k=1}^{r_m} \rho_{mkt|\ell_t}^{I(y_{mt}=r_{m,t})}, \quad (3.1)$$

donde:

$$\sum_{\ell_1=1}^C \delta_{\ell_1} = 1, \sum_{\ell_2=1}^C \tau_{\ell_2|\ell_1} = 1, \sum_{k=1}^{r_m} \rho_{mkt|\ell_t} = 1, \quad (3.2)$$

siendo  $\mathbf{y} = [y_{m,t}]$  la matriz de respuestas de un individuo al ítem  $m$  en el tiempo  $t$ ;  $C$  el número de clases latentes;  $M$  el número de ítems;  $\delta_{\ell_1}$  la probabilidad de pertenencia del individuo a la clase latente  $\ell_1$  en el tiempo 1;  $\tau_{\ell_2|\ell_1}$  la probabilidad de transición de que el individuo pase de la clase latente  $\ell_1$  en el tiempo 1 a una clase  $\ell_2$  en el tiempo 2;  $r_m$  la categoría de respuesta al ítem  $m$ ;  $I(y_{m,t} = k)$  una variable indicadora que es igual a 1 si la respuesta al ítem  $m$  en el tiempo  $t$  es  $k$  y 0 en caso contrario, y  $\rho_{mkt|\ell_t}$  la probabilidad de que su respuesta al ítem  $m$  en el tiempo  $t$  sea  $k$  condicionado a la membresía a la clase latente  $\ell_t$  en ese instante. Se asume que cada ítem  $m$  tiene  $r_m$  posibles categorías de respuestas. La consideración de un modelo con más de dos puntos en el tiempo es inmediata, pero este implica el incremento en el número de parámetros, no sólo en términos de las probabilidades de transición sino también de los parámetros dentro de las probabilidades  $\rho_{mkt|\ell_t}$  de respuesta al ítem del modelo.

Supongamos como ilustración que se tienen  $M = 6$  variables observables binarias, que han sido medidas en los tiempos  $t = 1, 2$ . Las variables observadas podrían ser 6 ítems de un cuestionario acerca de delincuencia, donde cada variable observada  $m$  tiene 2 categorías de respuesta (sí o no) en cada tiempo. La tabla de contingencia cruzando las 6 variables en los dos tiempos tiene  $2^{2 \times 6} = 4,096$  celdas o patrones de respuesta. Los patrones de respuesta en estas, son representados como  $(r_{1,1}, \dots, r_{6,2})$ . Por ejemplo, un patrón de respuesta podría ser (sí,no,no,no,no,sí,sí,no,no,no) que representa la respuesta de sí para el primer ítem del cuestionario en el tiempo 1 hasta una respuesta de no para el sexto ítem en el tiempo 2. En adelante denotaremos por  $L_t$  a la variable latente categórica que nos indica a que clase latente pertenece el individuo en el tiempo  $t$ .

Las prevalencias en las clases latentes son estimadas para cada punto en el tiempo. Así, por ejemplo, en un modelo con dos puntos en el tiempo y cuatro clases latentes, ocho prevalencias requerirán ser estimadas, mostrando ellas la proporción de personas en cada uno de los estados latentes. Por lo tanto, observando la prevalencia en estos estados podemos ver el aumento o disminución en las prevalencias entre los puntos en el tiempo considerados en el estudio, asimismo observar patrones de cambio y como las personas se mueven entre los estados.

Las probabilidades de respuesta al ítem representan las probabilidades de que las personas den respuestas correctas a cada ítem sujeto a su clase de pertenencia. El modelo de transición de clases latentes es particularmente usado más como un método exploratorio, en donde el etiquetado o codificación de los estados latentes se realiza evaluando las probabilidades de respuesta al ítem. Existe una probabilidad de respuesta al ítem para cada combinación ítem-estado. Por lo tanto, el número de celdas en una tabla de probabilidades de respuesta al ítem es el número de ítems multiplicado por el número de estados latentes.

Las probabilidades de transición son de relevante importancia en el modelo porque permiten identificar si ocurrió algún cambio entre los estados latentes a través del tiempo. Para un modelo con dos puntos en el tiempo ( $T = 2$ ), las probabilidades de transición permiten la creación de una tabla de clasificación cruzada con el número de estados latentes en el tiempo o periodo 1 y el número de estados latentes en el tiempo o periodo 2. En general, dado cierto número de puntos en el tiempo  $T$ , el modelo de transición de clases latentes estima  $T - 1$  matrices de probabilidades de transición. Cada examinado, se asume que puede ser miembro de uno y solo un estado latente en cada punto en el tiempo.

El modelo de transición de clases latentes asume la no existencia de datos faltantes en las variables indicadoras observadas. Sin embargo, los datos faltantes en el LTA son sobrellevados de la misma forma que en el modelo de clases latentes. Los datos pueden usarse a partir de las respuestas que dan los individuos a algunas de las preguntas en un tiempo en particular y para aquellos que solo están presentes en un subconjunto en el tiempo de medición. Se hace entonces la clásica suposición de datos faltantes de missing at random.

Los parámetros en los modelos de transición latente se pueden estimar por Máxima verosimilitud (ML) o por métodos bayesianos vía Cadenas de Markov de Montecarlo (MCMC). En este trabajo estudiaremos la estimación por ML a través del algoritmo de Esperanza Maximización (EM). Dada la gran cantidad de parámetros involucrados en el modelo, este es mucho más rápido que los métodos bayesianos; sin embargo, aquí la función de verosimilitud podría tener características inusuales que pueden causar dificultad al momento de hacer inferencias en la estimación, especialmente en muestras pequeñas. Por ejemplo, puede haber ciertas regiones continuas del espacio de parámetros donde la log-verosimilitud sea constante, lo que conduce a la indeterminación de algunos parámetros. En tales casos, sería conveniente optar por métodos bayesianos.

A las clases latentes del modelo de transición las llamaremos estados latentes teniendo en cuenta que las clases latentes son estados temporales, y que los individuos pueden moverse hacia dentro y hacia fuera de estos estados latentes.

La combinación de grandes grados de libertad y la extrema dispersión no permiten probar hipótesis tradicionales acerca del ajuste absoluto del modelo de transición de clases latentes. Esto se debe a que la distribución de la estadística  $G^2$  (radio de verosimilitud) no es muy bien aproximada por una distribución Chi-cuadrado y por consiguiente los valores  $p$  son bastante inexactos. Es por eso, que se prefiere enmarcar la selección del modelo en términos relativos donde sea posible, es decir, ajustar una serie de modelos y confiar en los AIC y BIC para tomar decisiones acerca del modelo que mejor represente a nuestros datos. Las pruebas de hipótesis son usadas cuando se desea comparar modelos de transición anidados. También la parsimonia y la conceptualización teórica son importantes criterios de selección en un modelo LTA, tal como lo es en un modelo de clase latente (Collins & Lanza, 2010).

Como fue mencionado, en el modelo LTA interesan estimar 3 conjuntos diferentes de parámetros que pasamos a detallar: las probabilidades de membresía, las probabilidades de transición y las probabilidades de respuesta al ítem.

### 3.1.1. La prevalencia de los estados latentes

La prevalencia del estado latente  $\ell$  en el tiempo  $t$  se denota por  $\delta_{\ell_t}$ , y viene dada por la probabilidad de membresía al estado latente  $\ell$  en el tiempo  $t$ . Debido a que los estados latentes son mutuamente excluyentes y exhaustivos en cada periodo de tiempo, es decir, que cada individuo es miembro de uno y solo un estado latente en el tiempo  $t$ , se debe cumplir que:

$$\sum_{\ell_t=1}^C \delta_{\ell_t} = 1. \quad (3.3)$$

En otras palabras, dentro de un tiempo en particular  $t$ , la prevalencia en los estados latentes debe sumar 1.

### 3.1.2. Las probabilidades de transición

La probabilidad de transición a un estado latente  $j$  en el tiempo  $t$ , condicionada a su membresía al estado latente  $i$  en el tiempo  $t-1$  se denota mediante  $\tau_{j|i}$ . Estas probabilidades por lo general son acomodadas en una matriz de probabilidades de transición de la siguiente manera:

$$\begin{bmatrix} \tau_{1|1} & \tau_{2|1} & \cdots & \tau_{C|1} \\ \tau_{1|2} & \tau_{2|2} & \cdots & \tau_{C|2} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{1|C} & \tau_{2|C} & \cdots & \tau_{C|C} \end{bmatrix}, \quad (3.4)$$

donde  $\tau_{j|i} = P(L_t = j | L_{t-1} = i)$ .

Los individuos que se encuentran en un estado latente  $\ell_t$  en el tiempo  $t$ , estarán en solo un estado latente en el tiempo  $t+1$ ,  $\ell_{t+1}$  el cual puede ser el mismo estado latente  $\ell_t$  o uno diferente. En cada uno de los estados latentes de la membresía estos son mutuamente excluyentes y exhaustivos, es decir, los individuos pertenecen solo a un estado latente en cada tiempo. Por consiguiente se cumple:

$$\sum_{\ell_{t+1}=1}^C \tau_{\ell_{t+1}|\ell_t} = 1, \quad (3.5)$$

En otras palabras, cada fila de la matriz de probabilidades de transición debe sumar 1.

### 3.1.3. Las probabilidades de respuesta al ítem

La probabilidad de obtener una respuesta  $k$  al ítem  $m$ , condicionada a la membresía en el estado latente  $\ell_t$  en el tiempo  $t$  se denota mediante  $\rho_{mkt|\ell_t}$ . Nuevamente cada individuo proporciona una y sola una respuesta a un ítem  $m$  en un tiempo específico  $t$  y por tanto se debe cumplir que:

$$\sum_{k=1}^{r_m} \rho_{mkt|\ell_t} = 1, \quad (3.6)$$

para todo  $m, t$ . En otras palabras, para los individuos en el estado latente  $\ell_t$  en el tiempo  $t$ , las probabilidades de cada respuesta alternativa a la variable  $m$  deben sumar 1.

### 3.1.4. Restricción de parámetros en el análisis de transición de clases latentes

Algunas restricciones en los parámetros pueden ser usadas en el modelo LTA. Estas se fijan en la prevalencia de los estados latentes, en las probabilidades de respuesta al ítem, o en las probabilidades de transición en el modelo de transición latente. Dos tipos diferentes de restricción de parámetros son comúnmente usados: los parámetros pueden ser fijos o restringidos. Un parámetro que es fijo en cierto valor particular no es estimado. Antes de comenzar la estimación, su valor debe ser especificado en el rango correspondiente. Este valor fijo está fuera de los límites del procedimiento de estimación. Cuando los parámetros están restringidos, estos son ubicados en un conjunto equivalente junto con otros parámetros. La estimación de todos los parámetros en un conjunto equivalente está restringida a ser igual al mismo valor, el cual puede ser cualquier valor en el rango correspondiente. Un modelo simple de transición de clases latentes puede contener cualquier combinación de parámetros fijos y restringidos. Debido a que los parámetros fijos no se estiman, estos no contribuyen al número total de parámetros a estimar. El conjunto equivalente cuenta como un único parámetro de estimación, independientemente de cuantos parámetros conforman el conjunto.

Dentro de las distintas restricciones es una buena idea restringir las probabilidades de respuesta al ítem en el modelo LTA para que estas sean iguales a través del tiempo donde

sea posible y razonable hacerlo, es decir, donde quiera una medida de invarianza a través del tiempo puede ser supuesta. Existen razones conceptuales y prácticas para hacer esto.

La razón conceptual es que estos modelos de transición latente son mucho más fáciles de interpretar si las probabilidades de respuesta al ítem son idénticas a través del tiempo. Se ha discutido mucho acerca de la medida de invarianza a través de los grupos y también como probar hipótesis acerca de la medida de invarianza. Si las probabilidades de respuesta al ítem son iguales a lo largo de los grupos, entonces la interpretación de las clases latentes es idéntica a lo largo de los grupos también. Esto significa que cualquier diferencia en los grupos observados en la prevalencia de las clases latentes pueden ser interpretadas simplemente como diferencias cuantitativas; ciertas clases latentes serán más grandes en algunos grupos que en otros. Por otro lado, si las probabilidades de respuesta al ítem no son iguales a lo largo de los grupos, entonces el significado de las clases latentes varía a lo largo de los grupos. Si esto sucede las comparaciones en los grupos de las prevalencias en las clases latentes llegar a ser menos sencillas. Cuando se interpreta las diferencias en la prevalencia de clases latentes se hace necesario tener en cuenta cualquier diferencia en el significado de las clases latentes en el mismo tiempo. Desde luego, dependiendo de las preguntas de investigación, algunas veces las diferencias cualitativas pueden ser interesantes en si mismas. En el mismo sentido, la matriz de probabilidades de transición en el LTA es fácil de interpretar si hay medidas equivalentes a través de los tiempos. Si las probabilidades de respuesta al ítem son idénticas a través del tiempo, el significado de los estados latentes permanece constante a través del tiempo. Esto significa por ejemplo, que un elemento en la diagonal de la matriz de probabilidades de transición refleja la probabilidad de la membresía en el estado latente  $s$  en el tiempo  $t + 1$  condicionada a la membresía en el mismo estado latente  $s$  en el tiempo  $t$ . Para extender de que las probabilidades de respuesta al ítem correspondientes al estado latente  $s$  cambie a través del tiempo, el significado del estado latente  $s$  cambiará. Luego, ya no es tan claro como interpretar esta probabilidad de transición, porque junto con la interpretación cuantitativa también cambia a través del tiempo la membresía del estado latente. Es necesario entonces interpretar el cambio a través del tiempo en el sentido de las clases latentes. Sin embargo, dependiendo de las preguntas de investigación este cambio cualitativo puede ser muy interesante, particularmente si es significativo para su desarrollo.

Por otro lado, la razón práctica para restringir las probabilidades de respuesta al ítem a que sean iguales a través del tiempo es para ayudar a estabilizar la estimación y mejorar el problema de identificabilidad. En los modelos de transición latente pueden haber un gran número de probabilidades de respuesta al ítem, particularmente en modelos con más de dos mediciones en el tiempo. Imponer restricciones a los parámetros en el tiempo puede reducir considerablemente el número de parámetros a estimar.

Al comparar diferentes modelos de transición de clases latentes usaremos los principales índices de ajuste como la estadística  $G^2$  de razón de verosimilitud, que en Estadística, evalúa la

bondad del ajuste de dos modelos estadísticos en competencia, específicamente uno saturado al maximizar en todo el espacio de parámetros y otro restringido después de imponer alguna restricción, en función de la razón de sus probabilidades. También se analizarán los modelos bajo los criterios de información.



## Capítulo 4

# Estimación del modelo

Los modelos de transición de clases latentes están basados en la teoría de clases latentes, la cual nos dice que podemos tener subgrupos homogéneos de unidades en una muestra o en una población que pueden ser identificados o clasificados basándonos en sus respuestas a un conjunto de variables manifiestas. En el LTA las variables manifiestas se miden repetidamente a través del tiempo con el fin de identificar clases latentes en cada ocasión, y además se estiman las probabilidades de transición a lo largo del tiempo en la membresía a las clases latentes. Esta clase de modelos es una excelente manera de estudiar cambios en el desarrollo y crecimiento, tales como las transiciones entre los estados descritos en la teoría de Piaget, según lo citado por Li et al. (2017). Por ejemplo, Chung et al. (2005) utilizaron el LTA para modelar patrones secuenciales en el consumo de cigarrillos y alcohol para investigar la relación existente entre la adolescencia y el uso de estas sustancias en rango de edades. Ellos identificaron 5 clases latentes: (No consumidores, Solo alcohol, Solo cigarrillos, Alcohol y cigarrillos, Cigarrillos y estado de ebriedad) y descubrieron que la etapa adolescente o de pubertad está altamente relacionada con el alto consumo de estas sustancias en jóvenes mujeres del séptimo al duodécimo grado. El modelo general del LTA con más de dos períodos en el tiempo, expresa como la probabilidad de observar un vector particular o patrón de respuestas es una función de las probabilidades de membresía en cada estado latente en el primer período ( $\delta_{\ell_1}$ ), las probabilidades de transición hacia un estado latente en un período específico  $t$  condicionado a la membresía del estado latente en el período anterior  $\tau_{j|i}^{(t)}$  y las probabilidades de observar la  $k$ -ésima respuesta de cada ítem  $m$  en cada período  $t$  condicionada a la membresía al estado latente  $\ell_t$  ( $\rho_{mkt|\ell_t}$ ). El modelo se plantea para un individuo  $i$  como:

$$P(\mathbf{Y}_{i1} = \mathbf{y}_{i1}, \mathbf{Y}_{i2} = \mathbf{y}_{i2}, \dots, \mathbf{Y}_{iT} = \mathbf{y}_{iT}) = \sum_{\ell_1=1}^C \dots \sum_{\ell_T=1}^C \delta_{\ell_1} \tau_{\ell_2|\ell_1}^{(2)} \dots \tau_{\ell_T|\ell_{T-1}}^{(T)} \prod_{t=1}^T \prod_{m=1}^M \prod_{k=1}^{r_m} \rho_{mkt|\ell_t}^{I(y_{imt}=k)}, \quad (4.1)$$

donde  $\tau_{c_b|c_a}^{(t)}$  representa la probabilidad de transición de una unidad en la clase  $c_a$  en el periodo  $t-1$  a la clase  $c_b$  en el período  $t$  e  $y_{imt}$  denota a la respuesta que el individuo  $i$  da al ítem  $m$

en el tiempo  $t$ , la cual es una componente del vector  $\mathbf{y}_{it}=(y_{i1t}, y_{i2t}, \dots, y_{iMt})$ .

Uno de los principales objetivos del LTA es estimar las probabilidades de pertenencia a posteriori y las probabilidades de transición. Estas brindan información individual y son usadas para evaluar el rendimiento de cada unidad basados en su pertenencia a una clase. Por otro lado, las probabilidades de transición consideran las probabilidades de permanencia en la misma clase o el movimiento entre las clases de un periodo  $t - 1$  a un periodo  $t$  para las unidades examinadas.

#### 4.1. Función de verosimilitud

Los parámetros en el modelo de transición de clases latentes incluyen la probabilidad de membresía a la clase latente en el período 1, las probabilidades de transición del periodo 1 al período 2, del período 2 al periodo 3 y así sucesivamente, y los parámetros involucrados en la probabilidad de respuesta al ítem condicionadas a la clase latente.

Sea  $\mathbf{L} = (L_1, \dots, L_T)$  el vector de pertenencia a la membresía de la clase latente desde el periodo inicial  $t = 1$  a un periodo  $T$ , donde  $L_t$  toma los valores  $1, \dots, C$ . Correspondientemente, sea  $\mathbf{Y}_{it} = (Y_{i1t}, \dots, Y_{iMt})$  un vector aleatorio  $M$  dimensional de las respuestas de un individuo  $i$  a los  $M$  ítems en el periodo  $t$ , donde cada  $Y_{imt}$  toma valores  $1, \dots, r_m$ , para  $t = 1, \dots, T$ . Entonces la probabilidad conjunta de que el  $i$ -ésimo individuo pertenezca al arreglo de clases  $\boldsymbol{\ell}=(\ell_1, \dots, \ell_T)$  y proporcione un cierto patrón de respuestas  $\mathbf{y}_{i1}=(y_{i11}, y_{i21}, \dots, y_{iM1}), \dots, \mathbf{y}_{iT}=(y_{i1T}, y_{i2T}, \dots, y_{iMT})$  a cada uno de los ítems en los tiempos  $1, 2, \dots, T$ , respectivamente, es:

$$P[\mathbf{Y}_{i1} = \mathbf{y}_{i1}, \dots, \mathbf{Y}_{iT} = \mathbf{y}_{iT}, \mathbf{L} = \boldsymbol{\ell}] =$$

$$P(\mathbf{Y}_{i1} = \mathbf{y}_{i1}, \mathbf{Y}_{i2} = \mathbf{y}_{i2}, \dots, \mathbf{Y}_{iT} = \mathbf{y}_{iT}, L_1 = \ell_1, L_2 = \ell_2, \dots, L_T = \ell_T)$$

$$= \left[ \delta_{\ell_1} \prod_{t=2}^T \tau_{\ell_t|\ell_{t-1}}^{(t)} \right] \times \left[ \prod_{t=1}^T \prod_{m=1}^M \prod_{k=1}^{r_m} \rho_{mkt|\ell_t}^{I(y_{imt}=k)} \right], \quad (4.2)$$

donde  $\delta_{\ell_1} = P(L_1 = \ell_1)$ ,  $\tau_{\ell_t|\ell_{t-1}}^{(t)} = P(L_t = \ell_t | L_{t-1} = \ell_{t-1})$  y  $\rho_{mkt|\ell_t} = P(Y_{imt} = k | L_t = \ell_t)$ . Se asume que las componentes de los vectores  $\mathbf{Y}_{it}$  son condicionalmente independientes dentro de cada clase  $\ell_t$  para  $t = 1, \dots, T$ . Este supuesto, llamado de independencia local, nos permite realizar inferencias acerca de las variables de clases latentes (Lazarsfeld & Henry, 1968). Se asume, también, que la secuencia  $\{L_t\}$ , donde  $L_t$  denota a la clase de pertenencia de un individuo en el tiempo  $t$ , constituye una cadena de Markov de primer orden para  $t=1, 2, \dots, T$ . En la ecuación (4.2) solo la probabilidad marginal de la clase de la membresía en el tiempo de inicio  $t = 1$ ,  $\delta_{\ell_1}$ , podrá ser directamente estimada; las probabilidades marginales de la clase de membresía en los tiempos  $t \geq 2$  no son estimadas directamente, pero están en función de

otros parámetros. La prevalencia marginal de cada clase en un tiempo  $t \geq 2$  es calculada, según (2.1.4), mediante

$$\begin{aligned}\delta_{\ell_t}^{(t)} &= P(L_t = \ell_t) \\ &= \sum_{\ell_1=1}^C \cdots \sum_{\ell_{t-1}=1}^C \delta_{\ell_1} \prod_{t=2}^T \tau_{\ell_t|\ell_{t-1}}^{(t)}\end{aligned}\quad (4.3)$$

La función de verosimilitud para un individuo  $i$ , o distribución conjunta incondicional de su respuesta viene dada por:

$$\begin{aligned}L_i(\boldsymbol{\theta}) &= P(\mathbf{Y}_{i1} = \mathbf{y}_{i1}, \dots, \mathbf{Y}_{iT} = \mathbf{y}_{iT}) \\ &= \sum_{\ell_1=1}^C \cdots \sum_{\ell_T=1}^C \left[ \delta_{\ell_1} \prod_{t=2}^T \tau_{\ell_t|\ell_{t-1}}^{(t)} \right] \times \left[ \prod_{t=1}^T \prod_{m=1}^M \prod_{k=1}^{r_m} \rho_{m,k,t|\ell_t}^{I(y_{imt}=k)} \right],\end{aligned}\quad (4.4)$$

donde  $\boldsymbol{\theta} = (\delta, \tau_1, \dots, \tau_C, \rho_1, \dots, \rho_C)$  es el vector de parámetros que incluye a todos los parámetros del modelo.

Si contamos ahora con la respuesta de  $n$  individuos, la función de verosimilitud vendrá dada por

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n L_i(\boldsymbol{\theta}) = \prod_{i=1}^n \left[ \sum_{\ell_1=1}^C \cdots \sum_{\ell_T=1}^C \left[ \delta_{\ell_1} \prod_{t=2}^T \tau_{\ell_t|\ell_{t-1}}^{(t)} \right] \times \left[ \prod_{t=1}^T \prod_{m=1}^M \prod_{k=1}^{r_m} \rho_{m,k,t|\ell_t}^{I(y_{imt}=k)} \right] \right]. \quad (4.5)$$

## 4.2. Algoritmo EM

El algoritmo EM es el método de estimación clásico más utilizado para estimar parámetros en modelos de clases latentes. El algoritmo EM proporciona estimaciones de máxima verosimilitud incluso para datos faltantes en el caso de que sigan un patrón aleatorio y puede producir estimaciones estables independientemente de las soluciones de los límites. Este algoritmo, como su nombre lo sugiere, consta de dos pasos: E y M, de Esperanza y Maximización (o Minimización), respectivamente. Dado un valor inicial para el vector de parámetros de interés, el algoritmo itera consecutivamente los pasos E y M hasta que las estimaciones convergan hacia puntos fijos. El objetivo del algoritmo EM será en nuestro modelo estimar el vector de parámetros  $\boldsymbol{\theta}$ .

El software Mplus (version 7) desarrollado por Muthen y Muthen, 2017 será usado para implementar la estimación del modelo LTA mediante el algoritmo EM.

El Paso E consiste en el cálculo de

$$Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(j)}) = E_j(\log(f(\mathbf{y}, \mathbf{L}; \boldsymbol{\theta}))), \quad (4.6)$$

donde la esperanza se toma bajo la función de probabilidad condicional conjunta de  $\mathbf{L}$ , dado el patrón de respuestas  $\mathbf{Y} = \mathbf{y}$ , cuando se tiene la estimación  $\hat{\boldsymbol{\theta}}^{(j)}$  en el algoritmo, es decir, este

esperado se toma bajo la distribución conjunta de la cadena de Markov, dadas las respuestas y la estimación actual  $\hat{\boldsymbol{\theta}}^{(j)}$  de  $\boldsymbol{\theta}$ . Aquí  $f(\mathbf{y}, \mathbf{L}; \boldsymbol{\theta})$  denota a la siguiente función de verosimilitud aumentada, que vimos restringida a solo un individuo en (4.2). Esta función evaluada en un arreglo de clases  $\boldsymbol{\ell} = (\ell_1, \ell_2, \dots, \ell_T)$  es:

$$\begin{aligned}
f(\mathbf{y}, \boldsymbol{\ell}; \boldsymbol{\theta}) &= \\
&= \prod_{i=1}^n P(\mathbf{Y}_{i1} = \mathbf{y}_{i1}, \mathbf{Y}_{i2} = \mathbf{y}_{i2}, \dots, \mathbf{Y}_{iT} = \mathbf{y}_{iT}, L_1 = \ell_1, L_2 = \ell_2, \dots, L_T = \ell_T) \\
&= \prod_{i=1}^n P(\mathbf{Y}_{i1} = \mathbf{y}_{i1}, \mathbf{Y}_{i2} = \mathbf{y}_{i2}, \dots, \mathbf{Y}_{iT} = \mathbf{y}_{iT} | \ell_1, \ell_2, \dots, \ell_T) P(L_1 = \ell_1, L_2 = \ell_2, \dots, L_T = \ell_T) \\
&= \prod_{i=1}^n \left( \prod_{t=1}^T \prod_{m=1}^M \rho_{m1t|\ell_t}^{I(y_{imt}=1)} \rho_{m2t|\ell_t}^{I(y_{imt}=2)} \dots \rho_{mr_m t|\ell_t}^{I(y_{imt}=r_m)} \right) \delta_{\ell_1} \prod_{t=2}^T \tau_{\ell_t|\ell_{t-1}}^{(t)} \quad (4.7)
\end{aligned}$$

Así, la correspondiente función de log-verosimilitud aumentada será

$$\begin{aligned}
\log f(\mathbf{y}, \boldsymbol{\ell}; \boldsymbol{\theta}) &= n \log(\delta_{\ell_1}) + n \sum_{t=2}^T \log(\tau_{\ell_t|\ell_{t-1}}^{(t)}) + \\
&\quad + \sum_{i=1}^n \sum_{t=1}^T \sum_{m=1}^M [I(y_{imt} = 1) \log(\rho_{m1t|\ell_t}) + I(y_{imt} = 2) \log(\rho_{m2t|\ell_t}) + \\
&\quad \dots + I(y_{imt} = r_m) \log(\rho_{mr_m t|\ell_t})] \quad (4.8)
\end{aligned}$$

Las probabilidades a posteriori asociadas a  $E_j$  se pueden obtener mediante el teorema de Bayes. Aquí esta medida de probabilidad se toma condicionada a que  $\mathbf{Y} = \mathbf{y}$ , donde  $\mathbf{Y}$  es un arreglo aleatorio con componentes  $Y_{imt}$ , y viene dada por:

$$\begin{aligned}
P_{j\mathbf{y}}(\ell_1, \ell_2, \dots, \ell_T) &= P(L_1 = \ell_1, L_2 = \ell_2, \dots, L_T = \ell_T | \mathbf{Y} = \mathbf{y}) \\
&= \frac{P(\mathbf{Y} = \mathbf{y} | L_1 = \ell_1, \dots, L_T = \ell_T) P(L_1 = \ell_1, \dots, L_T = \ell_T)}{P(\mathbf{Y} = \mathbf{y})} \\
&= \frac{P(\mathbf{Y}_1 = \mathbf{y}_1, \mathbf{Y}_2 = \mathbf{y}_2, \dots, \mathbf{Y}_n = \mathbf{y}_n | L_1 = \ell_1, L_2 = \ell_2, \dots, L_T = \ell_T) \hat{\delta}_{\ell_1} \prod_{t=2}^T \hat{\tau}_{\ell_t|\ell_{t-1}}^{(t)}}{P(\mathbf{Y} = \mathbf{y})},
\end{aligned}$$

siendo  $\mathbf{Y}_i$  la matriz aleatoria con las componentes de  $\mathbf{Y}$  para el sujeto  $i$  y el subíndice  $j$  indica que se están usando aquí las estimaciones de  $\hat{\boldsymbol{\theta}}^{(j)}$ . Así,

$$\begin{aligned}
P_{j\mathbf{y}}(\ell_1, \ell_2, \dots, \ell_T) &= \\
&= \frac{(\prod_{i=1}^n P(\mathbf{Y}_i = \mathbf{y}_i | L_1 = \ell_1, L_2 = \ell_2, \dots, L_T = \ell_T)) \hat{\delta}_{\ell_1} \prod_{t=2}^T \hat{\tau}_{\ell_t|\ell_{t-1}}^{(t)}}{P(\mathbf{Y} = \mathbf{y})} \\
&= \frac{\left( \prod_{i=1}^n \prod_{t=1}^T \prod_{m=1}^M \prod_{k=1}^{r_m} (1 - \hat{P}_{mkt|\ell_t})^{I(y_{imt}=k)} \hat{P}_{mkt|\ell_t}^{I(y_{imt}=k)} \right) \hat{\delta}_{\ell_1} \prod_{t=2}^T \hat{\tau}_{\ell_t|\ell_{t-1}}^{(t)}}{P(\mathbf{Y} = \mathbf{y})}, \quad (4.9)
\end{aligned}$$

donde:

$$P(\mathbf{Y} = \mathbf{y}) = \sum_{\ell_1=1}^C \cdots \sum_{\ell_T=1}^C \left( \prod_{i=1}^n \prod_{t=1}^T \prod_{m=1}^M \prod_{k=1}^{r_m} (1 - \hat{P}_{mkt|\ell_t})^{I(y_{imt}=k)} \hat{P}_{mkt|\ell_t}^{I(y_{imt}=k)} \right) \hat{\delta}_{\ell_1} \prod_{t=2}^T \hat{\tau}_{\ell_t|\ell_{t-1}}^{(t)}$$

Así, tomándose esperados con respecto a esta función de probabilidades, obtendremos la función

$$\begin{aligned} Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(j)}) &= E_j(f(\mathbf{y}, L, \boldsymbol{\theta})) = nE_j(\log(\delta_{\ell_1})) + n \sum_{t=2}^T E_j \left( \log \left( \tau_{\ell_t|\ell_{t-1}}^{(t)} \right) \right) \\ &\quad + \sum_{i=1}^n \sum_{t=1}^T \sum_{m=1}^M \sum_{k=1}^{r_m} [I(y_{imt} = k) E_j(\log(P_{mkt|\ell_t}))] \\ &= n \sum_{\ell_1=1}^L \log(\delta_{\ell_1}) P_{jy}(\ell_1) + n \sum_{t=2}^T \sum_{\ell_{t-1}=1}^C \sum_{\ell_t=1}^C \log \left( \tau_{\ell_t|\ell_{t-1}}^{(t)} \right) P_{jy}(\ell_{t-1}, \ell_t) \\ &\quad + \sum_{i=1}^n \sum_{t=1}^T \sum_{m=1}^M \sum_{k=1}^{r_m} \sum_{\ell_t=1}^C [I(y_{imt} = k) \log(P_{mkt|\ell_t})] P_{jy}(\ell_t), \end{aligned} \quad (4.10)$$

donde estamos denotando por  $P_{jy}(\ell_t)$  y  $P_{jy}(\ell_{t-1}, \ell_t)$  a las funciones de probabilidad marginal univariada y bivariada de (4.8).

El paso M se centra en la maximización del valor esperado de la función de log-verosimilitud aumentada anterior con respecto a los parámetros del modelo. Esto se consigue obteniendo las derivadas parciales y resolviendo las ecuaciones normales.

Para los parámetros  $\delta_\ell$  en donde omitiremos por simplicidad el subíndice 1, entendiéndose que son las probabilidades iniciales de pertenencia a las clases, debemos maximizar  $Q(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(j)})$  con respecto a  $\delta_\ell$  sujeto a la condición que  $\sum_{\ell=1}^C \delta_\ell = 1$ , lo cual equivale a maximizar el lagrangiano

$$\mathcal{L} = n \sum_{\ell=1}^C \log(\delta_\ell) P_{jy}(\ell) + \lambda \left( 1 - \sum_{\ell=1}^L \delta_\ell \right)$$

y para lo cual basta resolver las condiciones de primer orden. Así,

$$\frac{\partial \mathcal{L}}{\partial \delta_\ell} = 0 = \frac{n}{\delta_\ell} P_{jy}(\ell) - \lambda$$

Despejando,

$$\delta_\ell = \frac{n P_{jy}(\ell)}{\lambda}.$$

Como  $1 = \sum_{\ell=1}^C \delta_\ell = \frac{n}{\lambda} \sum_{\ell=1}^C P_{jy}(\ell)$ ,

$$\lambda = n \sum_{\ell=1}^C P_{jy}(\ell).$$

Por tanto, expresándose esta probabilidad de manera recursiva, se tendrá que

$$\delta_\ell^{(j+1)} = \frac{P_{jy}(\ell)}{\sum_{\ell=1}^C P_{jy}(\ell)}, \quad (4.11)$$

donde  $P_{jy}(\ell)$  denota la función de probabilidad marginal de  $L_1$  con respecto a la función de probabilidad conjunta  $P_{jy}(\ell_1, \ell_2, \dots, \ell_T)$  en (4.8), la cual recordemos se calcula en base a la estimación  $\hat{\theta}^{(j)}$ .

Análogamente, para las probabilidades de transición  $\tau_{b|a}^{(t)}$  y omitiendo por simplicidad los subíndices de los índices en las clases, bastará maximizarse el lagrangiano

$$\mathcal{L} = n \sum_{t=2}^T \sum_{a=1}^C \sum_{b=1}^C \log \left( \tau_{b|a}^{(t)} \right) P_{jy}(a, b) + \sum_{t=2}^T \lambda_{t-1} \left( 1 - \sum_{b=1}^C \tau_{b|a}^{(t)} \right) \quad (4.12)$$

donde  $\lambda_1, \lambda_2, \dots, \lambda_{t-1}$  son multiplicadores de Lagrange. Derivando con respecto a un  $\tau_{b|a}^{(t)}$  particular e igualando a 0, se tendrá que

$$0 = \frac{\partial \mathcal{L}}{\partial \tau_{b|a}^{(t)}} = \frac{n}{\tau_{b|a}^{(t)}} P_{jy}(a, b) - \lambda_{t-1}$$

Entonces,

$$\tau_{b|a}^{(t)} = n \frac{P_{jy}(a, b)}{\lambda_{t-1}}.$$

Como

$$1 = \sum_{b=1}^C \tau_{b|a}^{(t)} = \frac{n}{\lambda_{t-1}} \sum_{b=1}^C P_{jy}(a, b),$$

se tiene que

$$\lambda_{t-1} = n \sum_{b=1}^C P_{jy}(a, b).$$

Así,

$$\tau_{b|a}^{(t)(j+1)} = \frac{P_{jy}(a, b)}{\sum_{b=1}^C P_{jy}(a, b)}$$

o en su notación original

$$\tau_{\ell_t|\ell_{t-1}}^{(t)(j+1)} = \frac{P_{jy}(\ell_{t-1}, \ell_t)}{\sum_{\ell_t=1}^C P_{jy}(\ell_{t-1}, \ell_t)}. \quad (4.13)$$

De (4.9), el término a maximizar para actualizar las estimaciones de los parámetros de respuesta al ítem es:

$$B = \sum_{t=1}^T \sum_{m=1}^M \sum_{k=1}^{r_m} \sum_{\ell_t=1}^C r_{mt} \log \left( \rho_{mkt|\ell_t} \right) P_{jy}(\ell_t), \quad (4.14)$$

donde  $r_{mt} = \sum_{i=1}^n I(y_{imt} = 1)$  denota el número total de respuestas correctas de los  $n$  individuos al ítem  $m$  en el tiempo  $t$ .

Vale recordar que los parámetros  $\rho_{mkt|\ell_t} = P(Y_{mt} = k | L_t = \ell_t)$  hacen referencia a la probabilidad de que un individuo perteneciente a la clase  $\ell_t$  en el período  $t$  de una respuesta  $k$  al ítem  $m$ . Como su respuesta se asume que es única debe cumplirse que  $\sum_{k=1}^{r_m} \rho_{mkt|\ell_t} = 1$  y por tanto el lagrangiano de este problema viene dado por:

$$\mathcal{L} = B + \sum_{t=1}^T \lambda_t \left( 1 - \sum_{k=1}^{r_m} \rho_{mkt|\ell_t} \right) \quad (4.15)$$

donde los  $\lambda_t$  son multiplicadores de Lagrange.

Derivando  $L$  con respecto a  $\rho_{mkt|\ell_t}$  e igualando a 0 obtendremos que:

$$0 = \frac{\partial L}{\partial \rho_{mkt|\ell_t}} = \frac{1}{\rho_{mkt|\ell_t}} r_{mt} P_{jy}(\ell_t) - \lambda_t$$

y

$$\rho_{mkt|\ell_t} = \frac{r_{mt} P_{jy}(\ell_t)}{\lambda_t} \quad (4.16)$$

y como

$$1 = \sum_{k=1}^{r_m} \rho_{mkt|\ell_t} = \frac{1}{\lambda_t} \sum_{k=1}^{r_m} P_{jy}(\ell_t) r_{mt}, \quad (4.17)$$

$$\lambda_t = \sum_{k=1}^{r_m} P_{jy}(\ell_t) r_{mt}$$

Luego, reemplazando este resultado en (4.16) tenemos:

$$\rho_{mkt|\ell_t} = \frac{r_{mt} P_{jy}(\ell_t)}{\sum_{k=1}^{r_m} r_{mt} P_{jy}(\ell_t)}$$

$$= \frac{P_{jy}(\ell_t) \sum_{i=1}^n I(Y_{imt} = k)}{\sum_{k=1}^{r_m} r_{mt} P_{jy}(\ell_t)}$$

Luego, esto en términos recursivos se expresa por:

$$\rho_{mkt|\ell_t}^{(j+1)} = \frac{P_{jy}(\ell_t) \sum_{i=1}^n I(Y_{imt} = k)}{\sum_{k=1}^{r_m} P_{jy}(\ell_t) \sum_{i=1}^n I(Y_{imt} = k)}.$$

La iteración de los pasos E y M desarrollados previamente producirán una secuencia de parámetros estimados que convergerán a un máximo local o global de la función de verosimilitud. Los desafíos para la inferencia de máxima verosimilitud en muestras pequeñas para el

modelo de transición de clases latentes se encuentran principalmente en los parámetros que han de ser estimados y que podrían encontrarse en los límites del espacio paramétrico (es decir, cerca de 0 o 1), lo que dificulta la obtención adecuada de los errores estándares.

Sin embargo, las probabilidades de respuesta al ítem cercanas a cero o a 1 son altamente preferibles desde una perspectiva de medición, y cuando algunos de estos parámetros son estimados en el límite, es imposible obtener los errores estándar a partir de la inversa de la matriz hessiana  $H$ .

El algoritmo EM se utiliza para la estimación por máxima verosimilitud en problemas de datos incompletos. La matriz de información de Fisher observada, derivada de la función de verosimilitud  $L(\boldsymbol{\theta})$  según (4.5), puede emplearse para calcular los errores estándar de las estimaciones de parámetros que determinan el número de clases. Sin embargo, obtener las derivadas de segundo orden de la función log-verosimilitud para datos incompletos puede ser un desafío. (Louis, 1982) demostró que la matriz de información observada para datos incompletos puede expresarse en función de las derivadas parciales de primer y segundo orden de la función log-verosimilitud para datos completos, introducida en el marco teórico del EM. Por lo tanto, se adoptará el método de Louis para calcular la matriz de varianza-covarianza de las estimaciones de parámetros.

### 4.3. Criterios para la selección del modelo

Varios estadísticos de ajuste relativos están disponibles para comparar dos o más modelos. En general, los índices de criterio de información o estadísticos de prueba de razón de verosimilitud pueden ser usados dentro del contexto del modelamiento de variables latentes. El primero es apropiado para comparar modelos anidados o no anidados, mientras que el otro solo puede ser utilizado para modelos anidados. En este estudio nos basaremos principalmente en los índices de criterio de información, específicamente en el AIC, CAIC y BIC. A menudo múltiples índices se usan para comparar el ajuste de los datos al modelo entre un conjunto de modelos para los mismos datos cuando los estimadores de máxima verosimilitud de los parámetros se han obtenido. Los índices de criterio de información están basados en una forma de penalización de la función de verosimilitud. Valores pequeños de estos índices indican un mejor ajuste. Sin embargo, diferentes índices podrían seleccionar diferentes modelos de ajuste para los mismos datos debido a las diferencias en la función de penalización aplicada a la verosimilitud. El índice AIC está dado por:

$$\text{AIC} = -2 \log(L(\hat{\boldsymbol{\theta}})) + 2p \quad (4.18)$$

donde  $p$  es el número de parámetros estimados y  $2p$  es usada como una penalización para la sobreparametrización y  $L$  es la función de verosimilitud del modelo estimado en (4.5). Uno de los problemas con el AIC es que tiende a seleccionar modelos más complejos. La falta de penalización para el tamaño de muestra conduce a la inconsistencia en el desempeño o actuación del AIC y una tendencia a sobreestimar el correcto número de clases (Peel &

McLachlan, 2000). Sin embargo, el AIC es correcto y eficiente asintóticamente, si el verdadero modelo no está entre los modelos que se comparan (Vrieze, 2012). Para tener en cuenta el tamaño de la muestra, el BIC puede ser usado y viene dado por:

$$\text{BIC} = -2\log(L(\hat{\boldsymbol{\theta}})) + p\log(n) \quad (4.19)$$

donde  $L$  es la verosimilitud del modelo estimado con  $p$  parámetros libres y  $\log(n)$  es el logaritmo del tamaño de muestra total  $n$ . Como podemos observar en la ecuación (4.19), la función de penalización para el BIC está basada en el número de parámetros estimados así como también en el tamaño de muestra. El BIC tiende a aplicar mayor penalización a la función de verosimilitud que el AIC cuando modelos complejos son estimados. Como consecuencia, el BIC es más preferido a la hora de seleccionar modelos simples que el AIC debido a la inclusión del tamaño de muestra en la función de penalización. La penalización del BIC con  $n$  hace de la significación estadística más y más difícil de lograr a su vez que el tamaño de muestra aumenta (Vrieze, 2012).

En modelos de mixtura de variables latentes, la complejidad del modelo verdadero, la separación de las clases y la proporción de las clases han sido demostradas que afectan la utilidad de estos dos índices de ajuste (Lubke & Neale, 2006); (Nylund-Gibson et al., 2019); (Vrieze, 2012).

Otros índices relacionados al AIC y BIC son la versión consistente del AIC (CAIC) (Bozdogan, 1987):

$$\text{CAIC} = -2\log(L(\hat{\boldsymbol{\theta}})) + k(\log(n) + 1) \quad (4.20)$$

y el índice de tamaño de muestra ajustado: (Lanza & Cooper, 2016)

$$\text{ABIC} = -2\log(L(\hat{\boldsymbol{\theta}})) + q\log\left(\frac{n+2}{24}\right) \quad (4.21)$$

Otro indicador de ajuste también empleado es la entropía, cuyo valor se encuentra entre 0 y 1. Un mayor valor de entropía indica que hay una mejor separación entre las clases latentes (Lanza & Cooper, 2016). Una vez elegido el modelo, se definen las clases latentes en función de las probabilidades de respuesta a los diferentes ítems. El objetivo principal es buscar la mejor descripción de los datos observados a través de la identificación de clases latentes. La entropía mide la certeza al clasificar los estados latentes. Una entropía más alta denota una mejor separación de clases.

Finalmente, otra alternativa para saber que tan bien un modelo de transición de clases latentes se ajusta a los datos observados es mediante el cálculo de la estadística de razón de verosimilitud,

$$G^2 = 2 \sum_{\mathbf{y}} f_{\mathbf{y}} \log\left(\frac{f_{\mathbf{y}}}{\hat{f}_{\mathbf{y}}}\right), \quad (4.22)$$

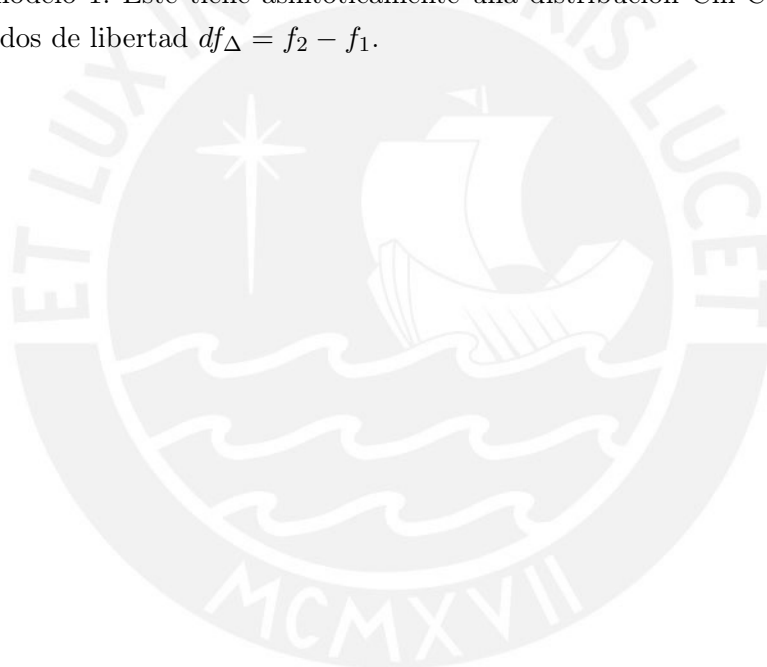
donde  $f_{\mathbf{y}}$  representa la frecuencia observada de casos que presentan un patrón de respuestas  $\mathbf{y}$  a las variables manifiestas y  $\hat{f}_{\mathbf{y}}$  representa la frecuencia esperada de este patrón bajo el modelo

a evaluar. Este es usado para evaluar el ajuste absoluto de un modelo y valores grandes de él nos evidencia de que el modelo debe ser rechazado. Más aún, si se tienen muestras grandes y frecuencias esperadas de patrones no muy pequeñas, es sabido que podríamos utilizar este estadístico para realizar pruebas de hipótesis, pues asintóticamente este tiene distribución Chi-cuadrado con  $df$  grados de libertad, siendo  $df$  igual al número de celdas o patrones de respuestas, menos el número de parámetros a estimar, menos 1. En un modelos de clases latentes e ítems dicotómicos, por ejemplo:

$$df = 2^p - C(1 + p),$$

donde  $C$  es el número de clases y  $p$  el número de ítems o variables manifiestas.

El test de diferencia de la razón de verosimilitud (LRDT) puede ser luego usado con el estadístico  $G_{\Delta}^2 = G_2^2 - G_1^2$  para comparar dos modelos anidados, por ejemplo, un modelo 2 anidado en un modelo 1. Este tiene asintóticamente una distribución Chi-Cuadrado con la diferencia de grados de libertad  $df_{\Delta} = f_2 - f_1$ .



## Capítulo 5

# Aplicación a un estudio de acoso escolar

En este capítulo se describe la implementación del modelo propuesto (LTA) en un conjunto de datos reales en el contexto educativo. En particular, se analiza una prueba sobre acoso escolar (bullying) realizada con 1,180 estudiantes, de quinto hasta noveno grado en 9 colegios del oeste de Estados Unidos durante un periodo de tres semestres. Esta aplicación está basada en el artículo de Ryoo et al. (2018).

La etapa de educación primaria y secundaria es un período importante para estudiar el acoso escolar (bullying). Curiosamente, a medida que aumenta este fenómeno durante la escuela primaria y secundaria, las actitudes del estudiante frente a esto se vuelven cada vez más arraigadas durante los años del nivel secundario en comparación con los años del nivel primario ((Oliver et al., 1994);(Pellegrini, 2002)). Además, los estudiantes son más propensos a encontrarse con compañeros agresivos durante la etapa de secundaria (Bukowski et al., 2000). Los adolescentes que tuvieron comportamientos agresivos durante la educación secundaria son usualmente calificados por sus compañeros no agresivos como populares, y son identificados como personajes importantes en las redes sociales según (Veenstra et al., 2005). Los estudiantes, calificados como peleadores son significativamente menos tímidos en comparación con el resto de sus compañeros, incluidos las víctimas y las víctimas de estos y los estudiantes no implicados. En resumen, los adolescentes pueden reunirse con compañeros agresivos o tolerar comportamientos de acoso con el fin de explorar nuevos roles sociales y desafiar normas sociales respaldadas por adultos, lo que proporciona un entorno favorable al aumento de comportamientos agresivos durante los años de la etapa secundaria.

Aún cuando las actitudes de los estudiantes hacía el acoso escolar se vuelven más positivas durante la etapa secundaria, algunos estudios han demostrado que tanto los estudiantes de primaria como de secundaria muestran preocupación por el acoso y seguridad escolar (Astor et al., 2001) y (Robers et al., 2013). Estos estudios muestran que los estudiantes preocupados por el acoso en la escuela, van desde el 19 % al 40 %, incluyendo el 27 % de los alumnos

de 3<sup>o</sup> al 5<sup>o</sup> grado preocupados también por el acoso cibernético (D'Antona et al., 2010). La preocupación y el miedo contribuyen a que muchos estudiantes se sientan inseguros, independientemente de la raza o credo.

Como resultado de esta preocupación por el acoso y la seguridad escolar, algunos estudiantes terminan abandonando la escuela, no asistiendo a ciertos cursos y evitando las actividades extracurriculares. Además, algunas investigaciones han mostrado que el efecto de la preocupación por el acoso escolar varía según los grados de los alumnos. Por ejemplo, (Hughes et al., 2015) demostraron en un estudio realizado en adolescentes que abandonaron la escuela, que sentirse inseguro debido a la intimidación es un factor significativo el abandono de esta en los estudiantes de 12<sup>o</sup> grado. Por otro lado, las preocupaciones sobre el acoso cibernético fueron un factor significativo para que los estudiantes de 9<sup>o</sup> grado abandonen la escuela por sentirse inseguros. La investigación también ha encontrado diferencias entre género y los distintos grados escolares. En resumen, menos acoso escolar contribuye a tener sentimientos de seguridad para los alumnos y alumnas de 5<sup>o</sup> grado, y para las chicas del 8<sup>o</sup> grado, pero no para los de 11<sup>o</sup>. (Bachman et al., 2011). Como consecuencia, de las preocupaciones sobre el acoso escolar, muchos estudiantes (65%) no creen que sus escuelas estén haciendo un buen trabajo en el tratamiento de la intimidación escolar (Marley, 2008).

Las preocupaciones generalizadas de los estudiantes sobre el acoso, su insatisfacción con las intervenciones por parte del colegio contra el acoso escolar y los resultados perjudiciales asociados con esas preocupaciones (por ejemplo, abandono escolar) destacan la importancia de estudiar las preocupaciones de los estudiantes sobre el acoso escolar a lo largo del tiempo, incluyendo variables como el género y el nivel de grado.

La mayoría de estudios que evalúan los cambios en los temores de los estudiantes frente al acoso escolar y también las tasas de prevalencias han utilizado diseños de corte transversales. Basado en la literatura, pocos estudios han utilizado diseños longitudinales para examinar los cambios en las preocupaciones de los estudiantes sobre el acoso escolar a lo largo del tiempo, lo que puede dar lugar a otra interpretación frente a este problema a lo largo del tiempo.

En el presente estudio, adaptado de Ryoo et al. (2018), se empleó el modelo LTA para estimar la prevalencia de las preocupaciones por el acoso escolar y el grado de transición entre los estados latentes a lo largo del tiempo. En lugar de definir directamente las clases latentes en la preocupación de los estudiantes con respecto al acoso escolar usando una medición observacional, los modelos de transición de clases latentes permiten clasificar subgrupos heterogéneos (o clases latentes) en relación con la preocupación de los alumnos por el acoso escolar. Esta aplicación busca responder las siguientes preguntas:

1. ¿Existen subgrupos de estudiantes cualitativamente distintos que muestren patrones particulares acerca de la preocupación en relación con el acoso escolar?
2. ¿Cuál es la probabilidad de que un estudiante se encuentre en un estado latente diferente?
3. ¿Hay algún cambio entre los estados latentes a lo largo del tiempo? En caso afirmativo, ¿Cómo se puede caracterizar este cambio?

4. ¿Se ve afectada la prevalencia en un estado latente por otras variables?

## 5.1. Metodología del recojo de la información

En esta aplicación, participaron 1180 alumnos de 5<sup>o</sup> al 9<sup>o</sup> que asistían a nueve colegios del oeste de Estados Unidos. Todos los alumnos recibieron formularios de consentimiento para entregar a sus padres, en los que se explicaba la naturaleza del estudio. Los participantes (97%) otorgaron su consentimiento para formar parte del estudio. Debido a los cambios y traslado de colegios, el número de estos se amplió a 22 durante los tres semestres en estudio. La edad promedio fué de 12.2 años (DE = 1.29 años) y el 9.9% indicó que el inglés no era su lengua materna. Los grados para el periodo inicial de estudio se distribuyeron como se aprecia en la Figura 5.4

Las evaluaciones se realizaron a lo largo de tres semestres. La tasa de deserción escolar fue de 5.59% en el segundo semestre y de 15.34% en el tercero.

## 5.2. Metodología de análisis

Cada alumno completó un cuestionario demográfico que incluía preguntas sobre género, edad, grado, lengua materna y raza o etnia. Estos alumnos respondieron luego un cuestionario sobre la medida de intimidación en la costa del pacifico (PRBm); (Konishi et al., 2009) cuyos ítems se aprecian en la Tabla 5.1. Este aborda sobre las experiencias y preocupaciones de los estudiantes en cuanto al acoso y victimización.

## 5.3. Preocupación por el acoso escolar

Los alumnos respondieron a seis preguntas acerca de la preocupación que tienen por ser sujetos de acoso escolar en una escala tipo Likert de cuatro puntos ("No, no en absoluto", "No, no mucho", "Sí, un poco", "Sí, mucho"). En lugar de utilizar las cuatro opciones de respuesta en esta escala, se utilizó aquí solo dos opciones agrupando las dos primeras categorías como un No y fusionando las dos últimas categorías como un Sí. (ver la Tabla 5.1)

La estimación de la consistencia interna, mediante el coeficiente  $\alpha$  de Cronbach, fué de 0.79 en el tiempo 1, 0.79 en el tiempo 2 y 0.80 en el tiempo 3.

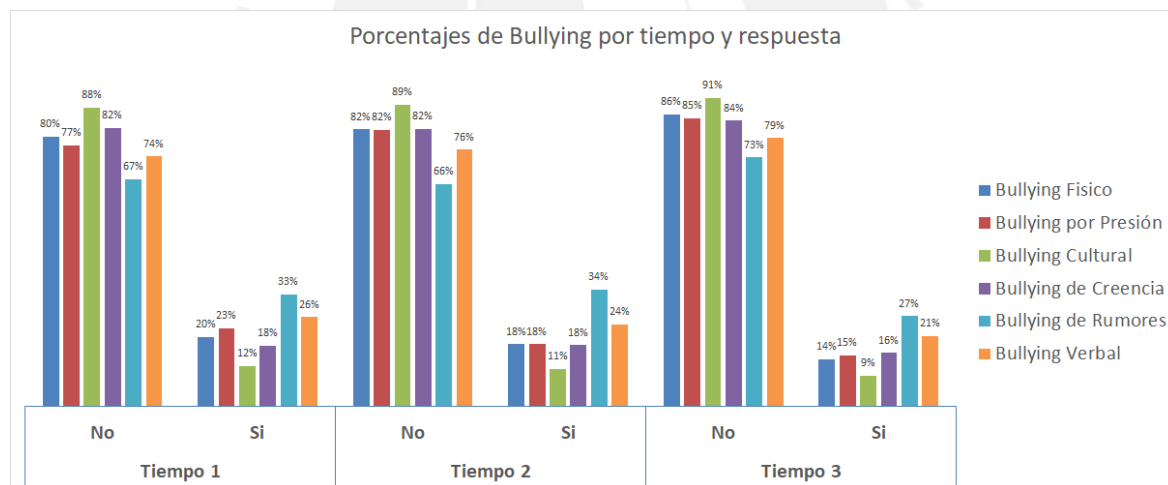
La Tabla 5.1 resume las proporciones de respuesta a los ítems en detalle y muestra que los

cambios en las proporciones marginales no son consistentes entre los ítems.

Tabla 5.1: Respuestas marginales para los 6 ítems en la encuesta sobre el acoso escolar PRBm

¿Qué tan preocupado o temeroso está Ud. de que pueda	Periodo 1		Periodo 2		Periodo 3	
	No	Si	No	Si	No	Si
Ser agredido o herido físicamente por otro alumno o grupo de alumnos. (B. Físico)	80	20	82	18	86	14
Ser convencido por otros alumnos a hacer cosas que no desea. (B. por presión)	77	23	82	18	85	15
Ser sujeto de burla o que lo dejen de lado por su cultura o raza. (B. Cultural)	88	12	89	11	91	9
Ser sujeto de burla o que lo dejen de lado por sus opiniones, creencias o ideas. (B. de Creencias)	82	18	82	18	84	16
Que se difundan chismes o rumores sobre su persona. (B. de Rumores)	67	33	66	35	74	27
Ser amenazado verbalmente o avergonzado en la escuela. (B. Verbal)	74	26	76	24	79	21

Figura 5.1: Porcentajes de Bullying según tiempo y respuesta



## 5.4. Resultados

Para abordar las preguntas de investigación en esta aplicación ajustaremos, como lo sugieren Ryoo et al. (2018), el modelo LTA siguiendo 4 pasos. Todos los análisis y resultados que se presentan a continuación se ejecutaron en el programa estadístico Mplus 7.4. Los comandos (sintaxis) necesarios para reproducir los resultados que aquí mostramos se muestran en el Anexo y la base de datos se encuentra en el material suplementario del artículo Ryoo et al. (2018).

**Paso 0: Diagnostico y exploración de datos transversales mediante un análisis de clases latentes.**

Como se muestra en la Figura 5.1, la proporción muestral de que los distintos tipos de acoso escolar cambia a lo largo del tiempo. En este paso, agruparemos las respuestas a los ítems, en cada periodo de tiempo, utilizando un Análisis de Clases Latentes (ACL).

La Tabla 5.2 resume los resultados del análisis de clases latentes, que muestran que los modelos de tres y cuatro soluciones resultaron favorecidos con respecto a los índices de ajuste de AIC, BIC, CAIC y ABIC siendo estos los más bajos. Aunque la entropía que mide la certeza en la clasificación de los estados latentes favoreció al modelo de dos soluciones, no se consideró este modelo, porque la simple separación alta/baja no proporcionaba mayores descripciones de las preocupaciones de los estudiantes sobre el acoso escolar que un análisis factorial.

Tabla 5.2: Resultados del análisis de clases latentes en cada periodo de tiempo.

	<b>G2</b>	<b>AIC</b>	<b>BIC</b>	<b>CAIC</b>	<b>ABIC</b>	<b>Entropía</b>	<b>GL</b>
<b>Periodo 1</b>							
ACL-2 soluciones	191	217	283	296	242	0.81	50
ACL-3 soluciones	53	93	194	214	131	0.79	43
ACL-4 soluciones	35	89	225	252	140	0.78	36
ACL-5 soluciones	23	91	263	297	155	0.76	29
ACL-6 soluciones	20	102	310	351	179	0.77	22
ACL-7 soluciones	17	113	356	404	203	0.67	15
<b>Periodo 2</b>							
ACL-2 soluciones	161	187	252	265	211	0.84	50
ACL-3 soluciones	59	99	199	219	135	0.82	43
ACL-4 soluciones	41	95	231	258	145	0.80	36
ACL-5 soluciones	30	98	268	302	160	0.84	29
ACL-6 soluciones	23	105	310	351	180	0.82	22
ACL-7 soluciones	18	114	355	403	203	0.82	15
<b>Periodo 3</b>							
ACL-2 soluciones	187	213	277	290	236	0.85	50
ACL-3 soluciones	75	115	213	233	149	0.84	43
ACL-4 soluciones	54	108	240	267	154	0.84	36
ACL-5 soluciones	41	109	276	310	168	0.82	29
ACL-6 soluciones	28	110	311	352	180	0.81	22
ACL-7 soluciones	21	117	352	400	200	0.76	15

**Paso 1: Probar la medición de invarianza longitudinal usando el análisis de transición de clases latentes LTA.**

En el paso anterior, las probabilidades de respuesta a los ítems no estaban restringidas, pues los análisis se hicieron independientemente en cada periodo de tiempo. Esto pueden causar ambigüedad al definir los estados latentes del próximo paso, porque las características de los estados latentes pueden explicarse por los ítems observados, así como por la varianza del

modelo de medición. Surge, entonces el concepto de invarianza, que significa que si un individuo se encuentra en un estado latente específico, la probabilidad de responder a un ítem no cambia con el tiempo. Esto implica que la función de respuesta al ítem es constante bajo cada estado latente a lo largo del tiempo. Cuando se habla de invarianza de las probabilidades de respuesta al ítem, se refiere a la idea de que las probabilidades de que un ítem (o pregunta) sea respondido de una cierta manera no deberían depender del momento en que se hace la observación, sino solo del estado latente en el que se encuentra el individuo. En un modelo de transición de clases latentes, se tiene un proceso en el que un individuo puede estar en diferentes estados latentes y la probabilidad de que responda a un ítem de una cierta manera está relacionada con su estado latente. La invarianza de las probabilidades de respuesta al ítem significa que esta probabilidad no cambia a lo largo del tiempo, o en diferentes momentos de la observación, si el estado latente del individuo es el mismo.

El objetivo de asumir invarianza a lo largo del tiempo es para tener una mejor y más fácil interpretación de los modelos de transición de clases latentes, para ayudar a estabilizar la estimación y por último para mejorar la identificabilidad del modelo. Otro de los objetivos de probar la medición de invarianza longitudinal es la de asegurar la medida de invarianza, de modo que las características de los estados latentes puedan ser explicadas por los ítems observados a lo largo del tiempo.

Entonces, lo que se hace en este trabajo, no es asumir la medición de invarianza sino probarla formalmente como se observa en la tabla (5.3).

Los resultados muestran, según la prueba de razón de verosimilitud y el valor-p observado, que tanto para el modelo de tres y cuatro soluciones se puede considerar que hay invarianza. Esto está también en concordancia con los valores más bajos de los índices AIC y BIC bajo invarianza de la Tabla 5.3.

Es importante aclarar que matemáticamente la restricción de invarianza de las probabilidades de respuesta al ítem en el algoritmo EM, se puede incorporar a este en (4.14) al fijar que los parámetros  $\rho_{mkt|\ell_t}$  no dependan de  $t$ .

Tabla 5.3: Resultados de la medida de invarianza longitudinal.

	<b>Medida de Invarianza</b>	<b>G2</b>	<b>AIC</b>	<b>BIC</b>	<b>DF</b>	<b>Dif. G</b>	<b>Dif. GL</b>	<b>Valor-p</b>
ACL3 soluciones	Sí	3824.65	3888.65	4050.99	262111			
	No	3774.91	3910.91	4255.89	262075	49.74	36	0.06349
ACL4 soluciones	Sí	3701.09	3803.09	4061.83	262092			
	No	3644.85	3842.85	4345.11	262044	56.24	48	0.19369

## Paso 2: Definir los estados latentes

Los pasos 0 y 1 sugirieron los modelos de tres y cuatro soluciones con invarianza de medida como modelos candidatos. Basándonos en las probabilidades de respuesta al ítem dentro de cada estado latente, podemos trazar entonces los perfiles de los modelos de tres y cuatro clases en la Figura 5.2. Según esta, seleccionaremos el modelo, con invarianza de medida, de cuatro clases latentes, los cuales creemos representan grupos significativos con respecto a las preocupaciones de los estudiantes sobre el acoso. Este gráfico nos permitiría también esbozar la siguiente interpretación de las clases:

- L1 = Estudiantes no preocupados por el Bullying
- L2 = Estudiantes preocupados por los rumores y en menor medida por el acoso verbal
- L3 = Estudiantes sumamente preocupados por todos los tipos de Bullying
- L4 = Estudiantes preocupados principalmente por un acoso relacionado a sus creencias y en menor medida a rumores.

Si bien al examinar los estadísticos de ajuste, alguien podría razonablemente argumentar que por parsimonia se debería preferir el modelo con tres clases; vale destacar que la parsimonia no es aquí necesariamente el objetivo del LTA. El objetivo del LTA, particularmente cuando se usa con carácter exploratorio, consiste en describir las clases latentes dentro de la población de observaciones en la data y la selección del modelo de cuatro soluciones proporciona una descripción más adecuada de los datos. Además, cuando examinamos qué clases cambian al pasar de un modelo de cuatro clases a un modelo de tres clases, vemos en las Figuras 5.2 y 5.3 que las dos clases situadas en el centro cambian, mientras que las clases situadas en los extremos de la distribución general de los datos permanecen casi inalterables.

Figura 5.2: Perfiles latentes para el modelo de 3 soluciones

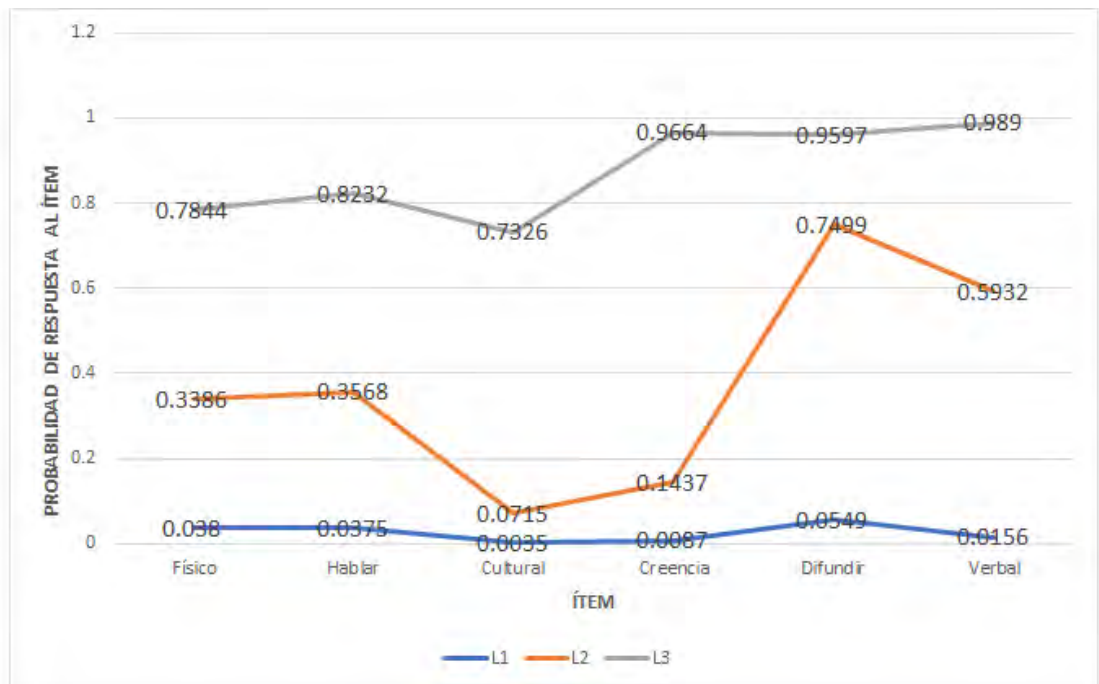
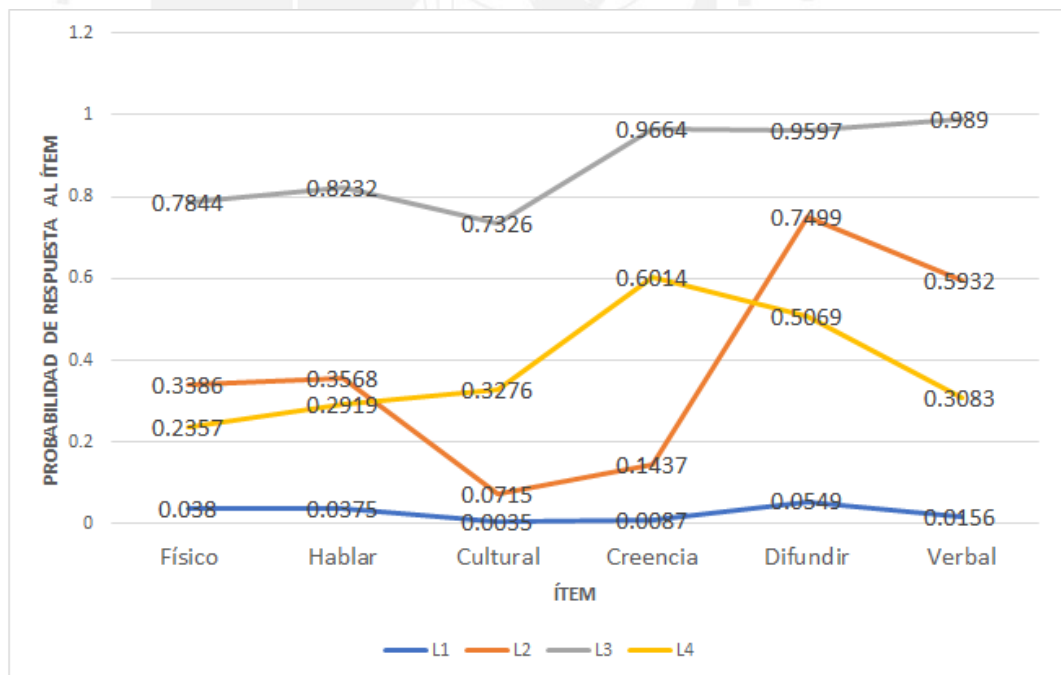


Figura 5.3: Perfiles latentes para el modelo de 4 soluciones



Dado que la muestra de nuestro análisis exploratorio proviene de una sola ciudad del oeste de Estados Unidos, es probable que las asociaciones entre estas sean mayores de lo que esperaríamos en el supuesto de un muestreo aleatorio simple. Con una muestra más amplia, la separación entre dos clases de colegios puede ser más diferente. Además, es importante seña-

lar que los estadísticos de ajuste no son pruebas de hipótesis estadísticas estrictas, sino guías para ayudar al investigador a determinar el número correcto de clases latentes. Los estadísticos de ajuste deben utilizarse junto con consideraciones teóricas. Dado que los estadísticos de ajuste apoyan al modelo de cuatro soluciones, cada una de las clases de este modelo tiene sentido desde el punto de vista teórico, el objetivo de este estudio es exploratorio, y no se ha realizado ningún análisis teórico. Existe la posibilidad de asociaciones confusas en los datos y se consideró que el modelo de cuatro clases era el preferido para avanzar en el análisis. Esta enumeración abordó la primera pregunta de investigación: ¿Existen subgrupos cualitativamente distintos de alumnos que muestren patrones particulares de preocupación en relación al acoso escolar?

La Tabla 5.4, por otro lado, resume la prevalencia del estado latente a lo largo de los tres periodos de tiempo para los estudiantes preocupados por el acoso escolar. Comenzando con las estimaciones podemos examinar las tasas de prevalencia en los tres momentos. Luego las Tabla 5.5 y 5.6 resumen la matriz de transición del Tiempo 1 al Tiempo 2 y la matriz de transición del Tiempo 2 al Tiempo 3.

Tabla 5.4: Estimaciones de las prevalencias en los estados latentes en cada punto en el tiempo

<b>PREVALENCIA DE LAS CLASES</b>					
Todas	<b>Tiempo</b>	<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>L4</b>
	<b>Tiempo 1</b>	0.5691	0.2395	0.0804	0.1109
	<b>Tiempo 2</b>	0.5961	0.2166	0.0719	0.1155
	<b>Tiempo 3</b>	0.6695	0.1625	0.0628	0.1052

La tasa de prevalencia indica que la probabilidad de pertenecer al grupo preocupado por los rumores y en menor medida por el acoso verbal (L2) es relativamente alta: 0.24 en el tiempo 1, 0.22 en el tiempo 2 y 0.16 en el tiempo 3, mientras que las probabilidades de pertenecer al grupo de preocupados principalmente por un acoso relacionado a sus creencias y en menor medida a rumores (L4) es de 0.11 en el tiempo 1, 0.12 en el tiempo 2 y 0.11 en el tiempo 3. Asimismo, las tasas de prevalencia indican que las probabilidades de pertenecer al grupo de los más preocupados por todos los tipos de Bullying (L3) eran relativamente bajas, 0.08 en el tiempo 1, 0.07 en el tiempo 2 y 0.06 en el tiempo 3, mientras que las probabilidades de pertenecer al grupo de los no preocupados por el Bullying (L1) son las más altas de los tres tiempos.

Esta prevalencia en los estados latentes responde a la segunda pregunta de investigación: ¿Cuál es la probabilidad de que un individuo se encuentre en un estado latente diferente?

Se observa, por ejemplo que del tiempo 1 al tiempo 2, el grupo de alumnos sumamente preocupados por el Bullying (L3) tiende a permanecer en este grupo o a pasar al preocupado; sin embargo, este patrón es ligeramente distinto entre el tiempo 2 y 3, pues aquí hay una mucho mayor probabilidad de que los alumnos que pertenecen a este grupo permanezcan allí. Se destaca también una alta probabilidad de transición de los alumnos preocupados por un acoso sus creencias y en menor medida a rumores (L4) al grupo de alumnos despreocupados

por el Bullying, entre los tiempos 2 y 3.

Tabla 5.5: Estimación de la matriz de transición del tiempo 1 al tiempo 2

<b>Estimación de las probabilidades de transición del tiempo 1 al 2</b>				
	<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>L4</b>
<b>L1</b>	0.8252	0.0925	0.0208	0.0615
<b>L2</b>	0.2780	0.5734	0.0893	0.0594
<b>L3</b>	0.2707	0.3309	0.3785	0.0198
<b>L4</b>	0.3432	0.0000	0.0739	0.5829

Tabla 5.6: Estimación de la matriz de transición del tiempo 2 al tiempo 3

<b>Estimación de las probabilidades de transición del tiempo 2 al 3</b>				
	<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>L4</b>
<b>L1</b>	0.8518	0.0703	0.0143	0.0637
<b>L2</b>	0.3666	0.5473	0.0861	0.0000
<b>L3</b>	0.2639	0.0287	0.4959	0.2114
<b>L4</b>	0.5492	0.0000	0.0000	0.4508

### **Paso 3: Evaluación de los estados latentes e invarianza en las probabilidades de transición.**

Se debe probar la invarianza en las probabilidades de transición para confirmar su necesidad. Esta se prueba con el modelo de 4 soluciones y los resultados se resumen en la Tabla 5.7. Los resultados indican que tanto el AIC como el BIC favorecen al modelo de invarianza en las probabilidades de transición mientras que la prueba de LRDT favorece al modelo de probabilidades de transición libres (ver tabla 5.8).

A pesar de la similitud entre las matrices de transición del tiempo 1 al 2 y del tiempo 2 al 3 vistas en las Tablas 5.5 y 5.6, se encontró probabilidades grandes y significativas en estas matrices de transición sugiriendo diferentes patrones de transición en diferentes periodos de tiempo. Por ejemplo, diferentes estimaciones de transición del estado latente 3 (Sumamente preocupados por todos los tipos de Bullying) al estado latente 2 (Preocupados por los rumores y en menor medida por el acoso verbal) en dos periodos de tiempos diferentes (0.33 vs 0.03). También, ocurre lo mismo del estado latente 3 (Sumamente preocupados por todos los tipos de Bullying) al estado latente 4 (Preocupados principalmente por un acoso relacionado a sus creencias y en menor medida a rumores) (0.02 vs 0.21).

Es por esto que se eligió el modelo de 4 soluciones con probabilidades de transiciones libres. Esta confirmación del modelo en las matrices de transición responde a la tercera pregunta de investigación planteada en esta aplicación: ¿Existe el cambio entre los estados latentes a lo largo del tiempo?.

Tabla 5.7: Resultados de la invarianza en las probabilidades de transición

Solución 4	G2	AIC	BIC	DF	Diff.G2	Diff.DF	Valor-p
Modelo 2 (Invarianza)	3722.93	3800.93	3998.79	262104			
Modelo 1	3701.09	3803.09	4061.83	262092	21.84	12	0.039

## 5.5. Acoso escolar y su relación con otras variables

Con el fin de estudiar con más detalle los diferentes tipos de bullying, se realizó un análisis de asociación con cuatro variables complementarias sociodemográficas como son el grado de estudios, el género, el idioma y la raza asignada por la escuela recogidos en la encuesta estudiantil PRBm. Para ello, presentamos en primer lugar, las distribuciones de frecuencias de estas variables así como también, la clasificación de frecuencias de las clases latentes según los periodos en estudio. Se excluyen aquí los casos perdidos.

Figura 5.4: Distribución de alumnos en los diferentes grados para el estudio de Bullying

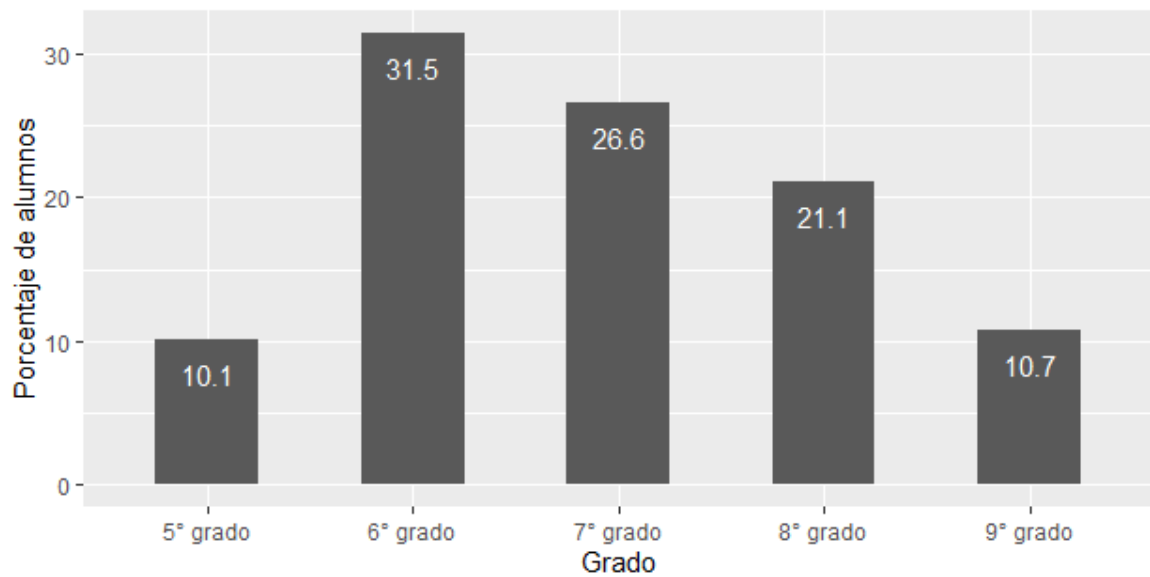


Figura 5.5: Distribución de los alumnos según género

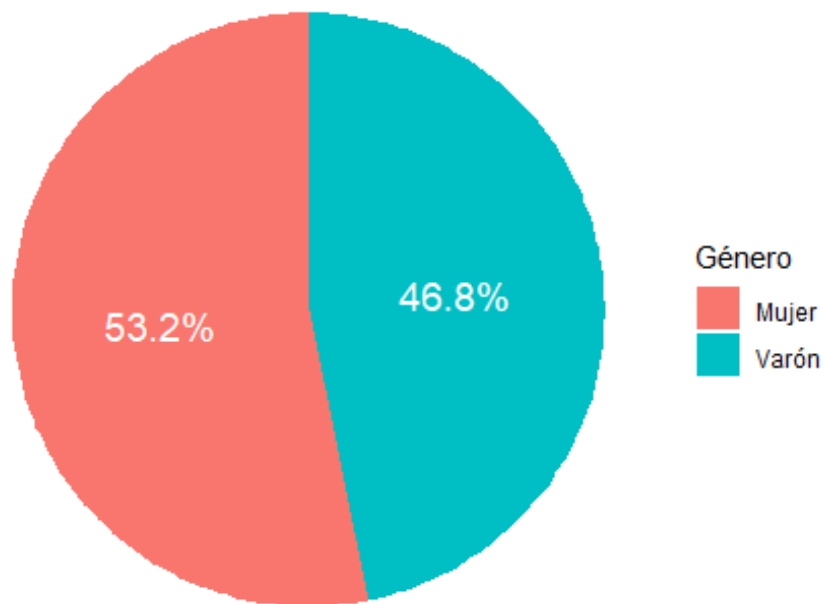


Figura 5.6: Distribución de los alumnos según si el idioma Inglés es (Sí) o no (No) nativo

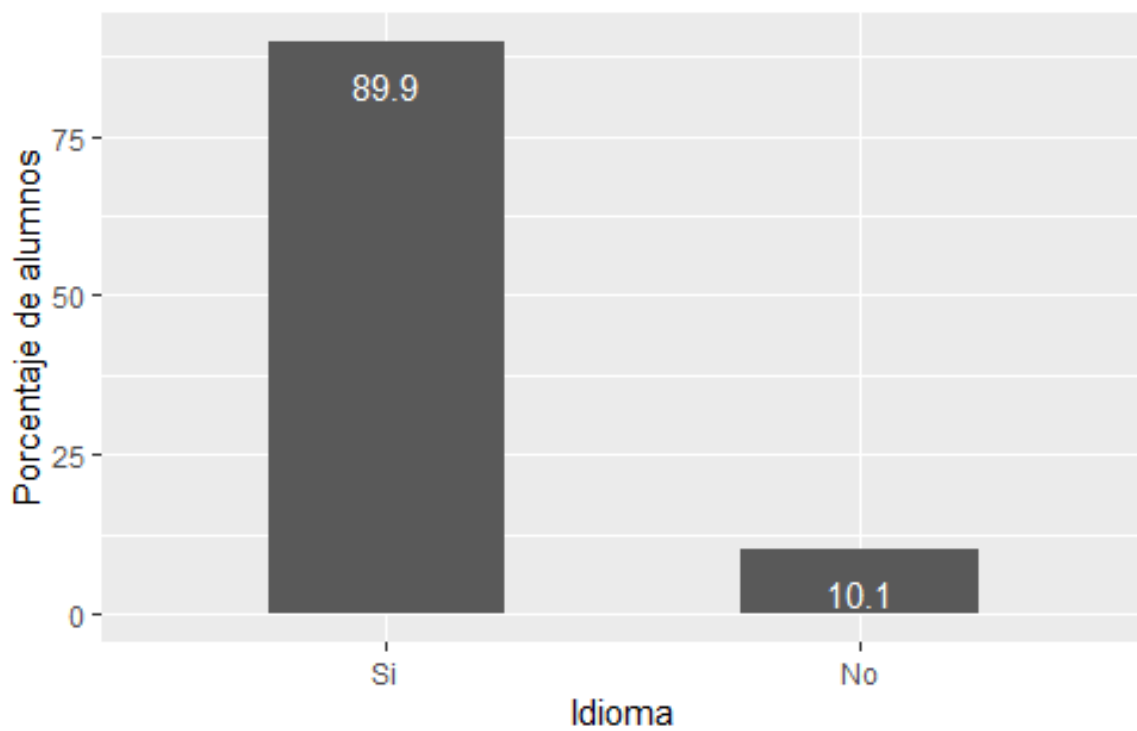


Figura 5.7: Distribución de los alumnos según Raza reportada por la escuela

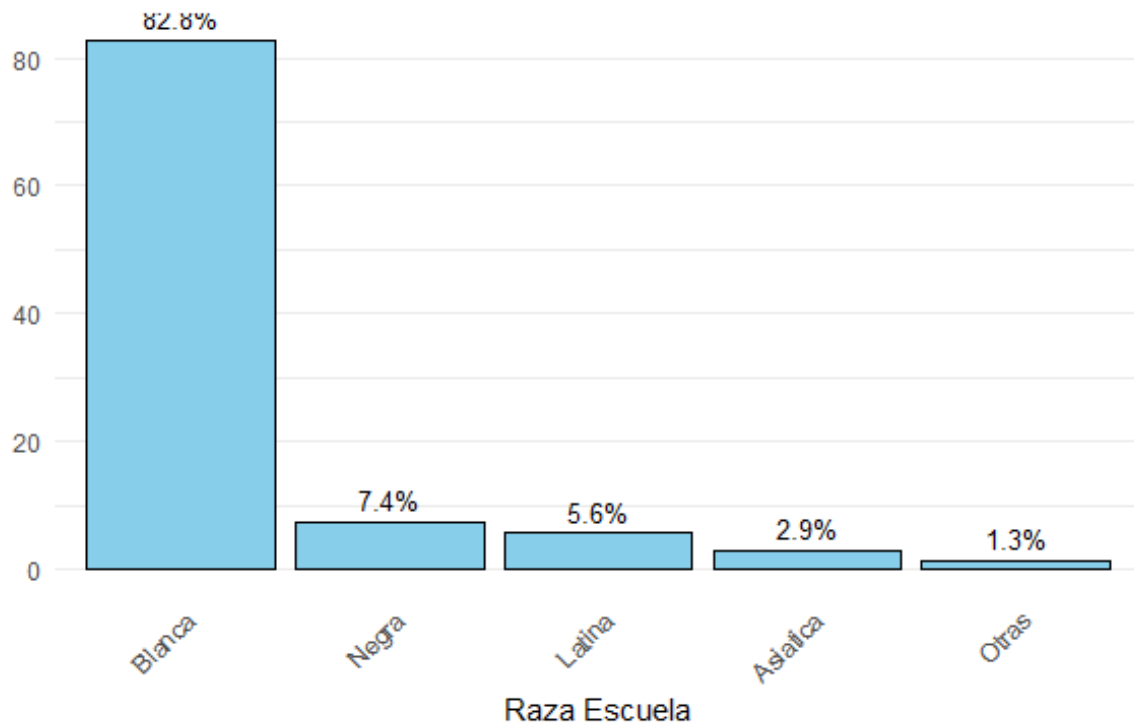


Tabla 5.8: Distribución porcentual de las clases latentes según periodos de estudio

Tiempos	Tiempo 1	Tiempo 2	Tiempo 3
Clases latentes	Porcentaje	Porcentaje	Porcentaje
Nada preocupados por el bullying	23.0 (L1.1)	21.4 (L2.1)	17.0 (L3.1)
Preocupados por los rumores y en menor medida por el acoso verbal	7.9 (L1.2)	6.7 (L2.2)	5.8 (L3.2)
Sumamente preocupados por todos los tipos de bullying	59.5 (L1.3)	62.6 (L2.3)	69.8 (L3.3)
Preocupados principalmente por un acoso relacionado a sus creencias y en menor medida a rumores	9.7 (L1.4)	9.3 (L2.4)	7.3 (L3.4)
Total	100.0	100.0	100.0

Seguidamente se realizaron, para cada periodo de tiempo, las pruebas de asociación Chi-cuadrado entre la clasificación latente y las variables complementarias.

Los resultados de las pruebas de independencia Chi-cuadrado en que la hipótesis nula indica que la clasificación latente es independiente de la covariable de interés, se muestran en las siguientes tablas, para cada periodo de tiempo.

Tabla 5.9: Prueba de asociación Chi cuadrado entre la clasificación latente y el grado de estudios

(Clasificación latente vs Grado de estudios)	Tiempo 1	Tiempo 2	Tiempo 3
Estadística $\chi^2$	1896	1784.6	1608.6
Grados de libertad	12	12	12
Valor-p	2.2e-16	2.2e-16	2.2e-16

Se aprecia que existe asociación significativa entre la clasificación latente con los cinco grados de estudios en los 3 periodos analizados.

Tabla 5.10: Prueba de asociación Chi cuadrado entre la clasificación latente y el género

(Clasificación latente vs Género del estudiante)	Tiempo 1	Tiempo 2	Tiempo 3
Estadística $\chi^2$	0.85236	0.9079	2.8485
Grados de libertad	3	3	3
Valor-p	0.8369	0.8235	0.4156

Se aprecia que no existe asociación significativa entre la clasificación latente y el género de los estudiantes en ningún periodo analizado.

Tabla 5.11: Prueba de asociación Chi cuadrado entre la clasificación latente y el idioma

(Clasificación latente vs Idioma)	Tiempo 1	Tiempo 2	Tiempo 3
Estadística $\chi^2$	15.929	14.168	5.1374
Grados de libertad	3	3	3
Valor-p	0.001173	0.002686	0.162

Se aprecia que existe asociación significativa entre las clasificación latente con el idioma en los dos primeros periodos, pero no en el tercer periodo.

Tabla 5.12: Prueba de asociación Chi cuadrado entre la clasificación latente y la raza

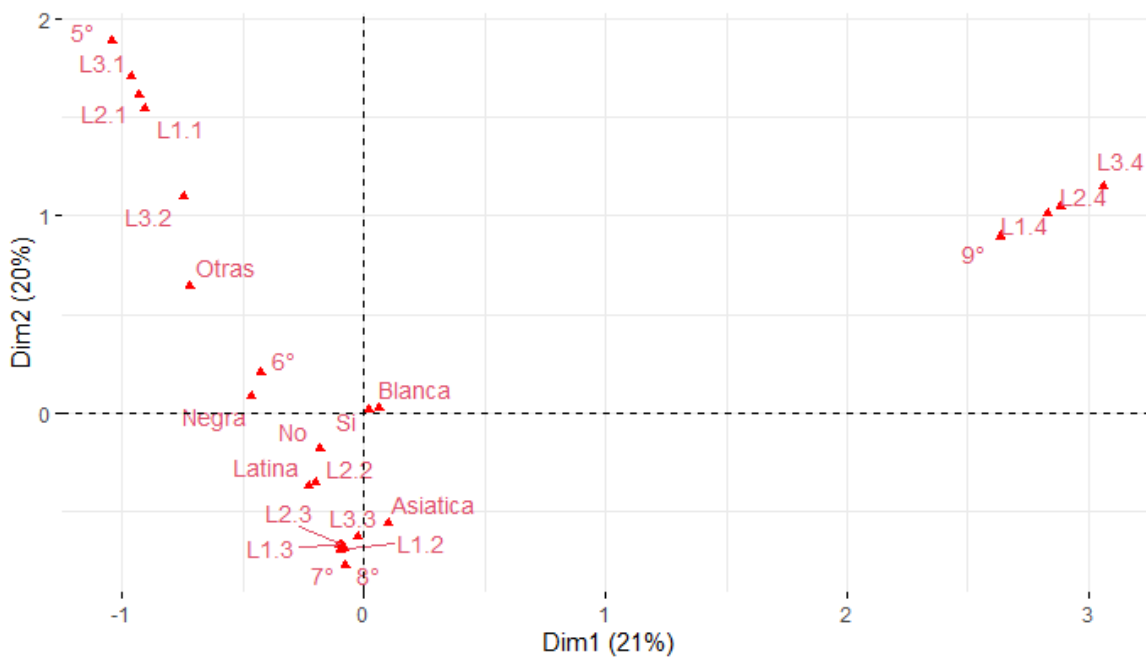
(Clasificación latente vs Raza)	Tiempo 1	Tiempo 2	Tiempo 3
Estadística $\chi^2$	45.234	43.439	30.761
Valor-p	0.0004998	0.0004998	0.003998

Se aprecia que existe asociación significativa entre la clasificación latente con las diferentes clasificaciones de razas en los 3 periodos analizados. Vale aquí aclarar que no se reportan los grados de libertad pues los valores p fueron calculados por el Método de Montecarlo bajo 2000 réplicas.

Finalmente, para describir mejor los resultados de asociación se realizó un análisis de correspondencia múltiple. El análisis de correspondencia múltiple es una técnica multivariante de reducción de la dimensionalidad para variables categóricas, que permite representar las distintas categorías de las distintas variables en un mismo espacio euclídeo bi-dimensional o tri-dimensional (biplot) de tal manera que la cercanía entre distintas categorías nos indiquen una alta frecuencia conjunta de casos para estas categorías. En las dimensiones de un biplot se pueden también representar simultáneamente las categorías de las variables analizadas y los individuos involucrados, aunque nuestro objetivo será estudiar solo la relación de nuestra clasificación latente con las demás variables. En su interpretación, la proximidad entre dos categorías indica que se presenta una alta frecuencia de alumnos que comparten ambas categorías, lo que a su vez indica una posible asociación entre las variables de las categorías involucradas. Además, mientras mas alejadas estén las categorías del origen del biplot, más atípicas serán estas.

En el biplot resultante, que se muestra en la Figura 5.8 y en donde no consideramos la variable Género, por no ser significativa, se utilizó para su creación las librerías FactoMineR y factorextra de R.

Figura 5.8: Análisis de correspondencias múltiples para la variables de estudio en el problema de Bullying



Si bien este biplot lo podríamos también haber evaluado para cada periodo de estudio, decidimos considerar simultáneamente los 3 periodos dada la uniformidad temporal de los resultados. En general, se aprecia claramente una menor preocupación por el acoso escolar en los alumnos del menor de los grados de estudio (5to grado) y una mayor preocupación por acoso sobre creencias y en menor medida rumores, en los alumnos del último grado de estudio (9no grado). Los casos críticos de víctimas de acoso escolar; es decir, los del grupo que tienen preocupación por todos los tipos de acoso, parecen darse con mayor frecuencia más en el 7mo y 8vo grado y principalmente se centra en alumnos de raza Latina o Asiática que no tienen al idioma inglés como lengua nativa.

## Capítulo 6

# Conclusiones y sugerencias

En este trabajo de tesis se ha desarrollado el modelo de transición de clases latentes LTA buscando estimar sus parámetros bajo un enfoque clásico y el uso del algoritmo EM. El objetivo principal de la investigación fue describir este modelo, estudiar sus propiedades, implementar su estimación y aplicarlo a un conjunto de datos reales. A diferencia de la poca literatura sobre el algoritmo EM en estos modelos, se explicitó en esta tesis cómo es que actúa este algoritmo en la estimación de los parámetros de interés. Se mostró también su aplicación, a través del software Mplus, a una base de datos de acoso escolar (Bullying). Si bien esta aplicación fue propuesta por Ryoo et al. (2018) para una muestra de 1180 estudiantes que respondieron al cuestionario sobre intimidación PRBm, aquí se complementó este estudio con un análisis de correspondencias múltiples para describir algunas relaciones que se pudieran establecer entre la clasificación latente del acoso y otras variables de interés disponibles en la base de datos.

Nosotros no encontramos problemas en la estimación de los parámetros del modelo LTA para el problema de aplicación sobre acoso escolar; sin embargo, se han reportado en la literatura algunos problemas con la estimación clásica. Ello se debe a que si la muestra es relativamente pequeña, la inferencia por máxima verosimilitud suele tener problemas en la estimación de algunos parámetros que se encuentran en los límites del espacio paramétrico (es decir, cerca de 0 o 1). Las probabilidades de respuesta al ítem cercanas a 0 o a 1, son altamente preferibles desde una perspectiva de medición, y cuando algunos de estos parámetros son estimados en el límite, es imposible obtener los errores estándar de estimación. Para solucionar ello, sería interesante optar, como lo proponen (Chung et al., 2007), por incorporar distribuciones de Dirichlet u otras a priori y encontrar las estimaciones por el máximo a posteriori. Otra posibilidad a estudiar es optar por un enfoque netamente Bayesiano. Queda pendiente un estudio de comparación de estos métodos de estimación.

Si bien, hemos estudiado la relación de la clasificación latente del acoso escolar con otras

covariables, mediante un análisis de correspondencias, sería también interesante extender este modelo, al igual que se lo hace con el de clases latentes, a un modelo de regresión. Por ejemplo, las asociaciones que encontramos con algunas covariables como el idioma nativo, la raza y el grado, fueron solo en la prevalencia del acoso sobre cada periodo. Un modelo de regresión aquí nos podría dar no solo información de estas prevalencias, sino también de cómo estas covariables afectan a las probabilidades de transición del modelo.

En relación a la aplicación, hemos aquí reducido por simplicidad la escala original de Likert de 4 puntuaciones a 2 para las variables manifiestas sobre el acoso del cuestionario PRBm. Sería interesante replicar este estudio con la escala de respuesta original y comparar si esta dicotomización conlleva o no a una pérdida significativa de información.

Desde el punto de vista computacional, sería también de mucho interés el implementar la estimación clásica de este modelo en R. Hemos revisado extensamente la literatura y hasta donde conocemos no encontramos ninguna implementación en R. Por esta razón, y dada la complejidad del modelo, optamos por realizarla en el software estadístico Mplus. Otra alternativa es el uso del paquete SAS.

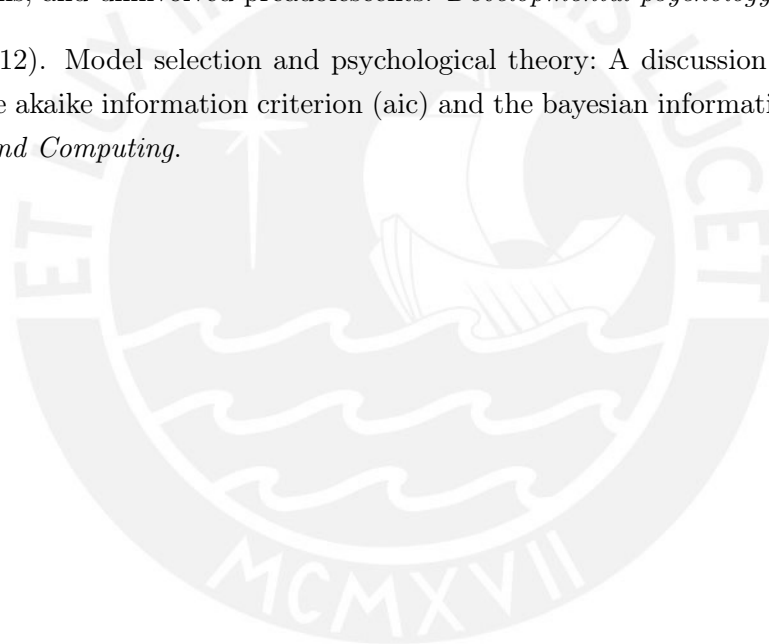
Finalmente, los LTA son en principio un modelo base que puede admitir múltiples extensiones. Recientemente, por ejemplo, Muthén & Asparouhov (2022) han propuesto un modelo LTA con intercepto aleatorio, que según estos autores es menos restrictivo y tiende a ajustar mejor a la data. Por otro lado, dado que los LTA incorporan probabilidades de respuesta al ítem generales, estas se pueden particularizar con modelos de diagnóstico cognitivo u otros, tales como los que se proponen en los artículos de Kaya & Leite (2017) y Li et al. (2017)

# Bibliografía

- Astor, R. A., Meyer, H. A., & Pitner, R. (2001). Elementary and middle school students' perceptions of safety: An examination of violence-prone school sub-contexts. *Elementary School Journal*, 101(5), 511–528.
- Bachman, R., Gunter, W. D., & Bakken, N. W. (2011). Predicting feelings of school safety for lower, middle, and upper school students: a gender specific analysis. *Applied Psychology in Criminal Justice*, 7, 59–76.
- Boscardin, C. K., Muthén, B., Francis, D. J., & Baker, E. L. (2008). Early identification of reading difficulties using heterogeneous developmental trajectories. *Journal of Educational Psychology*, 100(1), 192.
- Bozdogan, H. (1987). Model selection and akaike's information criterion (aic): The general theory and its analytical extensions. *Psychometrika*, 52(3), 345–370.
- Bukowski, W. M., Sippola, L. K., & Newcomb, A. F. (2000). Variations in patterns of attraction to same- and other-sex peers during early adolescence. *Developmental Psychology*, 36(2), 147–154.
- Chung, H., Park, Y., & Lanza, S. T. (2005). Latent transition analysis with covariates: Pubertal timing and substance use behaviours in adolescent females. *Statistics in Medicine*, 24(18), 2895–2910.
- Chung, H., Walls, T. A., & Park, Y. (2007). *A latent transition model with logistic regression*. Michigan, USA: Psychometrika.
- Collins, L., Fidler, P., Wugalter, S., & Long, J. (1993). Goodness-of-fit testing for latent class models. *Multivariate Behavioral Research - MULTIVARIATE BEHAV RES*, 28, 375–389.
- Collins, L. M. & Lanza, S. T. (2010). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. New York, NY: John Wiley & Sons.
- Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences*, 18(3), 329–337.

- D'Antona, R., Kevorkian, M., & Russom, A. (2010). Sexting, texting, cyberbullying and keeping youth safe online. *Journal of Social Sciences*, 6, 521–526.
- Hughes, M. R., Gaines, J. S., & Pryor, D. W. (2015). Staying away from school: adolescents who miss school due to feeling unsafe. *Youth Violence and Juvenile Justice*, 13, 270–290.
- Kaya, Y. & Leite, W. L. (2017). *Assessing change in latent skills across time with longitudinal cognitive diagnosis modeling: An evaluation of model performance*. Turkey and Gainesville, FL, USA: Educational and psychological measurement.
- Konishi, C., Hymel, S., Zumbo, B. D., Li, Z., Taki, M., Slee, P., & et al. (2009). Investigating the comparability of a self-report measure of childhood bullying across countries. *Can. J. School Psychol.*, 24, 82–93.
- Lanza, S., Flaherty, B., & Collins, L. (2003). *Latent Class and Latent Transition Analysis*.
- Lanza, S. T. & Cooper, B. R. (2016). Latent class analysis for developmental research. *Child Development Perspectives*, 10, 59–64.
- Lazarsfeld, P. F. & Henry, N. W. (1968). *Latent Structure Analysis*. Boston: Houghton Mifflin.
- Li, F., Cohen, A., Bottge, B., & Templin, J. (2017). *A latent transition analysis model for assessing change in cognitive skills*. Texas TX, Athens GA, Lexington KY, Lawrence KS, USA: Educational and Psychological Measurement.
- Louis, T. A. (1982). Finding the observed information matrix when using the em algorithm. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 44(2), 226–233.
- Lubke, G. & Neale, M. C. (2006). Distinguishing between latent classes and continuous factors: Resolution by maximum likelihood? *Multivariate Behavioral Research*, 41(4), 499–532.
- Marley, D. (2008). Pupils worry over exams, the future and their bodies. *Times Educational Supplement*, 4812, 12.
- Muthén, B. & Asparouhov, T. (2022). Latent transition analysis with random intercepts (ri-lta). *Psychol Methods*, 27(1), 1–16. Epub 2020 Nov 23.
- Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from latent classes: A demonstration of different approaches to include distal outcomes in mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(6), 967–985.
- Oliver, R., Hoover, J. H., & Hazler, R. (1994). The perceived roles of bullying in small-town midwestern schools. *Journal of Counseling & Development*, 72(4), 416–419.
- Parzen, E. (1979). *Teoría moderna de probabilidades y sus aplicaciones*. Limusa.
- Peel, D. & McLachlan, G. J. (2000). Robust mixture modelling using the t distribution. *Statistics and computing*, 10(4), 339–348.

- Pellegrini, A. D. (2002). Bullying, victimization, and sexual harassment during the transition to middle school. *Educational Psychology, 37*(3), 151–163.
- Robers, S., Kemp, J., Truman, J., & Snyder, T. D. (2013). Indicators of school crime and safety: 2012 (nces 2013-036/ncj 241446).
- Ryoo, J. H., Wang, C., Swearer, S. M., Hull, M., & Shi, D. (2018). Longitudinal model building using latent transition analysis: An example using school bullying data. *Frontiers in Psychology, 9*.
- Trentacosta, C. J., Criss, M. M., Shaw, D. S., Lacourse, E., Hyde, L. W., & Dishion, T. J. (2011). Antecedents and outcomes of joint trajectories of mother–son conflict and warmth during middle childhood and adolescence. *Child development, 82*(5), 1676–1690.
- Veenstra, R., Lindenberg, S., Oldehinkel, A. J., De Winter, A. F., Verhulst, F. C., & Ormel, J. (2005). Bullying and victimization in elementary schools: a comparison of bullies, victims, bully/victims, and uninvolved preadolescents. *Developmental psychology, 41*(4), 672–682.
- Vrieze, S. (2012). Model selection and psychological theory: A discussion of the differences between the akaike information criterion (aic) and the bayesian information criterion (bic). *Statistics and Computing*.



# Apéndice A

## Anexos

### A.1. Código Mplus para la estimación del modelo LTA

Title: Step 0: Diagnose and Explore Cross-sectional Data using LCA

Data: File = bullyingT.dat;

Variable: Names = core2711 core2721 core2731 core2741 core2751 core2761  
core2712 core2722 core2732 core2742 core2752 core2762 core2713 core2723  
core2733 core2743 core2753 core2763 Grade Gender Language  
RaceSch RaceStu;

! Time 1:

Usevar = core2711 core2721 core2731 core2741 core2751 core2761;

Categorical = core2711 core2721 core2731 core2741 core2751 core2761;

Missing = all(-999);

Classes = C1(2); ! Need to change from C1(2) to C1(7)

! Time 2:

! Usevar = core2712 core2722 core2732 core2742 core2752 core2762;

! Categorical = core2712 core2722 core2732 core2742 core2752 core2762;

! Missing = all(-999);

! Classes = C2(7); ! Need to change from C2(2) to C2(7)

! Time 3:

```

! Usevar = core2713 core2723 core2733 core2743 core2753 core2763;
! Categorical = core2713 core2723 core2733 core2743 core2753 core2763;

! Missing = all(-999);
! Classes = C3(7); ! Need to change from C3(2) to C3(7)

Analysis:      Type = mixture;

Plot:          type is plot3;
! Time 1:
series is core2711 (1) core2721 (2) core2731 (3)
core2741 (4) core2751 (5) core2761 (6);

! Time 2:
! series is core2712 (1) core2722 (2) core2732 (3)
! core2742 (4) core2752 (5) core2762 (6);
! Time 3:
! series is core2713 (1) core2723 (2) core2733 (3)
!core2743 (4) core2753 (5) core2763 (6);

Output:        Tech11 Tech14;

Title:         Step 1: Test Longitudinal Measurement Invariance using LTA

Data:          File = bullyingT.dat;

Variable: Names = core2711 core2721 core2731 core2741 core2751 core2761
core2712 core2722 core2732 core2742 core2752 core2762
core2713 core2723 core2733 core2743 core2753 core2763
Grade Gender Language RaceSch RaceStu;

! Time 1:
Usevar = core2711 core2721 core2731 core2741 core2751 core2761
core2712 core2722 core2732 core2742 core2752 core2762
core2713 core2723 core2733 core2743 core2753 core2763;
Categorical = core2711 core2721 core2731 core2741 core2751 core2761
core2712 core2722 core2732 core2742 core2752 core2762
core2713 core2723 core2733 core2743 core2753 core2763;

Missing = all(-999);
! Classes = C1(3) C2(3) C3(3); ! for three latent statuses (3LS) model
Classes = C1(4) C2(4) C3(4); ! for 4LS model

```

Analysis:           Type = mixture;

Model:             %OVERALL%  
                    C2 on C1;  
                    C3 on C2;

! Model needs to be set up with the number of assessment

Model C1:           %C1#1%

! For non-measurement invariance

! [core2711\$1 core2721\$1 core2731\$1 core2741\$1 core2751\$1 core2761\$1];

! For measurement invariance

[core2711\$1] (1);

[core2721\$1] (2);

[core2731\$1] (3);

[core2741\$1] (4);

[core2751\$1] (5);

[core2761\$1] (6);

%C1#2%

! [core2711\$1 core2721\$1 core2731\$1 core2741\$1 core2751\$1 core2761\$1];

[core2711\$1] (7);

[core2721\$1] (8);

[core2731\$1] (9);

[core2741\$1] (10);

[core2751\$1] (11);

[core2761\$1] (12);

%C1#3%

! [core2711\$1 core2721\$1 core2731\$1 core2741\$1 core2751\$1 core2761\$1];

[core2711\$1] (13);

[core2721\$1] (14);

[core2731\$1] (15);

[core2741\$1] (16);

[core2751\$1] (17);

[core2761\$1] (18);

! Remove %C1#4% when testing three latent statuses (3LS) model

%C1#4%

! [core2711\$1 core2721\$1 core2731\$1 core2741\$1 core2751\$1 core2761\$1];

[core2711\$1] (19);  
[core2721\$1] (20);  
[core2731\$1] (21);  
[core2741\$1] (22);  
[core2751\$1] (23);  
[core2761\$1] (24);

Model C2:           %C2#1%

! [core2712\$1 core2722\$1 core2732\$1 core2742\$1 core2752\$1 core2762\$1];  
[core2712\$1] (1);  
[core2722\$1] (2);  
[core2732\$1] (3);  
[core2742\$1] (4);  
[core2752\$1] (5);  
[core2762\$1] (6);

%C2#2%

! [core2712\$1 core2722\$1 core2732\$1 core2742\$1 core2752\$1 core2762\$1];  
[core2712\$1] (7);  
[core2722\$1] (8);  
[core2732\$1] (9);  
[core2742\$1] (10);  
[core2752\$1] (11);  
[core2762\$1] (12);

%C2#3%

! [core2712\$1 core2722\$1 core2732\$1 core2742\$1 core2752\$1 core2762\$1];  
[core2712\$1] (13);  
[core2722\$1] (14);  
[core2732\$1] (15);  
[core2742\$1] (16);  
[core2752\$1] (17);  
[core2762\$1] (18);

%C2#4%

! [core2712\$1 core2722\$1 core2732\$1 core2742\$1 core2752\$1 core2762\$1];  
[core2712\$1] (19);  
[core2722\$1] (20);  
[core2732\$1] (21);  
[core2742\$1] (22);  
[core2752\$1] (23);  
[core2762\$1] (24);

Model C3:            %C3#1%  
! [core2713\$1 core2723\$1 core2733\$1 core2743\$1 core2753\$1 core2763\$1];  
[core2713\$1] (1);  
[core2723\$1] (2);  
[core2733\$1] (3);  
[core2743\$1] (4);  
[core2753\$1] (5);  
[core2763\$1] (6);

%C3#2%  
! [core2713\$1 core2723\$1 core2733\$1 core2743\$1 core2753\$1 core2763\$1];  
[core2713\$1] (7);  
[core2723\$1] (8);  
[core2733\$1] (9);  
[core2743\$1] (10);  
[core2753\$1] (11);  
[core2763\$1] (12);

%C3#3%  
! [core2713\$1 core2723\$1 core2733\$1 core2743\$1 core2753\$1 core2763\$1];  
[core2713\$1] (13);  
[core2723\$1] (14);  
[core2733\$1] (15);  
[core2743\$1] (16);  
[core2753\$1] (17);  
[core2763\$1] (18);

%C3#4%  
! [core2713\$1 core2723\$1 core2733\$1 core2743\$1 core2753\$1 core2763\$1];  
[core2713\$1] (19);  
[core2723\$1] (20);  
[core2733\$1] (21);  
[core2743\$1] (22);  
[core2753\$1] (23);  
[core2763\$1] (24);

Output: Tech1 Tech8 Tech15;

Title: Step 2: Define Latent Statuses

Data: File = bullyingT.dat;

```
Variable: Names = core2711 core2721 core2731 core2741 core2751 core2761
core2712 core2722 core2732 core2742 core2752 core2762
core2713 core2723 core2733 core2743 core2753 core2763
Grade Gender Language RaceSch RaceStu;
```

```
Usevar = core2711 core2721 core2731 core2741 core2751 core2761
core2712 core2722 core2732 core2742 core2752 core2762
core2713 core2723 core2733 core2743 core2753 core2763;
Categorical = core2711 core2721 core2731 core2741 core2751 core2761
core2712 core2722 core2732 core2742 core2752 core2762
core2713 core2723 core2733 core2743 core2753 core2763;
```

```
Missing = all(-999);
Classes = C1(3) C2(3) C3(3); ! for three latent statuses (3LS) model
!Classes = C1(4) C2(4) C3(4); ! for 4LS model
Analysis: Type = mixture;
Model: %OVERALL%
```

```
C2 on C1;
C3 on C2;
```

```
! Model needs to be set up with the number of assessment
Model C1: %C1#1%
```

```
! For measurement invariance
```

```
[core2711$1] (1);
[core2721$1] (2);
[core2731$1] (3);
[core2741$1] (4);
[core2751$1] (5);
[core2761$1] (6);
%C1#2%
[core2711$1] (7);
[core2721$1] (8);
[core2731$1] (9);
[core2741$1] (10);
[core2751$1] (11);
[core2761$1] (12);
```

```
%C1#3%
[core2711$1] (13);
```

[core2721\$1] (14);  
[core2731\$1] (15);  
[core2741\$1] (16);  
[core2751\$1] (17);  
[core2761\$1] (18);

!Remove %C1#4% when testing three  
latent statuses (3LS) model  
!%C1#4%

![core2711\$1] (19);  
![core2721\$1] (20);  
![core2731\$1] (21);  
![core2741\$1] (22);  
![core2751\$1] (23);  
![core2761\$1] (24);

Model C2: %C2#1%

[core2712\$1] (1);  
[core2722\$1] (2);  
[core2732\$1] (3);  
[core2742\$1] (4);  
[core2752\$1] (5);  
[core2762\$1] (6);

%C2#2%

[core2712\$1] (7);  
[core2722\$1] (8);  
[core2732\$1] (9);  
[core2742\$1] (10);  
[core2752\$1] (11);  
[core2762\$1] (12);

%C2#3%

[core2712\$1] (13);  
[core2722\$1] (14);  
[core2732\$1] (15);  
[core2742\$1] (16);  
[core2752\$1] (17);  
[core2762\$1] (18);

! Remove %C1#4% when testing three latent statuses (3LS)  
model

!%C2#4%

![core2712\$1] (19);

![core2722\$1] (20);  
![core2732\$1] (21);  
![core2742\$1] (22);  
![core2752\$1] (23);  
![core2762\$1] (24);

Model C3:           %C3#1%

[core2713\$1] (1);  
[core2723\$1] (2);  
[core2733\$1] (3);  
[core2743\$1] (4);  
[core2753\$1] (5);  
[core2763\$1] (6);

%C3#2%

[core2713\$1] (7);  
[core2723\$1] (8);  
[core2733\$1] (9);  
[core2743\$1] (10);  
[core2753\$1] (11);  
[core2763\$1] (12);

%C3#3%

[core2713\$1] (13);  
[core2723\$1] (14);  
[core2733\$1] (15);  
[core2743\$1] (16);  
[core2753\$1] (17);  
[core2763\$1] (18);

! Remove %C1#4% when testing three latent statuses (3LS) model

!%C3#4%

![core2713\$1] (19);  
![core2723\$1] (20);  
![core2733\$1] (21);  
![core2743\$1] (22);  
![core2753\$1] (23);  
![core2763\$1] (24);

Output: Tech1 Tech8 Tech15;

Title: Step 3: Test Latent Statuses and transition probability invariance

Data: File = bullyingT.dat;

Variable: Names = core2711 core2721 core2731 core2741 core2751 core2761  
core2712 core2722 core2732 core2742 core2752 core2762  
core2713 core2723 core2733 core2743 core2753 core2763  
Grade Gender Language RaceSch RaceStu;

Usevar = core2711 core2721 core2731 core2741 core2751 core2761  
core2712 core2722 core2732 core2742 core2752 core2762  
core2713 core2723 core2733 core2743 core2753 core2763;  
Categorical = core2711 core2721 core2731 core2741 core2751 core2761  
core2712 core2722 core2732 core2742 core2752 core2762  
core2713 core2723 core2733 core2743 core2753 core2763;

Missing = all(-999);  
!Classes = C1(3) C2(3) C3(3); ! for three latent statuses (3LS) model  
Classes = C1(4) C2(4) C3(4); ! for 4LS model

Analysis: Type = mixture;  
Model: %OVERALL%

C2 on C1;  
C3 on C2;

!!! Constraining transition probabilities

[C2#1] (101);  
[C2#2] (102);  
[C2#3] (103);

[C3#1] (101);  
[C3#2] (102);  
[C3#3] (103);

! from Time 1 to Time 2

C2#1 on C1#1 (111);  
C2#1 on C1#2 (112);  
C2#1 on C1#3 (113);

C2#2 on C1#1 (114);  
C2#2 on C1#2 (115);  
C2#2 on C1#3 (116);

C2#3 on C1#1 (117);  
C2#3 on C1#2 (118);  
C2#3 on C1#3 (119);

! from Time 2 to Time 3  
C3#1 on C2#1 (111);  
C3#1 on C2#2 (112);  
C3#1 on C2#3 (113);

C3#2 on C2#1 (114);  
C3#2 on C2#2 (115);  
C3#2 on C2#3 (116);

C3#3 on C2#1 (117);  
C3#3 on C2#2 (118);  
C3#3 on C2#3 (119);

! Model needs to be set up with the number of assessment  
Model C1:           %C1#1%

! For measurement invariance

[core2711\$1] (1);  
[core2721\$1] (2);  
[core2731\$1] (3);  
[core2741\$1] (4);  
[core2751\$1] (5);  
[core2761\$1] (6);

%C1#2%

[core2711\$1] (7);  
[core2721\$1] (8);  
[core2731\$1] (9);  
[core2741\$1] (10);  
[core2751\$1] (11);  
[core2761\$1] (12);

%C1#3%

[core2711\$1] (13);

[core2721\$1] (14);  
[core2731\$1] (15);  
[core2741\$1] (16);  
[core2751\$1] (17);  
[core2761\$1] (18);

! Remove %C1#4% when testing three latent statuses (3LS) model  
%C1#4%

[core2711\$1] (19);  
[core2721\$1] (20);  
[core2731\$1] (21);  
[core2741\$1] (22);  
[core2751\$1] (23);  
[core2761\$1] (24);

Model C2:

%C2#1%

[core2712\$1] (1);  
[core2722\$1] (2);  
[core2732\$1] (3);  
[core2742\$1] (4);  
[core2752\$1] (5);  
[core2762\$1] (6);

%C2#2%

[core2712\$1] (7);  
[core2722\$1] (8);  
[core2732\$1] (9);  
[core2742\$1] (10);  
[core2752\$1] (11);  
[core2762\$1] (12);

%C2#3%

[core2712\$1] (13);  
[core2722\$1] (14);  
[core2732\$1] (15);  
[core2742\$1] (16);  
[core2752\$1] (17);  
[core2762\$1] (18);

!Remove %C1#4% when testing three latent statuses (3LS) model  
%C2#4%

[core2712\$1] (19);  
[core2722\$1] (20);  
[core2732\$1] (21);  
[core2742\$1] (22);  
[core2752\$1] (23);  
[core2762\$1] (24);

Model C3:

%C3#1%

[core2713\$1] (1);  
[core2723\$1] (2);  
[core2733\$1] (3);  
[core2743\$1] (4);  
[core2753\$1] (5);  
[core2763\$1] (6);

%C3#2%

[core2713\$1] (7);  
[core2723\$1] (8);  
[core2733\$1] (9);  
[core2743\$1] (10);  
[core2753\$1] (11);  
[core2763\$1] (12);

%C3#3%

[core2713\$1] (13);  
[core2723\$1] (14);  
[core2733\$1] (15);  
[core2743\$1] (16);  
[core2753\$1] (17);  
[core2763\$1] (18);

! Remove %C1#4% when testing three latent statuses (3LS) model  
%C3#4%

[core2713\$1] (19);  
[core2723\$1] (20);  
[core2733\$1] (21);  
[core2743\$1] (22);  
[core2753\$1] (23);

```
[core2763$1] (24);
```

```
Output:      Tech1 Tech8 Tech15;
```

## A.2. Código R para los gráficos biplot del análisis de correspondencias múltiple en los 3 períodos estudiados

```
library(haven)
BD <- read_sav("C:/Users/Usuario/Downloads/bullyingTetiquetado.sav")
BD <- as.data.frame(BD)
BD[BD==--999]<-NA
# Filtrando las variables de interés
BD <- BD[,c(1,56:63)]
names(BD)[4] = c("IdiomaIng")
names(BD)[7:9] = c("L1","L2","L3")

library(tidyverse)
BD$Grado <- factor(BD$Grado,
  levels = c(1,2,3,4,5),
  labels = c("5°","6°","7°","8°","9°"))

BD$Género <- factor(BD$Género,
  levels = c(1,2),
  labels = c("Mujer","Varon"))
BD$IdiomaIng <- factor(BD$IdiomaIng,
  levels = c(1,2),
  labels = c("Si","No"))
BD$Raza <- factor(BD$RazaEscuela,
  levels = c(1,2,3,4,8,5,6,10),
  labels = c("Blanca","Negra","Latina","Asiatica",
"Asiatica","Otras","Otras","Otras"))

# PERIODO 1
#=====

BD$Grado <- as.character(BD$Grado)
BD$Género <- as.character(BD$Género)
BD$Idioma <- as.character(BD$IdiomaIng)
BD$Raza <- as.character(BD$Raza)
BD$L1 <- as.character(BD$L1)
```

```

BD$L2    <- as.character(BD$L2)
BD$L3    <- as.character(BD$L3)

library(FactoMineR)
library(factoextra)
library(missMDA)
BD1 = BD[,c(2,4,10,7)]
BD1C <- imputeMCA(BD1,ncp=2)
res1 <-MCA(BD1,tab.disj=BD1C$tab.disj,graph=FALSE)
fviz_mca_var(res1,labelsize=2,repel=TRUE,title = "")

# PERIODO 2
#=====

BD2 = BD[,c(2,4,10,8)]
BD2C <- imputeMCA(BD2,ncp=2)
res2 <-MCA(BD2,tab.disj=BD2C$tab.disj,graph=FALSE)
fviz_mca_var(res2,labelsize=2,repel=TRUE,title = "")

# PERIODO 3
#=====

BD3 = BD[,c(2,4,10,9)]
BD3C <- imputeMCA(BD3,ncp=2)
res3 <-MCA(BD3,tab.disj=BD3C$tab.disj,graph=FALSE)
fviz_mca_var(res3,labelsize=2,repel=TRUE,title = "")

#=====

BDT = BD[,c(2,4,10,7:9)]
BDTC <- imputeMCA(BDT,ncp=2)
resT <-MCA(BDT,tab.disj=BDTC$tab.disj,graph=FALSE)
fviz_mca_var(resT,labelsize=2,repel=TRUE,title = "")

```