

EVALUACIÓN DE LA SALUD BUCAL A TRAVÉS DE LA TEORÍA DE LA RESPUESTA AL ÍTEM EN UN ESTUDIO POBLACIONAL EN URUGUAY

RAMÓN ÁLVAREZ-VAZ^a, FERNANDO MASSA^a

^aInstituto de Estadística, Departamento de Métodos Matemáticos Cuantitativos
Facultad de Ciencias Económicas y de Administración, Universidad de la República, Uruguay
e-mail: ramon@iesta.edu.uy, fmassa@iesta.edu.uy, unidad.biometria.iesta@gmail.com

En los estudios epidemiológicos es práctica habitual trabajar con variables binarias que reflejan la presencia de determinadas enfermedades, las que a su vez pueden estar asociadas con otro conjunto de variables, que en general se asumen como factores de riesgo de las primeras. En el ámbito de los estudios epidemiológicos referidos a la salud bucal, es común indagar sobre la relación entre la presencia de algunas patologías y ciertas características de los participantes del estudio a través de Modelos lineales generalizados (generalmente regresión logística). No obstante ese tipo de análisis suele ser llevado a cabo de manera separada para cada variable de interés y en ningún momento se obtiene una medida que resuma el estado de cada participante. En este estudio se propone el uso de modelos de Teoría de Respuesta al Ítem (TRI) (específicamente el modelo de Rasch) ya que estos permiten el análisis conjunto de varias variables obteniendo como subproducto una valoración individual, que en este caso se interpreta como “sickness proneness” (propensión a la enfermedad). Adicionalmente, el análisis presentado aquí extiende el modelo de Rasch incluyendo un predictor lineal que permite indagar sobre algunos de los factores determinantes de la tendencia de los individuos a padecer las distintas patologías. Los resultados de este estudio confirman el efecto de la falta de actividad física y el avance de la edad sobre la disposición a presentar enfermedades bucales.

Keywords: Enfermedades no transmisibles; Estudios Epidemiológicos; Modelos de Rasch; Salud Bucal

1. Introducción

En los estudios epidemiológicos es práctica habitual trabajar con variables binarias que reflejan la presencia o ausencia de determinadas enfermedades, las que a su vez pueden estar asociadas con otro conjunto de variables, que en general se asumen como factores de riesgo de las primeras. El estudio de las patologías bucales más comunes (caries, pérdida de inserción, bolsa periodontal, dentición funcional) puede llevarse a cabo a través de distintos indicadores, siendo muy común el uso de variables binarias que representan la presencia/ausencia de cada patología [1]. En el ámbito de las encuestas poblacionales, es práctica común que el análisis epidemiológico se encargue de indagar sobre los factores que propician la ocurrencia de dichas patologías valiéndose de Modelos Lineales Generalizados (GLM) [4]. De esta manera es posible determinar cuáles son los condicionantes de una determinada patología. Sin embargo, estos modelos relativamente sencillos y ampliamente usados no son capaces de considerar el

análisis simultáneamente en varias patologías.

Por esta razón, se propone utilizar modelos de la Teoría de Respuesta al Ítem (TRI) [2] en el ámbito epidemiológico debido a que:

- Los modelos TRI son capaces de analizar conjuntamente un grupo de variables de interés
- son capaces de proporcionar una valoración de cada individuo.

La forma que tiene el modelo TRI más sencillo es el que corresponde a una variable de tipo dicotómica θ_{ij} (manteniendo la notación que habitualmente se maneja en el ámbito de los TRI [2]), donde se presentan en la Figura 1 un modelo teórico de Rasch y en la Figura 2 curvas para modelos teóricos de Rasch con diferentes parámetros de dificultad δ , donde Y_{ij} es la respuesta correcta para el ítem j del individuo i

$$\mathbb{P}(Y_{ij} = 1|\theta_i, \delta_j) = \frac{e^{\theta_i - \delta_j}}{1 + e^{\theta_i - \delta_j}} \quad j = 1, 2, 3, 4 \quad (1)$$

$$i = 1, \dots, n = 602$$

que se conoce como el modelo de Rasch o modelo de un sólo parámetro.

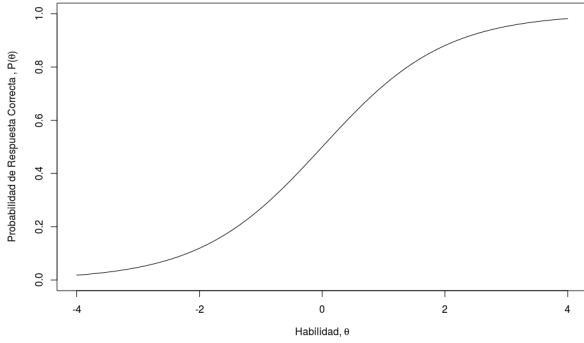


Figura 1. Modelo de un sólo parámetro o de Rasch.

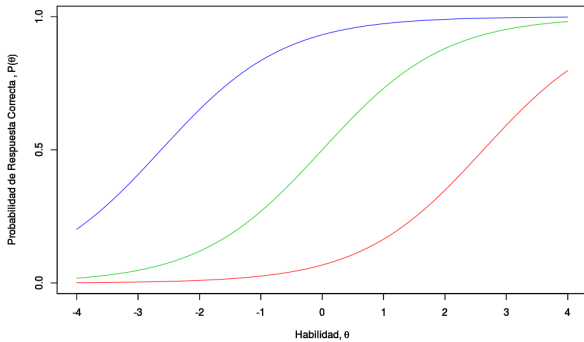


Figura 2. Comparación de Modelo de de Rasch con diferentes parámetros de dificultad.

Sin embargo, los modelos TRI más usualmente utilizados proporcionan índices que describen el comportamiento de cada variable sin considerar el posible efecto de otras variables explicativas. Tratando de superar esta dificultad se propone modelar las variables de interés a través de un modelo de Rasch donde el comportamiento del parámetro individual esté determinado por una distribución normal cuya media sea modelada por un predictor lineal (véase la Figura 3).

2. Metodología estadística

Como se mencionó en la sección anterior, dada la naturaleza binaria de las variables odontológicas consideradas, el modelo de Rasch surge como un punto

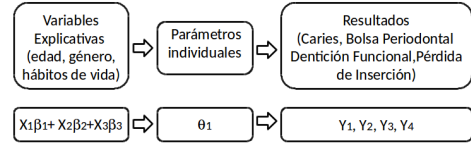


Figura 3. Diagrama de flujo del modelo considerado

de partida natural para el análisis. El modelo se generaliza al considerar un conjunto de predictores en la media del efecto aleatorio utilizado para modelar el comportamiento de cada individuo de un total de N , a los que se les evalúan k variables odontológicas.

$$\mathbb{P}(Y_{ij} = 1|\theta_i, \delta_j) = \frac{e^{\theta_i - \delta_j}}{1 + e^{\theta_i - \delta_j}} \quad j = 1, 2, 3, 4$$

$$\theta_i \sim N(X_i^T \beta, 1) \quad i = 1, \dots, n = N \quad (2)$$

donde en la Ecuación (2), Y_{ij} representa la ocurrencia de la enfermedad j en el participante i , θ_i es el parámetro individual (que en este contexto se puede interpretar como la propensión a la enfermedad (*sickness proneness*) de cada participante), δ_j es el parámetro de dificultad de cada variable (que aquí se relaciona con la prevalencia de cada patología), $X_i^T \beta$ es un predictor lineal que considera los efectos de las variables que habitualmente en otros contextos, donde se usan los modelos de variable de respuesta discreta, ofician de factores de riesgo, como por ejemplo en los modelos de regresión logística binaria [3].

La función de verosimilitud del modelo en este caso que considera 4 variables explicadas a través del TRI es la siguiente que aparece en la Ecuación (3)

$$\mathcal{L}(Y|X, \beta, \delta)$$

$$= \prod_{i=1}^n \int \prod_{j=1}^4 \mathbb{P}(\theta, \delta_j)^{Y_{ij}} (1 - \mathbb{P}(\theta, \delta_j))^{(1-Y_{ij})} * \quad (3)$$

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta - X_i \beta)^2}{2}} d\theta$$

La maximización de dicha función de verosimilitud se hace numéricamente y a través de una aproximación de la matriz hessiana se obtiene una aproximación de la varianza de cada estimador. Todos los cálculos son llevados a cabo en el software R, para los cuales se programaron diferentes funciones [5], [6].



3. Aplicación

Se trabaja con los datos provenientes del estudio sobre personas que demandan atención en la Facultad de Odontología de la Universidad de la República, Uruguay y que son evaluados por los odontólogos del Servicio de registros de la Facultad. Se aplica una muestra de 602 personas que consultan en el período corresponde a mayo 2015-junio 2016, los que se seleccionan mediante muestreo sistemático, a los que se les aplica un cuestionario sociodemográfico y un examen completo de la boca, en donde se evalúa el estado de las piezas dentales y de la mucosa. A partir de esta información se consideran para el análisis y el modelado los siguientes atributos que conforman 3 bloques de variables:

Tabla. 1. Bloques de Variables utilizadas

Variable	Descripción	Bloque
V1	Fuma a diario	1
V2	Consumo nocivo de alcohol	1
V3	Actividad física insuficiente	1
V4	IMC sobrepeso/obesidad,	2
V5	Razón de Cintura Cadera	2
V6	Hipertensión	2
V7	Diabetes	2
V8	Presencia bolsa	3
V9	Pérdida Dentaria	3
V10	Prevalencia de Caries	3
V11	Prevalencia de PIP	3

Las primeras 3 variables constituyen factores de riesgo que están en el (bloque 1), que son hábitos de vida o lo que se denomina variables conductuales, las que frecuentemente se asocian con las *Enfermedades no Transmisibles* (ENT) que aparecen en el (bloque 2), las que a su vez son patologías (enfermedades) y que también son factores de riesgo al mismo tiempo para las variables del bloque odontológico (bloque 3). La definición de presencia o ausencia de los factores de riesgo del (bloque 1) y del (bloque 2) siguen la lógica de la Encuesta nacional de factores de riesgo de Uruguay. En particular se usa un nuevo indicador llamado *Razón de cintura* (V5) que compara la *Circunferencia de cintura* para hombres y mujeres contra valores de referencia máximos, por lo cual razones de cintura mayor a 1 indican patología. En cuanto a las patologías bucales, se trabaja con la *Presencia de caries* (V10), la *Pérdida de inserción de las piezas*, también considerada como Prevalencia de pip, (V11), la *Pérdida dentaria* (V9) que este caso representa *Falta de dentición funcional* y por último la *Presencia de bolsa* que es un indicador relativo a la patología de la mucosa en cada sextante en los que se divide la boca.

Existen más de una forma de trabajar con el bloque de variables explicativas (bloque 2) y (bloque 3), además de considerar, atributos como la edad y el sexo, por lo

Tabla. 2. Prevalencias de variables estudiadas

Variable	Descripción	Prevalencia
V1	Fuma a diario	33.1
V2	Consumo nocivo de alcohol	9.8
V3	Actividad física insuficiente	44.7
V4	IMC sobrepeso/obesidad,	57.3
V5	Razón de cintura/cadera	56.3
V6	Hipertensión	43.2
V7	Diabetes	21.3
V8	Presencia bolsa	58.6
V9	Pérdida dentaria	59.6
V10	Prevalencia de caries	72.8
V11	Prevalencia de pip	63.6

cual, la estrategia que se propone luego de haber estimado varios modelos es la siguiente:

1. Trabajar con sexo, actividad física y un spline para la edad (modelo *m4*)
2. Considerar todas las comportamentales (y sexo edad) que llamaremos (modelo *m5*)
3. Trabajar con todas las ENT (y sexo edad) que llamaremos modelo (*m6*)

En la Tabla 3 se puede observar que en este modelo, debido a que 3 de los parámetros de dificultad son negativos, esas patologías presentan prevalencias relativamente altas en la población estudiada. Para los modelos *m5* y *m6*, en cuanto al predictor lineal se refiere, no existen diferencias por sexo, mientras que si la actividad física es un factor relevante para explicar las prevalencias de las 4 variables bucales del bloque 3. Para el caso de los modelos para la edad, el efecto se modela de manera no paramétrica utilizando splines cúbicos, siendo significativos.

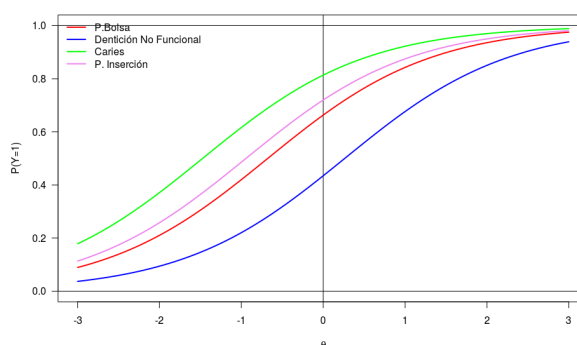


Figura. 4. Curva característica del ítem para cada patología en modelo (*m4*)

En la Tabla 4 pueden verse los coeficientes para los 3 modelos estimados para los predictores lineales.

Tabla. 3. Coeficientes de los parámetros de propensión a la enfermedad

	Estimación	Error Estandard	Z	p-value
Parámetros de dificultad - modelo 4				
P. bolsa	-0.68	0.09	-7.51	≤ 0.001
P. dentaria	0.26	0.09	2.93	≤ 0.001
P. caries	-1.47	0.10	-15.21	≤ 0.001
P. inserción	-0.94	0.09	-10.28	≤ 0.001
Parámetros de dificultad - modelo 5				
P. bolsa	-0.65	0.10	-6.69	≤ 0.001
P. dentaria	0.29	0.10	2.99	≤ 0.001
P. caries	-1.45	0.10	-13.89	≤ 0.001
P. inserción	-0.92	0.10	-9.27	≤ 0.001
Parámetros de dificultad - modelo 6				
P. bolsa	-0.74	0.09	-7.85	≤ 0.001
P. dentaria	0.20	0.09	2.13	0.03
P. caries	-1.53	0.10	-15.23	≤ 0.001
P. inserción	-1.00	0.10	-10.50	≤ 0.001

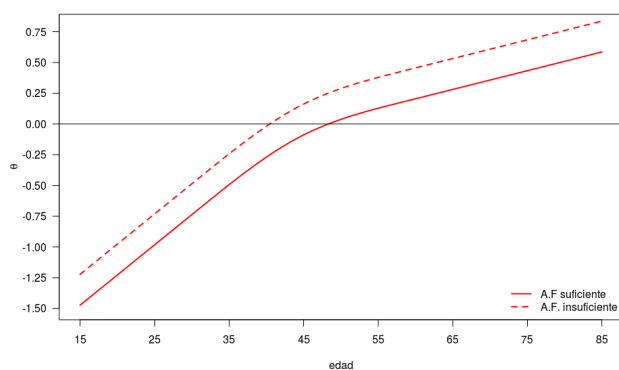


Figura. 5. Efecto de la edad en la propensión a la enfermedad (modelo *m4*).

Como ejemplo para el efecto de la edad, la Figura 5 muestra para el (modelo *m4*) el efecto no lineal de la misma, con un cambio en la velocidad de crecimiento de la *sickness proneness* a partir de los 45 años. Se puede observar que al aumentar la edad, la propensión a presentar alguna enfermedad bucal aumenta. Sin embargo no se detectó un efecto significativo del hábito de fumar sobre la propensión a tener enfermedades bucales.

4. Conclusión

A través de los diferentes modelos propuestos se logró determinar el nivel de prevalencia de las patologías estudiadas en forma conjunta así como también el efecto de algunas covariables de interés. Se observó que el hábito de fumar no presenta un efecto significativo, así como tampoco el sexo, mientras que si la actividad física

Tabla. 4. Coeficientes de los parámetros del predictor lineal

	Estimación	Error Estandard	Z	p-value
Parámetros de regresión - modelo 4				
spl.edad0	1.01	0.11	9.57	≤ 0.001
spl.edad1	-0.46	0.10	-4.75	≤ 0.001
act.fis	0.25	0.10	2.41	0.02
Parámetros de regresión - modelo 5				
spl.edad0	1.00	0.12	8.10	≤ 0.001
spl.edad1	-0.43	0.12	-3.61	≤ 0.001
sexo	-0.14	0.11	-1.28	0.20
act.fis	0.27	0.12	2.28	0.02
fuma	0.22	0.12	1.81	0.07
toma	0.12	0.20	0.60	0.55
Parámetros de regresión - modelo 6				
spl.edad0	0.97	0.12	8.22	≤ 0.001
spl.edad1	-0.45	0.12	-3.66	≤ 0.001
sexo	-0.14	0.12	-1.10	0.27
imc	0.12	0.17	0.71	0.48
cc	0.00	0.19	0.00	1.00
hta	0.12	0.14	0.85	0.39
diab	0.01	0.16	0.09	0.93

parece ser un atributo que en cierto manera modula la prevalencia de las enfermedades bucales aumentándolas. Si se observó un efecto no lineal y creciente de la edad sobre la tendencia a padecer enfermedades bucales.

Referencias

- [1] Álvarez Vaz, R. y Massa, F. (2012). Determinación de tipologías de infecciones parasitarias intestinales, en escolares mediante, técnicas de clustering sobre datos binarios. Documento de Trabajo Serie DT (12 / 05) - ISSN : 1688-6453, IESTA.
- [2] Baker, F. (2017). *The basics of item response theory using R*. Springer, Cham, Switzerland.
- [3] Hilbe, J. (2017). *Logistic Regression Models*. CRC Press, S.I.
- [4] McCullagh, P. (1989). *Generalized linear models*. Chapman and Hall, London New York.
- [5] R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [6] Rizopoulos, D. (2006). ltm: An R package for Latent Variable Modelling and Item Response Theory Analyses. *Journal of Statistical Software*, 17(5):1–25.