Introducción a Modelos Psicométricos

Clase 9–11

El Modelo de Rasch

Iwin Leenen y Ramsés Vázquez-Lira

Facultad de Psicología, UNAM

Programa de Licenciatura y Posgrado en Psicología
Semestre 2019–1

Índice

Índice

La ecuación básica del modelo de Rasch

La ecuación básica en el modelo de Rasch

- El modelo de Rasch es similar al modelo de Guttman en cuanto a:
  - ambos son modelos para ítems dicotómicos
  - suponen un parámetro para cada persona (θ_p; su nivel en el rasgo latente) y un parámetro para cada ítem (β_i; su grado de dificultad).

- Ambos modelos difieren respecto a cómo combinan los parámetros de la persona y el ítem para derivar la probabilidad de acertar.
  - El modelo de Rasch es un modelo probabilístico.
  - La ecuación básica del modelo de Rasch es:

$$\Pr(Y_{pi} = 1 | \theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

(e es la base de los logaritmos naturales; e $\approx$ 2.71828.)

La ecuación básica del modelo de Rasch

La ecuación básica en el modelo de Rasch

- El modelo de Rasch es similar al modelo de Guttman en cuanto a:
  - ambos son modelos para ítems dicotómicos
  - suponen un parámetro para cada persona (θ_p; su nivel en el rasgo latente) y un parámetro para cada ítem (β_i; su grado de dificultad).

- Ambos modelos difieren respecto a cómo combinan los parámetros de la persona y el ítem para derivar la probabilidad de acertar.
  - El modelo de Rasch es un modelo probabilístico.
  - La ecuación básica del modelo de Rasch es:

$$\Pr(Y_{pi} = 1 | \theta_p, \beta_i) \; = \; \frac{\mathrm{e}^{\theta_p - \beta_i}}{1 \; + \; \mathrm{e}^{\theta_p - \beta_i}}$$

(e es la base de los logaritmos naturales; $\mathrm{e} \approx 2.71828$.)

La ecuación básica del modelo de Rasch

Unos ejemplos

- Supongamos que un ítem *i* tiene grado de dificultad igual a 1 (es decir, $\beta_i = 1$).
  - Una persona *p* cuyo nivel en el rasgo latente es igual a -1 ($\theta_p = -1$) tendrá probabilidad de acertar el ítem:

$$\Pr(Y_{pi} = 1 | \theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{-1-1}}{1 + e^{-1-1}} = \frac{e^{-2}}{1 + e^{-2}} = .119$$

  - Una persona *p* cuyo nivel en el rasgo latente es igual a 2 ($\theta_p = 2$) tendrá probabilidad:

$$\Pr(Y_{pi} = 1 | \theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{2-1}}{1 + e^{2-1}} = \frac{e^{1}}{1 + e^{1}} = .731$$

- Para un ítem *i* con grado de dificultad igual a 0 (es decir, $\beta_i = 0$), tenemos:
  - Si la persona *p* tiene un nivel en el rasgo latente de 0 ($\theta_p = 0$), entonces:

$$\Pr(Y_{pi} = 1 | \theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{0-0}}{1 + e^{0-0}} = \frac{e^{0}}{1 + e^{0}} = .500$$

  - Una persona *p* cuyo nivel en el rasgo latente es igual a 1 ($\theta_p = 1$) tendrá probabilidad:

$$\Pr(Y_{pi} = 1 | \theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{1-0}}{1 + e^{1-0}} = \frac{e^{1}}{1 + e^{1}} = .731$$

La ecuación básica del modelo de Rasch

Unos ejemplos

- Supongamos que un ítem *i* tiene grado de dificultad igual a 1 (es decir, $\beta_i = 1$).
  - Una persona *p* cuyo nivel en el rasgo latente es igual a -1 ($\theta_p = -1$) tendrá probabilidad de acertar el ítem:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{-1-1}}{1 + e^{-1-1}} = \frac{e^{-2}}{1 + e^{-2}} = .119$$

  - Una persona *p* cuyo nivel en el rasgo latente es igual a 2 ($\theta_p = 2$) tendrá probabilidad:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{2-1}}{1 + e^{2-1}} = \frac{e^{1}}{1 + e^{1}} = .731$$

- Para un ítem *i* con grado de dificultad igual a 0 (es decir, $\beta_i = 0$), tenemos:
  - Si la persona *p* tiene un nivel en el rasgo latente de 0 ($\theta_p = 0$), entonces:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{0-0}}{1 + e^{0-0}} = \frac{e^{0}}{1 + e^{0}} = .500$$

  - Una persona *p* cuyo nivel en el rasgo latente es igual a 1 ($\theta_p = 1$) tendrá probabilidad:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{1-0}}{1 + e^{1-0}} = \frac{e^{1}}{1 + e^{1}} = .731$$

La ecuación básica del modelo de Rasch

Unos ejemplos

- Supongamos que un ítem *i* tiene grado de dificultad igual a 1 (es decir, $\beta_i = 1$).
  - Una persona *p* cuyo nivel en el rasgo latente es igual a -1 ($\theta_p = -1$) tendrá probabilidad de acertar el ítem:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{-1-1}}{1 + e^{-1-1}} = \frac{e^{-2}}{1 + e^{-2}} = .119$$

  - Una persona *p* cuyo nivel en el rasgo latente es igual a 2 ($\theta_p = 2$) tendrá probabilidad:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{2-1}}{1 + e^{2-1}} = \frac{e^{1}}{1 + e^{1}} = .731$$

- Para un ítem *i* con grado de dificultad igual a 0 (es decir, $\beta_i = 0$), tenemos:
  - Si la persona *p* tiene un nivel en el rasgo latente de 0 ($\theta_p = 0$), entonces:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{0-0}}{1 + e^{0-0}} = \frac{e^{0}}{1 + e^{0}} = .500$$

  - Una persona *p* cuyo nivel en el rasgo latente es igual a 1 ($\theta_p = 1$) tendrá probabilidad:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{1-0}}{1 + e^{1-0}} = \frac{e^{1}}{1 + e^{1}} = .731$$

La ecuación básica del modelo de Rasch

Unos ejemplos

- Supongamos que un ítem *i* tiene grado de dificultad igual a 1 (es decir, $\beta_i = 1$).
  - Una persona *p* cuyo nivel en el rasgo latente es igual a -1 ($\theta_p = -1$) tendrá probabilidad de acertar el ítem:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{-1-1}}{1 + e^{-1-1}} = \frac{e^{-2}}{1 + e^{-2}} = .119$$

  - Una persona *p* cuyo nivel en el rasgo latente es igual a 2 ($\theta_p = 2$) tendrá probabilidad:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{2-1}}{1 + e^{2-1}} = \frac{e^{1}}{1 + e^{1}} = .731$$

- Para un ítem *i* con grado de dificultad igual a 0 (es decir, $\beta_i = 0$), tenemos:
  - Si la persona *p* tiene un nivel en el rasgo latente de 0 ($\theta_p = 0$), entonces:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{0-0}}{1 + e^{0-0}} = \frac{e^{0}}{1 + e^{0}} = .500$$

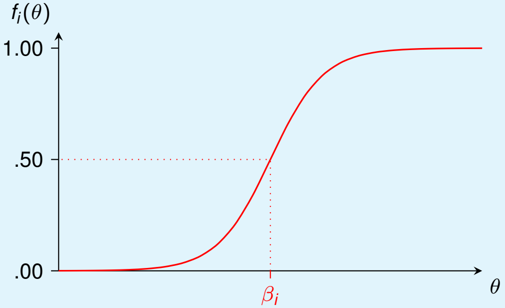
  - Una persona *p* cuyo nivel en el rasgo latente es igual a 1 ($\theta_p = 1$) tendrá probabilidad:

$$\Pr(Y_{pi} = 1|\theta_p, \beta_i) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}} = \frac{e^{1-0}}{1 + e^{1-0}} = \frac{e^{1}}{1 + e^{1}} = .731$$

La curva característica del ítem en el modelo de Rasch

En el modelo de Rasch:

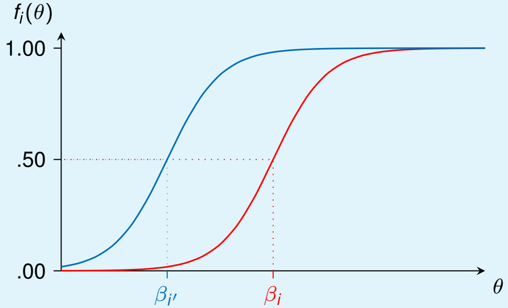
$$f_i(\theta) = \frac{e^{\theta - \beta_i}}{1 + e^{\theta - \beta_i}}$$



La curva característica del ítem en el modelo de Rasch

En el modelo de Rasch:

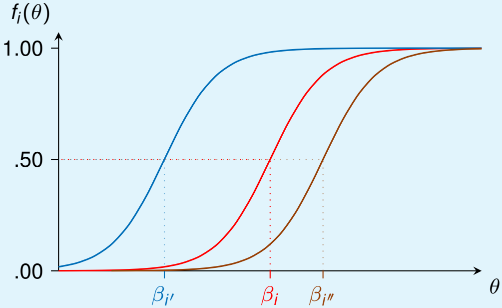
$$f_i(\theta) = \frac{e^{\theta-\beta_i}}{1 + e^{\theta-\beta_i}}$$



La curva característica del ítem en el modelo de Rasch

En el modelo de Rasch:

$$f_i(\theta) = \frac{e^{\theta - \beta_i}}{1 + e^{\theta - \beta_i}}$$



La función logística

Nota: La función logística

En las matemáticas, la función

$$f(x) = \frac{e^x}{1 + e^x}$$

se conoce como la función logística.

Por lo tanto, el modelo de Rasch se conoce como un modelo logístico de la TRI.

Nótese que, a veces, se escribe la función logística como:

$$f(x) = \frac{1}{1 + e^{-x}}.$$

Sin embargo, es la misma función ya que algebraicamente:

$$\frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}}.$$

La función logística

Nota: La función logística

En las matemáticas, la función

$$f(x) = \frac{e^x}{1 + e^x}$$

se conoce como la función logística.

Por lo tanto, el modelo de Rasch se conoce como un modelo logístico de la TRI.

Nótese que, a veces, se escribe la función logística como:

$$f(x) = \frac{1}{1 + e^{-x}}.$$

Sin embargo, es la misma función ya que algebraicamente:

$$\frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}}.$$

La función logística

Nota: La función logística

En las matemáticas, la función

$$f(x) = \frac{e^x}{1 + e^x}$$

se conoce como la función logística.

Por lo tanto, el modelo de Rasch se conoce como un modelo logístico de la TRI.

Nótese que, a veces, se escribe la función logística como:

$$f(x) = \frac{1}{1 + e^{-x}}.$$

Sin embargo, es la misma función ya que algebraicamente:

$$\frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}}.$$

La función logística

Nota: La función logística

La función logística $f(x)$ tiene las siguientes propiedades:

- Para cualquier número real x:
$$0 < f(x) < 1$$
- Es una función monótonamente creciente
- Se cumplen los siguientes límites:
$$\lim_{x \to -\infty} f(x) = 0$$
$$\lim_{x \to +\infty} f(x) = 1$$

En el modelo de Rasch, el argumento de la función logística es la diferencia $\theta_p - \beta_i$.

La función logística

Nota: La función logística

La función logística $f(x)$ tiene las siguientes propiedades:

- Para cualquier número real x:
$$0 < f(x) < 1$$
- Es una función monótonamente creciente
- Se cumplen los siguientes límites:
$$\begin{aligned} \lim_{x \to -\infty} f(x) &= 0 \\ \lim_{x \to +\infty} f(x) &= 1 \end{aligned}$$

En el modelo de Rasch, el argumento de la función logística es la diferencia $\theta_p - \beta_i$.

La función logística

Nota: La función logística

La función logística $f(x)$ tiene las siguientes propiedades:

- Para cualquier número real x:
$$0 < f(x) < 1$$
- Es una función monótonamente creciente
- Se cumplen los siguientes límites:
$$\lim_{x \to -\infty} f(x) = 0$$
$$\lim_{x \to +\infty} f(x) = 1$$

En el modelo de Rasch, el argumento de la función logística es la diferencia $\theta_p - \beta_i$.

La función logística

Nota: La función logística

La función logística $f(x)$ tiene las siguientes propiedades:

- Para cualquier número real x:
$$0 < f(x) < 1$$
- Es una función monótonamente creciente
- Se cumplen los siguientes límites:
$$\begin{aligned} \lim_{x \to -\infty} f(x) &= 0 \\ \lim_{x \to +\infty} f(x) &= 1 \end{aligned}$$

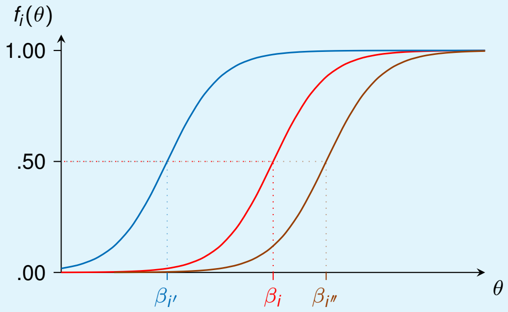
En el modelo de Rasch, el argumento de la función logística es la diferencia $\theta_p - \beta_i$.

La curva característica del ítem en el modelo de Rasch

En el modelo de Rasch:

$$f_i(\theta) = \frac{e^{\theta - \beta_i}}{1 + e^{\theta - \beta_i}}$$

Propiedades de las CCIs en el modelo de Rasch:



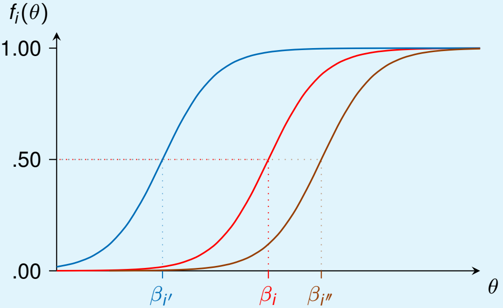
La curva característica del ítem en el modelo de Rasch

En el modelo de Rasch:

$$f_i(\theta) = \frac{e^{\theta - \beta_i}}{1 + e^{\theta - \beta_i}}$$

Propiedades de las CCIs en el modelo de Rasch:

- Son continuas



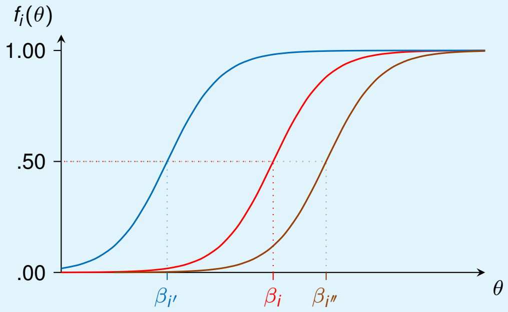
La curva característica del ítem en el modelo de Rasch

En el modelo de Rasch:

$$f_i(\theta) = \frac{e^{\theta - \beta_i}}{1 + e^{\theta - \beta_i}}$$

Propiedades de las CCIs en el modelo de Rasch:

- Son monótonamente crecientes



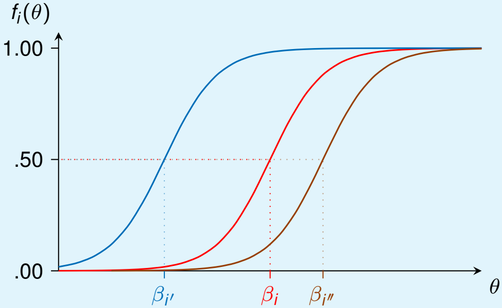
La curva característica del ítem en el modelo de Rasch

En el modelo de Rasch:

$$f_i(\theta) = \frac{e^{\theta - \beta_i}}{1 + e^{\theta - \beta_i}}$$

Propiedades de las CCIs en el modelo de Rasch:

- $\lim_{\theta \to -\infty} f_i(\theta) = 0$ y $\lim_{\theta \to +\infty} f_i(\theta) = 1$



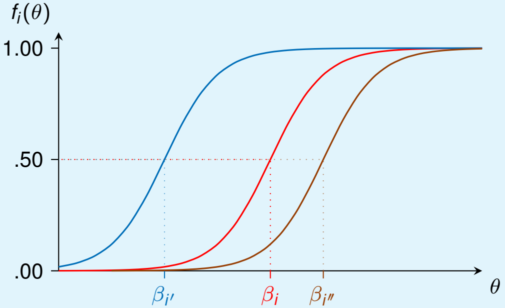
La curva característica del ítem en el modelo de Rasch

En el modelo de Rasch:

$$f_i(\theta) = \frac{e^{\theta-\beta_i}}{1 + e^{\theta-\beta_i}}$$

Propiedades de las CCIs en el modelo de Rasch:

- Nunca se cruzan



Una indeterminación del modelo

Ejemplo

Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

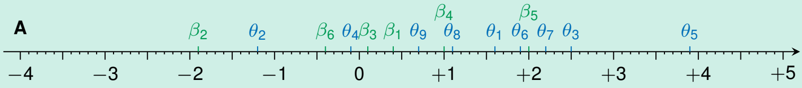
	A
θ_1	$+1.6$
θ_2	-1.2
θ_3	$+2.5$
θ_4	-0.1
θ_5	$+3.9$
θ_6	$+1.9$
θ_7	$+2.2$
θ_8	$+1.1$
θ_9	$+0.7$
β_1	$+0.4$
β_2	-1.9
β_3	$+0.1$
β_4	$+1.0$
β_5	$+2.0$
β_6	-0.4

Una indeterminación del modelo

Ejemplo

Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

	A
θ_1	$+1.6$
θ_2	-1.2
θ_3	$+2.5$
θ_4	-0.1
θ_5	$+3.9$
θ_6	$+1.9$
θ_7	$+2.2$
θ_8	$+1.1$
θ_9	$+0.7$
β_1	$+0.4$
β_2	-1.9
β_3	$+0.1$
β_4	$+1.0$
β_5	$+2.0$
β_6	-0.4



Una indeterminación del modelo

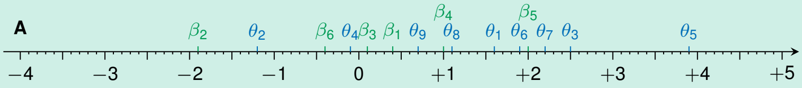
Ejemplo

Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

	A
θ_1	$+1.6$
θ_2	-1.2
θ_3	$+2.5$
θ_4	-0.1
θ_5	$+3.9$
θ_6	$+1.9$
θ_7	$+2.2$
θ_8	$+1.1$
θ_9	$+0.7$
β_1	$+0.4$
β_2	-1.9
β_3	$+0.1$
β_4	$+1.0$
β_5	$+2.0$
β_6	-0.4

Modelo:

$$\forall p, i : f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$



Una indeterminación del modelo

Ejemplo

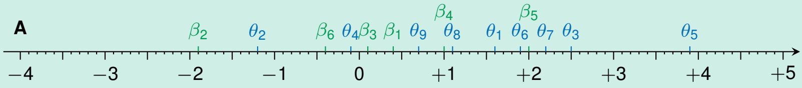
Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

	A	B
θ_1	$+1.6$	$+2.6$
θ_2	-1.2	-0.2
θ_3	$+2.5$	$+3.5$
θ_4	-0.1	$+0.9$
θ_5	$+3.9$	$+4.9$
θ_6	$+1.9$	$+2.9$
θ_7	$+2.2$	$+3.2$
θ_8	$+1.1$	$+2.1$
θ_9	$+0.7$	$+1.7$
β_1	$+0.4$	$+1.4$
β_2	-1.9	-0.9
β_3	$+0.1$	$+1.1$
β_4	$+1.0$	$+2.0$
β_5	$+2.0$	$+3.0$
β_6	-0.4	$+0.6$

Modelo:

$$\forall p, i : f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

A



Una indeterminación del modelo

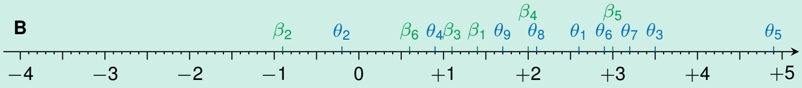
Ejemplo

Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

	A	B
θ_1	$+1.6$	$+2.6$
θ_2	-1.2	-0.2
θ_3	$+2.5$	$+3.5$
θ_4	-0.1	$+0.9$
θ_5	$+3.9$	$+4.9$
θ_6	$+1.9$	$+2.9$
θ_7	$+2.2$	$+3.2$
θ_8	$+1.1$	$+2.1$
θ_9	$+0.7$	$+1.7$
β_1	$+0.4$	$+1.4$
β_2	-1.9	-0.9
β_3	$+0.1$	$+1.1$
β_4	$+1.0$	$+2.0$
β_5	$+2.0$	$+3.0$
β_6	-0.4	$+0.6$

Modelo:

$$\forall p, i : f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$



Una indeterminación del modelo

Ejemplo

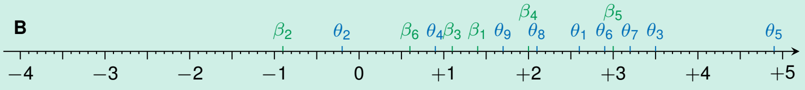
Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

	A	B
θ_1	$+1.6$	$+2.6$
θ_2	-1.2	-0.2
θ_3	$+2.5$	$+3.5$
θ_4	-0.1	$+0.9$
θ_5	$+3.9$	$+4.9$
θ_6	$+1.9$	$+2.9$
θ_7	$+2.2$	$+3.2$
θ_8	$+1.1$	$+2.1$
θ_9	$+0.7$	$+1.7$
β_1	$+0.4$	$+1.4$
β_2	-1.9	-0.9
β_3	$+0.1$	$+1.1$
β_4	$+1.0$	$+2.0$
β_5	$+2.0$	$+3.0$
β_6	-0.4	$+0.6$

Modelo:

$$\forall p, i : f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

Se obtienen las mismas probabilidades en $f_j(\theta)$.



Una indeterminación del modelo

Ejemplo

Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

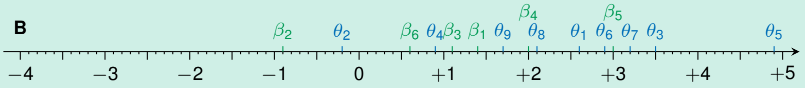
	A	B
θ_1	$+1.6$	$+2.6$
θ_2	-1.2	-0.2
θ_3	$+2.5$	$+3.5$
θ_4	-0.1	$+0.9$
θ_5	$+3.9$	$+4.9$
θ_6	$+1.9$	$+2.9$
θ_7	$+2.2$	$+3.2$
θ_8	$+1.1$	$+2.1$
θ_9	$+0.7$	$+1.7$
β_1	$+0.4$	$+1.4$
β_2	-1.9	-0.9
β_3	$+0.1$	$+1.1$
β_4	$+1.0$	$+2.0$
β_5	$+2.0$	$+3.0$
β_6	-0.4	$+0.6$

Modelo:

$$\forall p, i : f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

Se obtienen las mismas probabilidades en $f_j(\theta)$.

$\longrightarrow$ **indeterminación**.



Una indeterminación del modelo

Ejemplo

Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

	A	B	C
θ_1	$+1.6$	$+2.6$	$+1.2$
θ_2	-1.2	-0.2	-1.6
θ_3	$+2.5$	$+3.5$	$+2.1$
θ_4	-0.1	$+0.9$	-0.5
θ_5	$+3.9$	$+4.9$	$+3.5$
θ_6	$+1.9$	$+2.9$	$+1.5$
θ_7	$+2.2$	$+3.2$	$+1.8$
θ_8	$+1.1$	$+2.1$	$+0.7$
θ_9	$+0.7$	$+1.7$	$+0.3$
β_1	$+0.4$	$+1.4$	0.0
β_2	-1.9	-0.9	-2.3
β_3	$+0.1$	$+1.1$	-0.3
β_4	$+1.0$	$+2.0$	$+0.6$
β_5	$+2.0$	$+3.0$	$+1.6$
β_6	-0.4	$+0.6$	-0.8

Modelo:

$$\forall p, i : f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

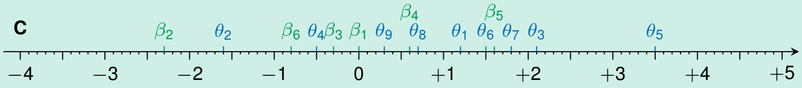
Se obtienen las mismas probabilidades en $f_j(\theta)$.

$\longrightarrow$ **indeterminación**.

Tres maneras para resolverla:

- Fijar $\beta_1 = 0$.

C



Una indeterminación del modelo

Ejemplo

Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

	A	B	C	D
θ_1	$+1.6$	$+2.6$	$+1.2$	$+1.4$
θ_2	-1.2	-0.2	-1.6	-1.4
θ_3	$+2.5$	$+3.5$	$+2.1$	$+2.3$
θ_4	-0.1	$+0.9$	-0.5	-0.3
θ_5	$+3.9$	$+4.9$	$+3.5$	$+3.7$
θ_6	$+1.9$	$+2.9$	$+1.5$	$+1.7$
θ_7	$+2.2$	$+3.2$	$+1.8$	$+2.0$
θ_8	$+1.1$	$+2.1$	$+0.7$	$+0.9$
θ_9	$+0.7$	$+1.7$	$+0.3$	$+0.5$
β_1	$+0.4$	$+1.4$	0.0	$+0.2$
β_2	-1.9	-0.9	-2.3	-2.1
β_3	$+0.1$	$+1.1$	-0.3	-0.1
β_4	$+1.0$	$+2.0$	$+0.6$	$+0.8$
β_5	$+2.0$	$+3.0$	$+1.6$	$+1.8$
β_6	-0.4	$+0.6$	-0.8	-0.6

Modelo:

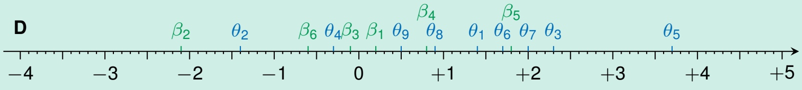
$$\forall p, i : f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

Se obtienen las mismas probabilidades en $f_i(\theta)$.

$\longrightarrow$ **indeterminación**.

Tres maneras para resolverla:

- Fijar $\beta_1 = 0$.
- Restringir $\sum_i \beta_i = 0$.



Una indeterminación del modelo

Ejemplo

Supongamos que el modelo de Rasch se cumple para 9 personas y 6 ítems con θ's y β's igual a:

	A	B	C	D	E
θ_1	$+1.6$	$+2.6$	$+1.2$	$+1.4$	$+0.2$
θ_2	-1.2	-0.2	-1.6	-1.4	-2.6
θ_3	$+2.5$	$+3.5$	$+2.1$	$+2.3$	$+1.1$
θ_4	-0.1	$+0.9$	-0.5	-0.3	-1.5
θ_5	$+3.9$	$+4.9$	$+3.5$	$+3.7$	$+2.5$
θ_6	$+1.9$	$+2.9$	$+1.5$	$+1.7$	$+0.5$
θ_7	$+2.2$	$+3.2$	$+1.8$	$+2.0$	$+0.8$
θ_8	$+1.1$	$+2.1$	$+0.7$	$+0.9$	-0.3
θ_9	$+0.7$	$+1.7$	$+0.3$	$+0.5$	-0.7
β_1	$+0.4$	$+1.4$	0.0	$+0.2$	-1.0
β_2	-1.9	-0.9	-2.3	-2.1	-3.3
β_3	$+0.1$	$+1.1$	-0.3	-0.1	-1.3
β_4	$+1.0$	$+2.0$	$+0.6$	$+0.8$	-0.4
β_5	$+2.0$	$+3.0$	$+1.6$	$+1.8$	$+0.6$
β_6	-0.4	$+0.6$	-0.8	-0.6	-1.8

Modelo:

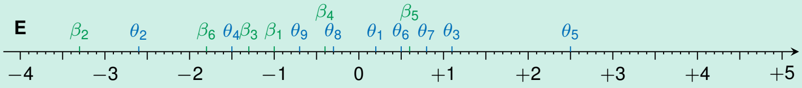
$$\forall p, i : f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

Se obtienen las mismas probabilidades en $f_i(\theta)$.

$\longrightarrow$ **indeterminación**.

Tres maneras para resolverla:

- Fijar $\beta_1 = 0$.
- Restringir $\sum_i \beta_i = 0$.
- Restringir $\sum_p \theta_p = 0$.



Una indeterminación del modelo

Implicaciones de esta indeterminación

- Si en dos soluciones para el modelo de Rasch, los parámetros difieren únicamente por una constante

  $\Longrightarrow$ Son dos soluciones equivalentes;
  $\Longrightarrow$ En esencia, son idénticas.

  Compara con:

  $$10^\circ C \;=\; 50^\circ F \;=\; 283.15\,K$$

  Son diferentes números, pero temperaturas idénticas.

- El nivel de medición de los parámetros en el modelos de Rasch es

Una indeterminación del modelo

Implicaciones de esta indeterminación

- Si en dos soluciones para el modelo de Rasch, los parámetros difieren únicamente por una constante

  $\Longrightarrow$ Son dos soluciones equivalentes;
  $\Longrightarrow$ En esencia, son idénticas.

  Compara con:

  $$10^\circ C \;=\; 50^\circ F \;=\; 283.15\,K$$

  Son diferentes números, pero temperaturas idénticas.

- El nivel de medición de los parámetros en el modelos de Rasch es

Una indeterminación del modelo

Implicaciones de esta indeterminación

- Si en dos soluciones para el modelo de Rasch, los parámetros difieren únicamente por una constante

  $\Longrightarrow$ Son dos soluciones equivalentes;
  $\Longrightarrow$ En esencia, son idénticas.

  Compara con:

  $$10^\circ C \; = \; 50^\circ F \; = \; 283.15\,K$$

  Son diferentes números, pero temperaturas idénticas.

- El nivel de medición de los parámetros en el modelos de Rasch es

Una indeterminación del modelo

Implicaciones de esta indeterminación

- Si en dos soluciones para el modelo de Rasch, los parámetros difieren únicamente por una constante

  $\Longrightarrow$ Son dos soluciones equivalentes;
  $\Longrightarrow$ En esencia, son idénticas.

  Compara con:

  $$10^\circ C \;=\; 50^\circ F \;=\; 283.15\,K$$

  Son diferentes números, pero temperaturas idénticas.

- El nivel de medición de los parámetros en el modelos de Rasch es intervalar.

Índice

El modelo de Rasch: supuestos

Los supuestos del modelo de Rasch

Gerhard Fischer (1987) mostró que el modelo Rasch es matemáticamente equivalente al siguiente conjunto de supuestos:

1. θ is unidimensional y la CCI $f_i(\theta)$ es continua y monótonamente creciente para cualquier ítem i;
2. Para todos los ítems, se cumple que:
$$\lim_{\theta \to -\infty} f_i(\theta) = 0 \quad \text{y} \quad \lim_{\theta \to +\infty} f_i(\theta) = 1;$$
3. Las respuestas de cada persona son *localmente independientes*;
4. El número de aciertos es un estadístico suficiente para θ.

El modelo de Rasch y estos supuestos son matemáticamente equivalentes. Esto quiere decir:

El modelo de Rasch se cumple.

$$\Downarrow$$

Todos estos supuestos se cumplen.

El modelo de Rasch: supuestos

Los supuestos del modelo de Rasch

Gerhard Fischer (1987) mostró que el modelo Rasch es matemáticamente equivalente al siguiente conjunto de supuestos:

1. θ is unidimensional y la CCI $f_i(\theta)$ es continua y monótonamente creciente para cualquier ítem i;
2. Para todos los ítems, se cumple que:
$$\lim_{\theta \to -\infty} f_i(\theta) = 0 \quad \text{y} \quad \lim_{\theta \to +\infty} f_i(\theta) = 1;$$
3. Las respuestas de cada persona son *localmente independientes*;
4. El número de aciertos es un estadístico suficiente para θ.

El modelo de Rasch y estos supuestos son matemáticamente equivalentes. Esto quiere decir:

El modelo de Rasch se cumple.

$\Downarrow$

Todos estos supuestos se cumplen.

El modelo de Rasch: supuestos

Los supuestos del modelo de Rasch

Gerhard Fischer (1987) mostró que el modelo Rasch es matemáticamente equivalente al siguiente conjunto de supuestos:

1. θ is unidimensional y la CCI $f_i(\theta)$ es continua y monótonamente creciente para cualquier ítem i;
2. Para todos los ítems, se cumple que:
$$\lim_{\theta \to -\infty} f_i(\theta) = 0 \quad \text{y} \quad \lim_{\theta \to +\infty} f_i(\theta) = 1;$$
3. Las respuestas de cada persona son *localmente independientes*;
4. El número de aciertos es un estadístico suficiente para θ.

El modelo de Rasch y estos supuestos son matemáticamente equivalentes. Esto quiere decir:

El modelo de Rasch se cumple.

$\Downarrow$

Todos estos supuestos se cumplen.

El modelo de Rasch: supuestos

Los supuestos del modelo de Rasch

Gerhard Fischer (1987) mostró que el modelo Rasch es matemáticamente equivalente al siguiente conjunto de supuestos:

1. θ is unidimensional y la CCI $f_i(\theta)$ es continua y monótonamente creciente para cualquier ítem i;
2. Para todos los ítems, se cumple que:
$$\lim_{\theta \to -\infty} f_i(\theta) = 0 \quad \text{y} \quad \lim_{\theta \to +\infty} f_i(\theta) = 1;$$
3. Las respuestas de cada persona son *localmente independientes*;
4. El número de aciertos es un estadístico suficiente para θ.

El modelo de Rasch y estos supuestos son matemáticamente equivalentes. Esto quiere decir:

El modelo de Rasch se cumple.

$\Downarrow$

Todos estos supuestos se cumplen.

El modelo de Rasch: supuestos

Los supuestos del modelo de Rasch

Gerhard Fischer (1987) mostró que el modelo Rasch es matemáticamente equivalente al siguiente conjunto de supuestos:

1. θ is unidimensional y la CCI $f_i(\theta)$ es continua y monótonamente creciente para cualquier ítem i;
2. Para todos los ítems, se cumple que:
$$\lim_{\theta \to -\infty} f_i(\theta) = 0 \quad \text{y} \quad \lim_{\theta \to +\infty} f_i(\theta) = 1;$$
3. Las respuestas de cada persona son *localmente independientes*;
4. El número de aciertos es un estadístico suficiente para θ.

El modelo de Rasch y estos supuestos son matemáticamente equivalentes. Esto quiere decir:

El modelo de Rasch se cumple.

$\Downarrow$

Todos estos supuestos se cumplen.

El modelo de Rasch: supuestos

Los supuestos del modelo de Rasch

Gerhard Fischer (1987) mostró que el modelo Rasch es matemáticamente equivalente al siguiente conjunto de supuestos:

1. θ is unidimensional y la CCI $f_i(\theta)$ es continua y monótonamente creciente para cualquier ítem i;
2. Para todos los ítems, se cumple que:
$$\lim_{\theta\to-\infty} f_i(\theta) = 0 \quad \text{y} \quad \lim_{\theta\to+\infty} f_i(\theta) = 1;$$
3. Las respuestas de cada persona son *localmente independientes*;
4. El número de aciertos es un estadístico suficiente para θ.

El modelo de Rasch y estos supuestos son matemáticamente equivalentes. Esto quiere decir:

El modelo de Rasch se cumple.

$\Downarrow$

Todos estos supuestos se cumplen.

El modelo de Rasch: supuestos

Los supuestos del modelo de Rasch

Gerhard Fischer (1987) mostró que el modelo Rasch es matemáticamente equivalente al siguiente conjunto de supuestos:

1. θ is unidimensional y la CCI $f_i(\theta)$ es continua y monótonamente creciente para cualquier ítem i;
2. Para todos los ítems, se cumple que:
$$\lim_{\theta \to -\infty} f_i(\theta) = 0 \quad \text{y} \quad \lim_{\theta \to +\infty} f_i(\theta) = 1;$$
3. Las respuestas de cada persona son *localmente independientes*;
4. El número de aciertos es un estadístico suficiente para θ.

El modelo de Rasch y estos supuestos son matemáticamente equivalentes. Esto quiere decir:

El modelo de Rasch se cumple.

$$\Updownarrow$$

Todos estos supuestos se cumplen.

El supuesto de independencia local

Independencia local

El supuesto de independencia local tiene que ver con las siguientes preguntas:

> *¿A qué se debe la correlación entre los ítems?*

> *¿Por qué, al comparar una persona que acertó el ítem i con otra persona que no lo acertó, la primera persona tiene mayor probabilidad de acertar otro ítem i'?*

Si el supuesto de independencia local se cumple, entonces la respuesta a las preguntas anteriores es:

> *La **única** razón es que las personas difieren respecto al valor de su parámetro θ.*

Dentro de un grupo de personas que todos tienen el mismo valor para θ, no habrá correlación entre ítems.

El supuesto de independencia local

Independencia local

El supuesto de independencia local tiene que ver con las siguientes preguntas:

> *¿A qué se debe la correlación entre los ítems?*
>
> *¿Por qué, al comparar una persona que acertó el ítem i con otra persona que no lo acertó, la primera persona tiene mayor probabilidad de acertar otro ítem i'?*

Si el supuesto de independencia local se cumple, entonces la respuesta a las preguntas anteriores es:

> *La **única** razón es que las personas difieren respecto al valor de su parámetro θ.*

Dentro de un grupo de personas que todos tienen el mismo valor para θ, no habrá correlación entre ítems.

El supuesto de independencia local

Independencia local

El supuesto de independencia local tiene que ver con las siguientes preguntas:

> *¿A qué se debe la correlación entre los ítems?*
>
> *¿Por qué, al comparar una persona que acertó el ítem i con otra persona que no lo acertó, la primera persona tiene mayor probabilidad de acertar otro ítem i'?*

Si el supuesto de independencia local se cumple, entonces la respuesta a las preguntas anteriores es:

> *La **única** razón es que las personas difieren respecto al valor de su parámetro θ.*

Dentro de un grupo de personas que todos tienen el mismo valor para θ, no habrá correlación entre ítems.

El supuesto de independencia local

Independencia local

El supuesto de independencia local tiene que ver con las siguientes preguntas:

> *¿A qué se debe la correlación entre los ítems?*
>
> *¿Por qué, al comparar una persona que acertó el ítem i con otra persona que no lo acertó, la primera persona tiene mayor probabilidad de acertar otro ítem i′ ?*

Si el supuesto de independencia local se cumple, entonces la respuesta a las preguntas anteriores es:

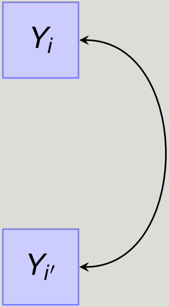
> *La **única** razón es que las personas difieren respecto al valor de su parámetro θ.*

Dentro de un grupo de personas que todos tienen el mismo valor para θ, no habrá correlación entre ítems.

El supuesto de independencia local

Independencia local

Gráficamente:

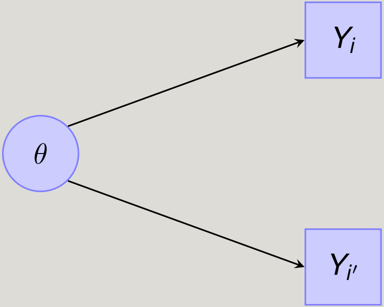


Existe correlación entre Y_i y $Y_{i'}$.

El supuesto de independencia local

Independencia local

Gráficamente:



Al "controlar" por θ, la correlación entre Y_i y $Y_{i'}$ desaparece.

El supuesto de independencia local

Definición

Independencia local en el modelo de Rasch se define formalmente como:

Para cualquier par de ítems i e i' y cualquier valor para θ:

$$\Pr(Y_i = 1|\theta \text{ y } Y_{i'} = 1) = \Pr(Y_i = 1|\theta)$$

o, de forma equivalente:

$$\begin{aligned} \Pr(Y_i = 1 \text{ y } Y_{i'} = 1 \mid \theta) &= \Pr(Y_i = 1 \mid \theta) \times \Pr(Y_{i'} = 1 \mid \theta) \\ &= f_i(\theta) \times f_{i'}(\theta) \end{aligned}$$

La idea anterior se generaliza a más de dos respuestas.

El supuesto de independencia local

El supuesto de independencia local en la TRI.

El supuesto de independencia local es parte de todos los modelos de la TRI.

¿Cuáles pueden ser razones por las que se invalida el supuesto de independencia local?

> *En general, cualquier factor que influye en algunos ítems y en otros no.*

Ejemplos específicos:

- Casos a los que pertenecen múltiples preguntas
- Diferentes formatos de respuesta
- Preguntas que se refieren a diferentes temas (y, entonces a diferentes habilidades latentes)
- ...

El supuesto de independencia local

El supuesto de independencia local en la TRI.

El supuesto de independencia local es parte de todos los modelos de la TRI.

¿Cuáles pueden ser razones por las que se invalida el supuesto de independencia local?

En general, cualquier factor que influye en algunos ítems y en otros no.

Ejemplos específicos:

- Casos a los que pertenecen múltiples preguntas
- Diferentes formatos de respuesta
- Preguntas que se refieren a diferentes temas (y, entonces a diferentes habilidades latentes)
- ...

El supuesto de independencia local

El supuesto de independencia local en la TRI.

El supuesto de independencia local es parte de todos los modelos de la TRI.

¿Cuáles pueden ser razones por las que se invalida el supuesto de independencia local?

> *En general, cualquier factor que influye en algunos ítems y en otros no.*

Ejemplos específicos:

- Casos a los que pertenecen múltiples preguntas
- Diferentes formatos de respuesta
- Preguntas que se refieren a diferentes temas (y, entonces a diferentes habilidades latentes)
- ...

El supuesto de independencia local

El supuesto de independencia local en la TRI.

El supuesto de independencia local es parte de todos los modelos de la TRI.

¿Cuáles pueden ser razones por las que se invalida el supuesto de independencia local?

> *En general, cualquier factor que influye en algunos ítems y en otros no.*

Ejemplos específicos:

- Casos a los que pertenecen múltiples preguntas
- Diferentes formatos de respuesta
- Preguntas que se refieren a diferentes temas (y, entonces a diferentes habilidades latentes)
- ...

El supuesto de independencia local

El supuesto de independencia local en la TRI.

El supuesto de independencia local es parte de todos los modelos de la TRI.

¿Cuáles pueden ser razones por las que se invalida el supuesto de independencia local?

> *En general, cualquier factor que influye en algunos ítems y en otros no.*

Ejemplos específicos:

- Casos a los que pertenecen múltiples preguntas
- Diferentes formatos de respuesta
- Preguntas que se refieren a diferentes temas (y, entonces a diferentes habilidades latentes)
- ...

El supuesto de independencia local

El supuesto de independencia local en la TRI.

El supuesto de independencia local es parte de todos los modelos de la TRI.

¿Cuáles pueden ser razones por las que se invalida el supuesto de independencia local?

> *En general, cualquier factor que influye en algunos ítems y en otros no.*

Ejemplos específicos:

- Casos a los que pertenecen múltiples preguntas
- Diferentes formatos de respuesta
- Preguntas que se refieren a diferentes temas (y, entonces a diferentes habilidades latentes)
- ...

El supuesto de independencia local

El supuesto de independencia local en la TRI.

El supuesto de independencia local es parte de todos los modelos de la TRI.

¿Cuáles pueden ser razones por las que se invalida el supuesto de independencia local?

> *En general, cualquier factor que influye en algunos ítems y en otros no.*

Ejemplos específicos:

- Casos a los que pertenecen múltiples preguntas
- Diferentes formatos de respuesta
- Preguntas que se refieren a diferentes temas (y, entonces a diferentes habilidades latentes)
- ...

El supuesto de suficiencia del número de aciertos para θ

El número de aciertos es un estadístico suficiente para θ

- En el modelo de Rasch, para estimar la θ_p de una persona, es suficiente conocer el número de ítems que acierta.
  Es decir, no importan otros aspectos de su patrón de respuestas.
- Consecuencia: Todas las personas con el mismo número de aciertos tendrán la misma estimación de su θ.

El supuesto de suficiencia del número de aciertos para θ

El número de aciertos es un estadístico suficiente para θ

- En el modelo de Rasch, para estimar la θ_p de una persona, es suficiente conocer el número de ítems que acierta.

  Es decir, no importan otros aspectos de su patrón de respuestas.

- Consecuencia: Todas las personas con el mismo número de aciertos tendrán la misma estimación de su θ.

El supuesto de suficiencia del número de aciertos para θ

El número de aciertos es un estadístico suficiente para θ

- En el modelo de Rasch, para estimar la θ_p de una persona, es suficiente conocer el número de ítems que acierta.

  Es decir, no importan otros aspectos de su patrón de respuestas.

- Consecuencia: Todas las personas con el mismo número de aciertos tendrán la misma estimación de su θ.

Índice

Estimación de parámetros

Estimar los parámetros en el modelo de Rasch

- Después de haber especificado el modelo de Rasch, el siguiente paso es obtener valores para los parámetros del modelo.
- En la estadística, existen varios métodos para estimar parámetros.
- Para los modelos de la Teoría de Respuesta al Ítem, el método más utilizado es Estimación por Máxima Versolimilitud.
  En inglés: *Maximum Likelihood Estimation*.

Estimación de parámetros

Estimar los parámetros en el modelo de Rasch

- Después de haber especificado el modelo de Rasch, el siguiente paso es obtener valores para los parámetros del modelo.
- En la estadística, existen varios métodos para estimar parámetros.
- Para los modelos de la Teoría de Respuesta al Ítem, el método más utilizado es Estimación por Máxima Versolimilitud.
  En inglés: *Maximum Likelihood Estimation*.

Estimación de parámetros

Estimar los parámetros en el modelo de Rasch

- Después de haber especificado el modelo de Rasch, el siguiente paso es obtener valores para los parámetros del modelo.
- En la estadística, existen varios métodos para estimar parámetros.
- Para los modelos de la Teoría de Respuesta al Ítem, el método más utilizado es Estimación por Máxima Versolimilitud.
  En inglés: *Maximum Likelihood Estimation*.

Índice

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

Consideremos el siguiente escenario:

- Disponemos de una moneda trucada y
  nos interesa conocer la probabilidad de que salga “aguila” o “sol” al lanzarla.

- Con este objetivo, realizamos un experimento que consiste en:
  - lanzar 5 veces la moneda y
  - observar cada vez el resultado (“aguila” o “sol”).

- Definimos 5 variables aleatorias: X_1, X_2, X_3, X_4, X_5, donde

$$X_i = \begin{cases} 1 & \text{si en el lanzamiento } i \text{ se observa “aguila”} \\ 0 & \text{si en el lanzamiento } i \text{ se observa “sol”} \end{cases} \qquad (i = 1, \ldots, 5)$$

- Se observan los siguientes datos:

$$\mathbf{x} = (1\ 1\ 0\ 0\ 1)$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

Consideremos el siguiente escenario:

- Disponemos de una moneda trucada y
  nos interesa conocer la probabilidad de que salga “aguila” o “sol” al lanzarla.

- Con este objetivo, realizamos un experimento que consiste en:
  - lanzar 5 veces la moneda y
  - observar cada vez el resultado (“aguila” o “sol”).

- Definimos 5 variables aleatorias: X_1, X_2, X_3, X_4, X_5, donde

$$X_i = \begin{cases} 1 & \text{si en el lanzamiento } i \text{ se observa “aguila”} \\ 0 & \text{si en el lanzamiento } i \text{ se observa “sol”} \end{cases} \qquad (i = 1, \ldots, 5)$$

- Se observan los siguientes datos:

$$\mathbf{x} = (1\ 1\ 0\ 0\ 1)$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

Consideremos el siguiente escenario:

- Disponemos de una moneda trucada y
  nos interesa conocer la probabilidad de que salga “aguila” o “sol” al lanzarla.
- Con este objetivo, realizamos un experimento que consiste en:
  - lanzar 5 veces la moneda y
  - observar cada vez el resultado (“aguila” o “sol”).
- Definimos 5 variables aleatorias: X_1, X_2, X_3, X_4, X_5, donde

$$X_i = \begin{cases} 1 & \text{si en el lanzamiento } i \text{ se observa “aguila”} \\ 0 & \text{si en el lanzamiento } i \text{ se observa “sol”} \end{cases} \qquad (i = 1, \ldots, 5)$$

- Se observan los siguientes datos:

$$\mathbf{x} = (1\ 1\ 0\ 0\ 1)$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

Consideremos el siguiente escenario:

- Disponemos de una moneda trucada y
  nos interesa conocer la probabilidad de que salga “aguila” o “sol” al lanzarla.

- Con este objetivo, realizamos un experimento que consiste en:
  - lanzar 5 veces la moneda y
  - observar cada vez el resultado (“aguila” o “sol”).

- Definimos 5 variables aleatorias: X_1, X_2, X_3, X_4, X_5, donde

$$X_i = \begin{cases} 1 & \text{si en el lanzamiento } i \text{ se observa “aguila”} \\ 0 & \text{si en el lanzamiento } i \text{ se observa “sol”} \end{cases} \qquad (i = 1, \ldots, 5)$$

- Se observan los siguientes datos:

$$\mathbf{x} = (1\ 1\ 0\ 0\ 1)$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- Para lograr el objetivo (estimar la probabilidad de que salga “aguila” o “sol” al lanzar la moneda), especificamos un modelo estadístico, con los siguientes supuestos:
  - la probabilidad de que salga aguila es la misma en cada lanzamiento;
  - esta probabilidad es igual a π (donde π es un número entre 0 y 1).
  - los resultados de los 5 lanzamientos en el experimento son independientes.
- Bajo los supuestos de este modelo, podemos calcular la probabilidad del resultado (1 1 0 0 1):

$$\begin{aligned}\Pr(X_1 = 1 \text{ y } X_2 = 1 \text{ y } X_3 = 0 \text{ y } X_4 = 0 \text{ y } X_5 = 1) &= \pi\,\pi\,(1-\pi)\,(1-\pi)\,\pi \\ &= \pi^3\,(1-\pi)^2\end{aligned}$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- Para lograr el objetivo (estimar la probabilidad de que salga “aguila” o “sol” al lanzar la moneda), especificamos un modelo estadístico, con los siguientes supuestos:
  - la probabilidad de que salga aguila es la misma en cada lanzamiento;
  - esta probabilidad es igual a π (donde π es un número entre 0 y 1).
  - los resultados de los 5 lanzamientos en el experimento son independientes.
- Bajo los supuestos de este modelo, podemos calcular la probabilidad del resultado (1 1 0 0 1):

$$\begin{aligned}\Pr(X_1 = 1 \text{ y } X_2 = 1 \text{ y } X_3 = 0 \text{ y } X_4 = 0 \text{ y } X_5 = 1) &= \pi\, \pi\, (1-\pi)\, (1-\pi)\, \pi \\ &= \pi^3\, (1-\pi)^2\end{aligned}$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- Para lograr el objetivo (estimar la probabilidad de que salga “aguila” o “sol” al lanzar la moneda), especificamos un modelo estadístico, con los siguientes supuestos:
  - la probabilidad de que salga aguila es la misma en cada lanzamiento;
  - esta probabilidad es igual a π (donde π es un número entre 0 y 1).
  - los resultados de los 5 lanzamientos en el experimento son independientes.
- Bajo los supuestos de este modelo, podemos calcular la probabilidad del resultado (1 1 0 0 1):

$$\begin{aligned}\Pr(X_1 = 1 \text{ y } X_2 = 1 \text{ y } X_3 = 0 \text{ y } X_4 = 0 \text{ y } X_5 = 1) &= \pi\,\pi\,(1-\pi)\,(1-\pi)\,\pi \\ &= \pi^3\,(1-\pi)^2\end{aligned}$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- Para lograr el objetivo (estimar la probabilidad de que salga "aguila" o "sol" al lanzar la moneda), especificamos un modelo estadístico, con los siguientes supuestos:
  - la probabilidad de que salga aguila es la misma en cada lanzamiento;
  - esta probabilidad es igual a π (donde π es un número entre 0 y 1).
  - los resultados de los 5 lanzamientos en el experimento son independientes.
- Bajo los supuestos de este modelo, podemos calcular la probabilidad del resultado (1 1 0 0 1):

$$\begin{aligned}\Pr(X_1 = 1 \text{ y } X_2 = 1 \text{ y } X_3 = 0 \text{ y } X_4 = 0 \text{ y } X_5 = 1) &= \pi\,\pi\,(1-\pi)\,(1-\pi)\,\pi \\ &= \pi^3\,(1-\pi)^2\end{aligned}$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- Para lograr el objetivo (estimar la probabilidad de que salga “aguila” o “sol” al lanzar la moneda), especificamos un modelo estadístico, con los siguientes supuestos:
  - la probabilidad de que salga aguila es la misma en cada lanzamiento;
  - esta probabilidad es igual a π (donde π es un número entre 0 y 1).
  - los resultados de los 5 lanzamientos en el experimento son independientes.

- Bajo los supuestos de este modelo, podemos calcular la probabilidad del resultado (1 1 0 0 1):

$$\begin{aligned}\Pr(X_1 = 1 \text{ y } X_2 = 1 \text{ y } X_3 = 0 \text{ y } X_4 = 0 \text{ y } X_5 = 1) &= \pi\,\pi\,(1-\pi)\,(1-\pi)\,\pi \\ &= \pi^3\,(1-\pi)^2\end{aligned}$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- En la expresión

$$\Pr(X_1 = 1 \text{ y } X_2 = 1 \text{ y } X_3 = 0 \text{ y } X_4 = 0 \text{ y } X_5 = 1) \; = \; \pi^3\,(1-\pi)^2,$$

π es un parámetro del modelo: Un número fijo entre 0 y 1 (aunque desconocido).

- Aunque π es un desconocido, podemos evaluar cuál sería la probabilidad del resultado observado si π fuera igual a cierto valor.

Por ejemplo:

$$\text{Si } \pi = 0.1, \text{ entonces } \Pr\left[\mathbf{X} = (1\,1\,0\,0\,1)\right] \; = \; 0.10^3 \times 0.90^2 \; = \; 0.00081$$

$$\text{Si } \pi = 0.2, \text{ entonces } \Pr\left[\mathbf{X} = (1\,1\,0\,0\,1)\right] \; = \; 0.20^3 \times 0.80^2 \; = \; 0.00512$$

$$\vdots$$

Esta probabilidad indica qué tan verosimil es este valor de π a la luz de los datos observados:

$$\ell\left[\pi; (1\,1\,0\,0\,1)\right] \; = \; \pi^3\,(1-\pi)^2$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- En la expresión

$$\Pr(X_1 = 1 \text{ y } X_2 = 1 \text{ y } X_3 = 0 \text{ y } X_4 = 0 \text{ y } X_5 = 1) \;=\; \pi^3\,(1-\pi)^2,$$

π es un parámetro del modelo: Un número fijo entre 0 y 1 (aunque desconocido).

- Aunque π es un desconocido, podemos evaluar cuál sería la probabilidad del resultado observado si π fuera igual a cierto valor.

  Por ejemplo:

$$\text{Si } \pi = 0.1, \text{ entonces } \Pr\left[\mathbf{X} = (1\,1\,0\,0\,1)\right] \;=\; 0.10^3 \;\times\; 0.90^2 \;=\; 0.00081$$
$$\text{Si } \pi = 0.2, \text{ entonces } \Pr\left[\mathbf{X} = (1\,1\,0\,0\,1)\right] \;=\; 0.20^3 \;\times\; 0.80^2 \;=\; 0.00512$$
$$\vdots$$

  Esta probabilidad indica qué tan verosimil es este valor de π a la luz de los datos observados:

$$\ell\left[\pi; (1\,1\,0\,0\,1)\right] \;=\; \pi^3\,(1-\pi)^2$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- En la expresión

$$\Pr(X_1 = 1 \text{ y } X_2 = 1 \text{ y } X_3 = 0 \text{ y } X_4 = 0 \text{ y } X_5 = 1) \;=\; \pi^3\,(1-\pi)^2,$$

  π es un parámetro del modelo: Un número fijo entre 0 y 1 (aunque desconocido).

- Aunque π es un desconocido, podemos evaluar cuál sería la probabilidad del resultado observado si π fuera igual a cierto valor.

  Por ejemplo:

$$\text{Si } \pi = 0.1,\ \text{entonces } \Pr\left[\mathbf{X} = (1\,1\,0\,0\,1)\right] \;=\; 0.10^3 \times 0.90^2 \;=\; 0.00081$$
$$\text{Si } \pi = 0.2,\ \text{entonces } \Pr\left[\mathbf{X} = (1\,1\,0\,0\,1)\right] \;=\; 0.20^3 \times 0.80^2 \;=\; 0.00512$$
$$\vdots$$

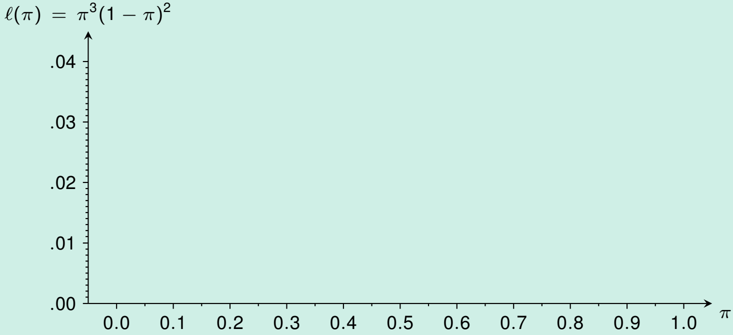
  Esta probabilidad indica qué tan verosimil es este valor de π a la luz de los datos observados:

$$\ell\left[\pi; (1\,1\,0\,0\,1)\right] \;=\; \pi^3\,(1-\pi)^2$$

El principio de máxima verosimilitud

Ejemplo: la moneda trucada

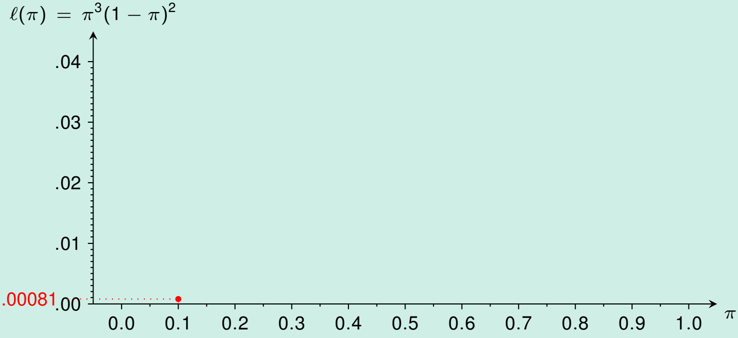
- En principio, podemos “probar” y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

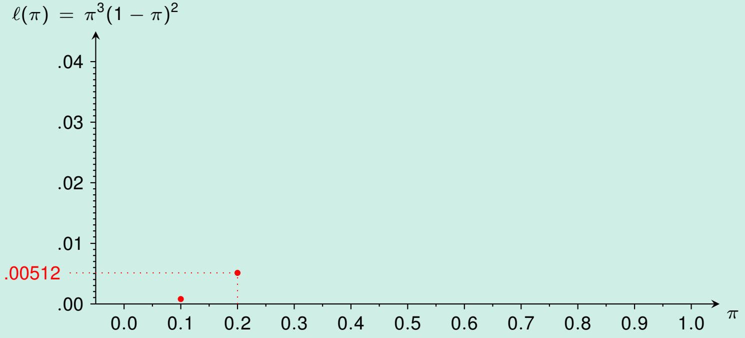
- En principio, podemos "probar" y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

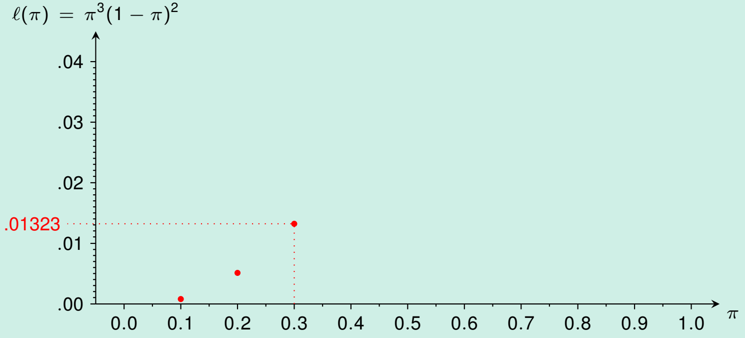
- En principio, podemos “probar” y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

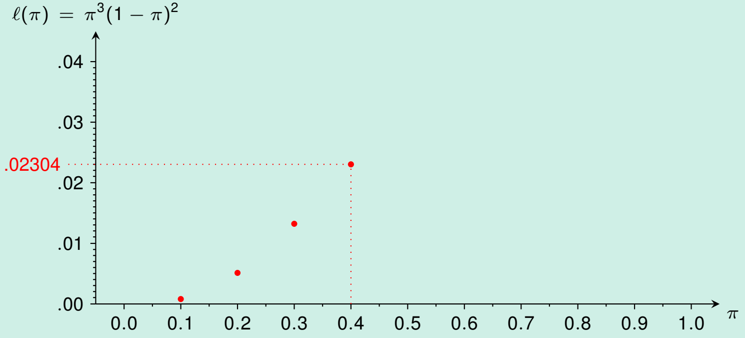
- En principio, podemos "probar" y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

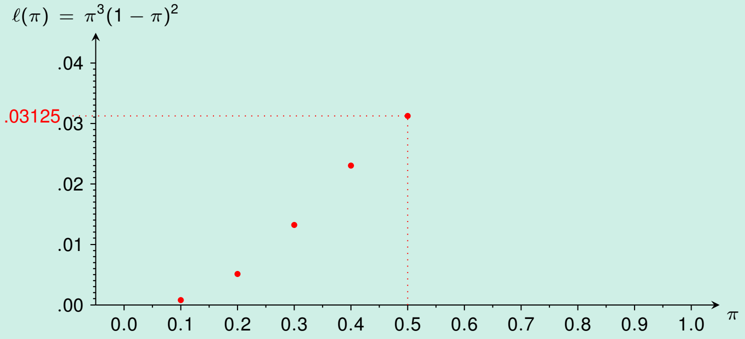
- En principio, podemos "probar" y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- En principio, podemos “probar” y graficar la verosimilitud de todos los posibles valores para π:

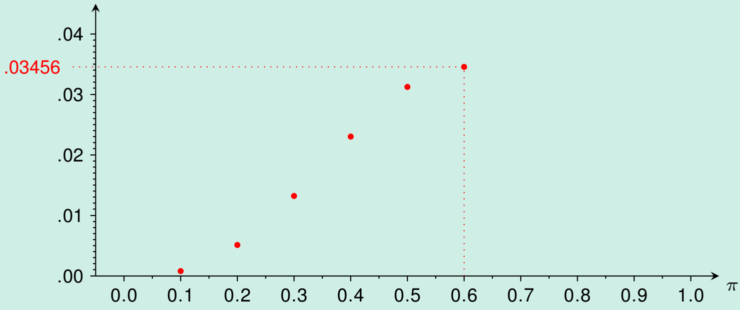


El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- En principio, podemos “probar” y graficar la verosimilitud de todos los posibles valores para π:

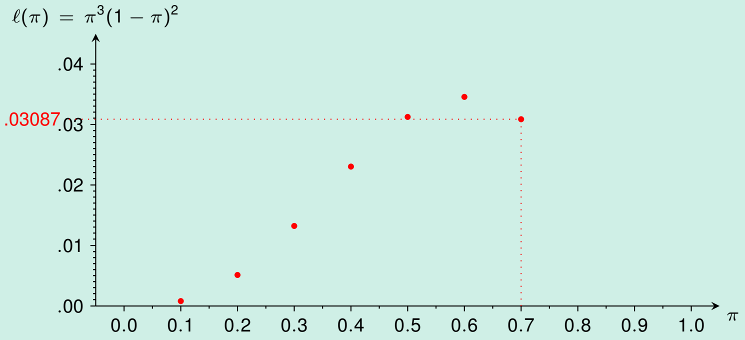
$$\ell(\pi) = \pi^3(1-\pi)^2$$



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

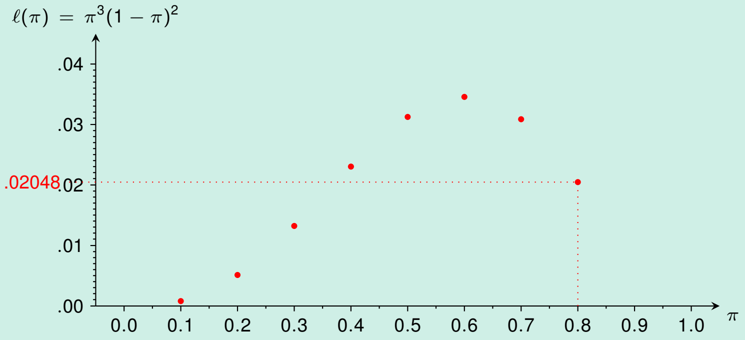
- En principio, podemos "probar" y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

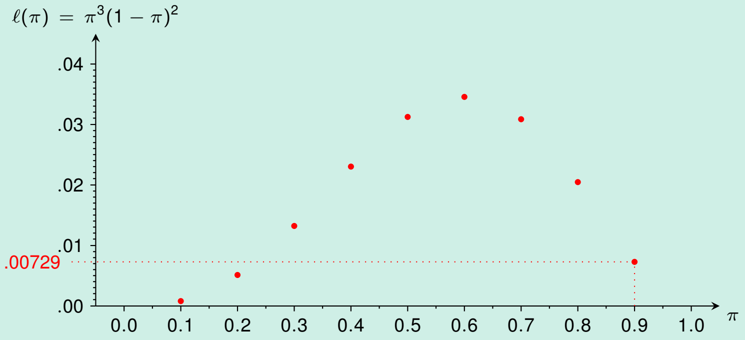
- En principio, podemos “probar” y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

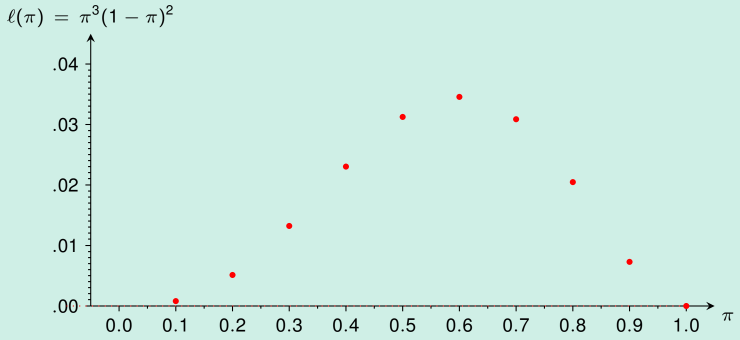
- En principio, podemos "probar" y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

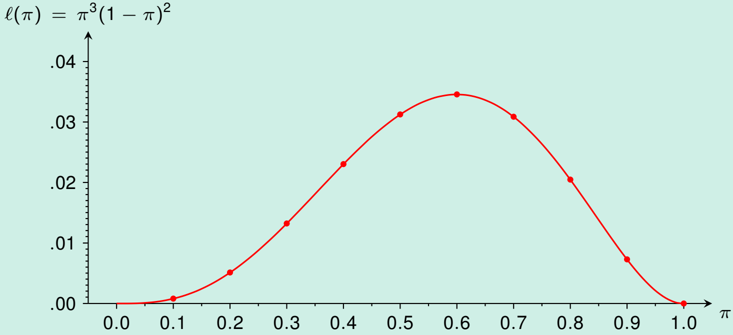
- En principio, podemos “probar” y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

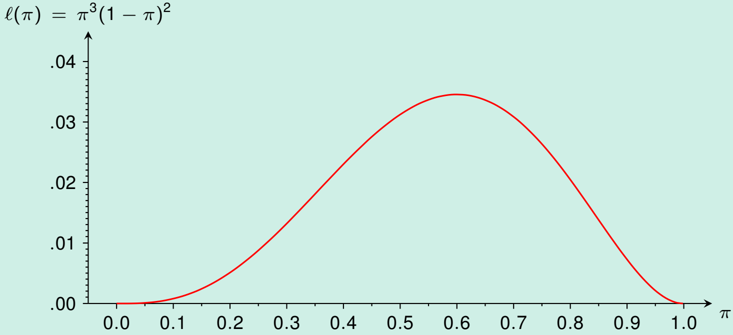
- En principio, podemos "probar" y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

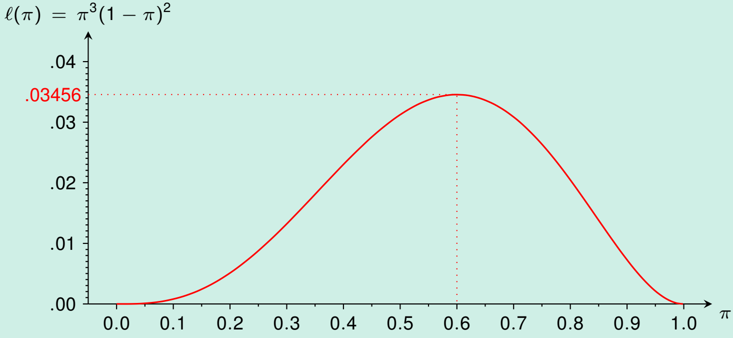
- En principio, podemos “probar” y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

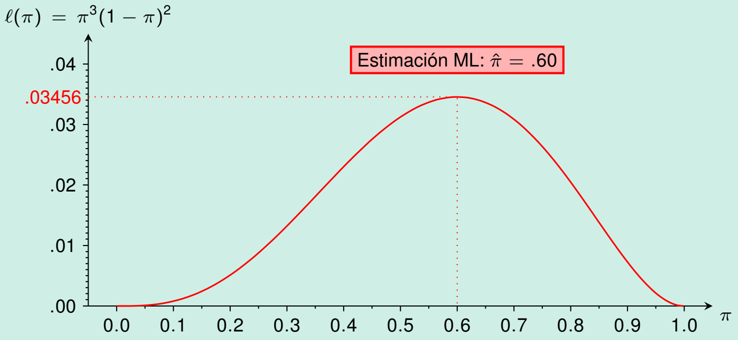
- En principio, podemos "probar" y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Ejemplo: la moneda trucada

- En principio, podemos "probar" y graficar la verosimilitud de todos los posibles valores para π:



El principio de máxima verosimilitud

Definición

La estimación por máxima verosimilitud asigna aquel valor al parámetro para el cual la probabilidad de los datos observados sea máxima.

El principio de máxima verosimilitud

Nota: Buscar el máximo de una función

- En general, en el ejemplo anterior, si el experimento consistiera en n lanzamientos, la función de verosimilitud de π a la luz de un resultado observado $\mathbf{x}$ sería:

$$\ell(\pi; \mathbf{x}) \;=\; \pi^{y}\,(1-\pi)^{n-y},$$

donde y es el número de águilas en el resultado observado.

- ¿Cómo se encuentra el máximo de una función?

  En la matemática, y particularmente el cálculo diferencial, se ha mostrado que:

  La función $\ell(\pi; \mathbf{x})$ llega a un máximo en aquel valor de π para el cual:
  - la primera derivada evaluada en este valor de π sea igual a 0;
  - la segunda derivada evaluada en este valor de π sea menor que 0.

- En la práctica resulta más conveniente buscar el máximo del logaritmo de la función de verosimilitud: $\log \ell(\pi; \mathbf{x})$.

  Recuerda que el logaritmo es una transformación monónota. Esto implica que:

  $\ell(\pi; \mathbf{x})$ y $\log \ell(\pi; \mathbf{x})$ llegan a su máximo en el mismo punto.

El principio de máxima verosimilitud

Nota: Buscar el máximo de una función

- En general, en el ejemplo anterior, si el experimento consistiera en n lanzamientos, la función de verosimilitud de π a la luz de un resultado observado $\mathbf{x}$ sería:

$$\ell(\pi; \mathbf{x}) \;=\; \pi^{y}\,(1-\pi)^{n-y},$$

  donde y es el número de águilas en el resultado observado.

- ¿Cómo se encuentra el máximo de una función?

  En la matemática, y particularmente el cálculo diferencial, se ha mostrado que:

  La función $\ell(\pi; \mathbf{x})$ llega a un máximo en aquel valor de π para el cual:

  - la primera derivada evaluada en este valor de π sea igual a 0;
  - la segunda derivada evaluada en este valor de π sea menor que 0.

- En la práctica resulta más conveniente buscar el máximo del logaritmo de la función de verosimilitud: $\log \ell(\pi; \mathbf{x})$.

  Recuerda que el logaritmo es una transformación monónota. Esto implica que:

  $\ell(\pi; \mathbf{x})$ y $\log \ell(\pi; \mathbf{x})$ llegan a su máximo en el mismo punto.

El principio de máxima verosimilitud

Nota: Buscar el máximo de una función

- En general, en el ejemplo anterior, si el experimento consistiera en n lanzamientos, la función de verosimilitud de π a la luz de un resultado observado $\mathbf{x}$ sería:

$$\ell(\pi; \mathbf{x}) = \pi^y (1 - \pi)^{n-y},$$

  donde y es el número de águilas en el resultado observado.

- ¿Cómo se encuentra el máximo de una función?

  En la matemática, y particularmente el cálculo diferencial, se ha mostrado que:

  La función $\ell(\pi; \mathbf{x})$ llega a un máximo en aquel valor de π para el cual:
  - la primera derivada evaluada en este valor de π sea igual a 0;
  - la segunda derivada evaluada en este valor de π sea menor que 0.

- En la práctica resulta más conveniente buscar el máximo del logaritmo de la función de verosimilitud: $\log \ell(\pi; \mathbf{x})$.

  Recuerda que el logaritmo es una transformación monónota. Esto implica que:

  $\ell(\pi; \mathbf{x})$ y $\log \ell(\pi; \mathbf{x})$ llegan a su máximo en el mismo punto.

El principio de máxima verosimilitud

Nota: Buscar el máximo de una función

- En general, en el ejemplo anterior, si el experimento consistiera en *n* lanzamientos, la función de verosimilitud de π a la luz de un resultado observado $\mathbf{x}$ sería:

  $$\ell(\pi; \mathbf{x}) \;=\; \pi^{y}\,(1-\pi)^{n-y},$$

  donde *y* es el número de águilas en el resultado observado.

- ¿Cómo se encuentra el máximo de una función?

  En la matemática, y particularmente el cálculo diferencial, se ha mostrado que:

  La función $\ell(\pi; \mathbf{x})$ llega a un máximo en aquel valor de π para el cual:
  - la primera derivada evaluada en este valor de π sea igual a 0;
  - la segunda derivada evaluada en este valor de π sea menor que 0.

- En la práctica resulta más conveniente buscar el máximo del logaritmo de la función de verosimilitud: $\log \ell(\pi; \mathbf{x})$.

  Recuerda que el logaritmo es una transformación monónota. Esto implica que:
  $\ell(\pi; \mathbf{x})$ y $\log \ell(\pi; \mathbf{x})$ llegan a su máximo en el mismo punto.

El principio de máxima verosimilitud

Ejemplo: Buscar el máximo de una función

Por lo anterior, buscar el máximo de la función de verosimilitud implica los siguientes pasos:

1. Obtener el logaritmo de la función de verosimilitud:

$$\begin{aligned} \log \ell(\pi; \mathbf{x}) &= \log \left[\pi^y (1-\pi)^{n-y} \right] \\ &= y \log \pi + (n-y) \log(1-\pi) \end{aligned}$$

2. Obtener la primera derivada del logaritmo de la función de verosimilitud:

$$\frac{\mathrm{d} \log \ell(\pi; \mathbf{x})}{\mathrm{d}\pi} = \frac{y}{\pi} - \frac{n-y}{1-\pi}$$

3. Encontrar el valor de π para el cual la primera derivada es 0:

$$\begin{aligned} \frac{\mathrm{d} \log \ell(\pi; \mathbf{x})}{\mathrm{d}\pi} &= 0 \\ \iff \quad \frac{y}{\pi} - \frac{n-y}{1-\pi} &= 0 \\ \iff \quad \pi &= \frac{y}{n} \end{aligned}$$

El principio de máxima verosimilitud

Ejemplo: Buscar el máximo de una función

Por lo anterior, buscar el máximo de la función de verosimilitud implica los siguientes pasos:

1. Obtener el logaritmo de la función de verosimilitud:

$$\begin{aligned} \log \ell(\pi; \mathbf{x}) &= \log \left[\pi^{y} (1-\pi)^{n-y} \right] \\ &= y \log \pi + (n-y) \log(1-\pi) \end{aligned}$$

2. Obtener la primera derivada del logaritmo de la función de verosimilitud:

$$\frac{\mathrm{d} \log \ell(\pi; \mathbf{x})}{\mathrm{d}\pi} = \frac{y}{\pi} - \frac{n-y}{1-\pi}$$

3. Encontrar el valor de π para el cual la primera derivada es 0:

$$\begin{aligned} \frac{\mathrm{d} \log \ell(\pi; \mathbf{x})}{\mathrm{d}\pi} &= 0 \\ \Longleftrightarrow \quad \frac{y}{\pi} - \frac{n-y}{1-\pi} &= 0 \\ \Longleftrightarrow \quad \pi &= \frac{y}{n} \end{aligned}$$

El principio de máxima verosimilitud

Ejemplo: Buscar el máximo de una función

Por lo anterior, buscar el máximo de la función de verosimilitud implica los siguientes pasos:

1. Obtener el logaritmo de la función de verosimilitud:

$$\begin{aligned}\log \ell(\pi; \mathbf{x}) &= \log \left[\pi^{y} (1-\pi)^{n-y}\right] \\ &= y \log \pi + (n-y) \log(1-\pi)\end{aligned}$$

2. Obtener la primera derivada del logaritmo de la función de verosimilitud:

$$\frac{\mathrm{d} \log \ell(\pi; \mathbf{x})}{\mathrm{d}\pi} = \frac{y}{\pi} - \frac{n-y}{1-\pi}$$

3. Encontrar el valor de π para el cual la primera derivada es 0:

$$\begin{aligned} \frac{\mathrm{d} \log \ell(\pi; \mathbf{x})}{\mathrm{d}\pi} &= 0 \\ \Longleftrightarrow \quad \frac{y}{\pi} - \frac{n-y}{1-\pi} &= 0 \\ \Longleftrightarrow \quad \pi &= \frac{y}{n}\end{aligned}$$

El principio de máxima verosimilitud

Ejemplo: Buscar el máximo de una función (continuación)

4. Obtener la segunda derivada del logaritmo de la función de verosimilitud:

$$\frac{\mathrm{d}^2 \log \ell(\pi; \mathbf{x})}{\mathrm{d}\pi^2} = -\frac{y}{\pi^2} - \frac{n-y}{(1-\pi)^2}$$

5. Verificar que la segunda derivada, cuando se evalua para $\pi = \frac{y}{n}$, es negativa.

   Efectivamente, si $\pi = \frac{y}{n}$, entonces

$$-\frac{y}{\pi^2} - \frac{n-y}{(1-\pi)^2} = -\frac{n^2}{y} - \frac{n^2}{n-y},$$

   lo cual es negativo para cualquier valor de y que cumple $0 < y < n$.

El principio de máxima verosimilitud

Ejemplo: Buscar el máximo de una función (continuación)

4. Obtener la segunda derivada del logaritmo de la función de verosimilitud:

$$\frac{\mathrm{d}^2 \log \ell(\pi; \mathbf{x})}{\mathrm{d}\pi^2} = -\frac{y}{\pi^2} - \frac{n-y}{(1-\pi)^2}$$

5. Verificar que la segunda derivada, cuando se evalua para $\pi = \frac{y}{n}$, es negativa.

   Efectivamente, si $\pi = \frac{y}{n}$, entonces

$$-\frac{y}{\pi^2} - \frac{n-y}{(1-\pi)^2} = -\frac{n^2}{y} - \frac{n^2}{n-y},$$

   lo cual es negativo para cualquier valor de y que cumple $0 < y < n$.

El principio de máxima verosimilitud

Ejemplo: Buscar el máximo de una función

Conclusión:

- En este ejemplo, $\frac{y}{n}$ es el estimador por máxima verosimilitud del parámetro π.

  Se escribe:

$$\hat{\pi} = \frac{y}{n}$$

Índice

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

Consideremos el siguiente escenario:

- Una persona p responde una prueba que consiste en 4 ítems.
- La puntuación en cada ítem es 0 (fallar) o 1 (acertar).
- El modelo de Rasch se cumple para las respuestas de esta persona en estos ítems.
- Conocemos los valores de los parámetros β_1, β_2, β_3 y β_4.

Objetivo: Estimar el parámetro θ_p de esta persona a partir de sus respuestas en los 4 ítems.

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

Consideremos el siguiente escenario:

- Una persona p responde una prueba que consiste en 4 ítems.
- La puntuación en cada ítem es 0 (fallar) o 1 (acertar).
- El modelo de Rasch se cumple para las respuestas de esta persona en estos ítems.
- Conocemos los valores de los parámetros β_1, β_2, β_3 y β_4.

Objetivo: Estimar el parámetro θ_p de esta persona a partir de sus respuestas en los 4 ítems.

	Ítems			
	1	2	3	4
Persona p				

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

Consideremos el siguiente escenario:

- Una persona p responde una prueba que consiste en 4 ítems.
- La puntuación en cada ítem es 0 (fallar) o 1 (acertar).
- El modelo de Rasch se cumple para las respuestas de esta persona en estos ítems.
- Conocemos los valores de los parámetros β_1, β_2, β_3 y β_4.

Objetivo: Estimar el parámetro θ_p de esta persona a partir de sus respuestas en los 4 ítems.

	Ítems			
	1	2	3	4
Persona p	0	1	1	0

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

Consideremos el siguiente escenario:

- Una persona p responde una prueba que consiste en 4 ítems.
- La puntuación en cada ítem es 0 (fallar) o 1 (acertar).
- El modelo de Rasch se cumple para las respuestas de esta persona en estos ítems.
- Conocemos los valores de los parámetros β_1, β_2, β_3 y β_4.

Objetivo: Estimar el parámetro θ_p de esta persona a partir de sus respuestas en los 4 ítems.

	Ítems			
	1	2	3	4
Persona p	0	1	1	0

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

Consideremos el siguiente escenario:

- Una persona p responde una prueba que consiste en 4 ítems.
- La puntuación en cada ítem es 0 (fallar) o 1 (acertar).
- El modelo de Rasch se cumple para las respuestas de esta persona en estos ítems.
- Conocemos los valores de los parámetros β_1, β_2, β_3 y β_4.

Objetivo: Estimar el parámetro θ_p de esta persona a partir de sus respuestas en los 4 ítems.

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

Consideremos el siguiente escenario:

- Una persona p responde una prueba que consiste en 4 ítems.
- La puntuación en cada ítem es 0 (fallar) o 1 (acertar).
- El modelo de Rasch se cumple para las respuestas de esta persona en estos ítems.
- Conocemos los valores de los parámetros β_1, β_2, β_3 y β_4.

Objetivo: Estimar el parámetro θ_p de esta persona a partir de sus respuestas en los 4 ítems.

	Ítems				
	1	2	3	4	
Persona p	0	1	1	0	$\theta_p = ?$
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$	

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems				
	1	2	3	4	
Persona p	0	1	1	0	$\theta_p = ?$
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$	

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \;\text{ y }\; Y_2 = 1 \;\text{ y }\; Y_3 = 1 \;\text{ y }\; Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems				
	1	2	3	4	
Persona *p*	0	1	1	0	$\theta_p = ?$
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$	

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \ \text{ y } \ Y_2 = 1 \ \text{ y } \ Y_3 = 1 \ \text{ y } \ Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

$$\ell(\theta_p; \mathbf{y}) = \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4)$$

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems				
	1	2	3	4	
Persona p	0	1	1	0	$\theta_p = ?$
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$	

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \text{ y } Y_2 = 1 \text{ y } Y_3 = 1 \text{ y } Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

$$\begin{aligned}\ell(\theta_p; \mathbf{y}) &= \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4) \\ &= \frac{1}{1 + e^{\theta_p - 2.0}} \times \frac{e^{\theta_p - 0.0}}{1 + e^{\theta_p - 0.0}} \times \frac{e^{\theta_p - 1.0}}{1 + e^{\theta_p - 1.0}} \times \frac{1}{1 + e^{\theta_p - 0.5}}\end{aligned}$$

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems				
	1	2	3	4	
Persona p	0	1	1	0	$\theta_p = ?$
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$	

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \ \text{ y } \ Y_2 = 1 \ \text{ y } \ Y_3 = 1 \ \text{ y } \ Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

$$\begin{aligned}\ell(\theta_p; \mathbf{y}) &= \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4) \\ &= \frac{1}{1 + e^{\theta_p - 2.0}} \times \frac{e^{\theta_p - 0.0}}{1 + e^{\theta_p - 0.0}} \times \frac{e^{\theta_p - 1.0}}{1 + e^{\theta_p - 1.0}} \times \frac{1}{1 + e^{\theta_p - 0.5}}\end{aligned}$$

Probemos diferentes valores para θ_p.

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$ $\rightarrow$ $\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \text{ y } Y_2 = 1 \text{ y } Y_3 = 1 \text{ y } Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

$$\begin{aligned}
\ell(\theta_p; \mathbf{y}) &= \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4) \\
&= \frac{1}{1 + e^{0.0 - 2.0}} \times \frac{e^{0.0 - 0.0}}{1 + e^{0.0 - 0.0}} \times \frac{e^{0.0 - 1.0}}{1 + e^{0.0 - 1.0}} \times \frac{1}{1 + e^{0.0 - 0.5}} \\
&= .881 \times .500 \times .269 \times .622 \\
&= .074
\end{aligned}$$

Probemos diferentes valores para θ_p. Por ejemplo, $\theta_p = 0$.

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \text{ y } Y_2 = 1 \text{ y } Y_3 = 1 \text{ y } Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

$$\begin{aligned}
\ell(\theta_p; \mathbf{y}) &= \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4) \\
&= \frac{1}{1 + e^{0.5 - 2.0}} \times \frac{e^{0.5 - 0.0}}{1 + e^{0.5 - 0.0}} \times \frac{e^{0.5 - 1.0}}{1 + e^{0.5 - 1.0}} \times \frac{1}{1 + e^{0.5 - 0.5}} \\
&= .818 \times .622 \times .378 \times .500 \\
&= .096
\end{aligned}$$

Probemos diferentes valores para θ_p. Por ejemplo, $\theta_p = 0.5$.

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$

$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \text{ y } Y_2 = 1 \text{ y } Y_3 = 1 \text{ y } Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

$$\begin{aligned}
\ell(\theta_p; \mathbf{y}) &= \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4) \\
&= \frac{1}{1 + e^{1.0 - 2.0}} \times \frac{e^{1.0 - 0.0}}{1 + e^{1.0 - 0.0}} \times \frac{e^{1.0 - 1.0}}{1 + e^{1.0 - 1.0}} \times \frac{1}{1 + e^{1.0 - 0.5}} \\
&= .731 \times .731 \times .500 \times .378 \\
&= .101
\end{aligned}$$

Probemos diferentes valores para θ_p. Por ejemplo, $\theta_p = 1$.

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$

$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$

$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \text{ y } Y_2 = 1 \text{ y } Y_3 = 1 \text{ y } Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

$$\begin{aligned}
\ell(\theta_p; \mathbf{y}) &= \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4) \\
&= \frac{1}{1 + e^{1.5-2.0}} \times \frac{e^{1.5-0.0}}{1 + e^{1.5-0.0}} \times \frac{e^{1.5-1.0}}{1 + e^{1.5-1.0}} \times \frac{1}{1 + e^{1.5-0.5}} \\
&= .622 \times .818 \times .622 \times .269 \\
&= .085
\end{aligned}$$

Probemos diferentes valores para θ_p. Por ejemplo, $\theta_p = 1.5$.

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$

$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$

$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$

$\theta_p = 2.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .059$

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \text{ y } Y_2 = 1 \text{ y } Y_3 = 1 \text{ y } Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

$$\begin{aligned}
\ell(\theta_p; \mathbf{y}) &= \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4) \\
&= \frac{1}{1 + e^{2.0 - 2.0}} \times \frac{e^{2.0 - 0.0}}{1 + e^{2.0 - 0.0}} \times \frac{e^{2.0 - 1.0}}{1 + e^{2.0 - 1.0}} \times \frac{1}{1 + e^{2.0 - 0.5}} \\
&= .500 \times .881 \times .731 \times .182 \\
&= .059
\end{aligned}$$

Probemos diferentes valores para θ_p. Por ejemplo, $\theta_p = 2.0$.

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$

$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$

$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$

$\theta_p = 2.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .059$

$\theta_p = 2.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .034$

La probabilidad de este patrón de respuestas es:

$$\Pr(Y_1 = 0 \ \text{ y } \ Y_2 = 1 \ \text{ y } \ Y_3 = 1 \ \text{ y } \ Y_4 = 0 \mid \theta_p, \beta_1, \ldots, \beta_4) \equiv \ell(\theta_p; \mathbf{y})$$

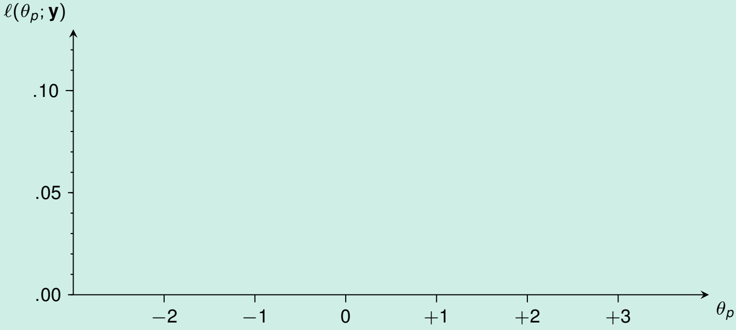
$$\begin{aligned}
\ell(\theta_p; \mathbf{y}) &= \Pr(Y_1 = 0 \mid \theta_p, \beta_1) \times \Pr(Y_2 = 1 \mid \theta_p, \beta_2) \times \Pr(Y_3 = 1 \mid \theta_p, \beta_3) \times \Pr(Y_4 = 0 \mid \theta_p, \beta_4) \\
&= \frac{1}{1 + e^{2.5-2.0}} \times \frac{e^{2.5-0.0}}{1 + e^{2.5-0.0}} \times \frac{e^{2.5-1.0}}{1 + e^{2.5-1.0}} \times \frac{1}{1 + e^{2.5-0.5}} \\
&= .378 \times .924 \times .818 \times .119 \\
&= .034
\end{aligned}$$

Probemos diferentes valores para θ_p. Por ejemplo, $\theta_p = 2.5$.

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems				
	1	2	3	4	
Persona *p*	0	1	1	0	$\theta_p = ?$
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$	



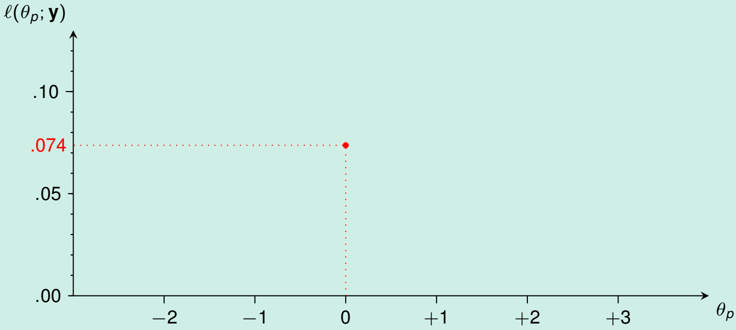
Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$



Estimación por máxima verosimilitud del parámetro θ_p

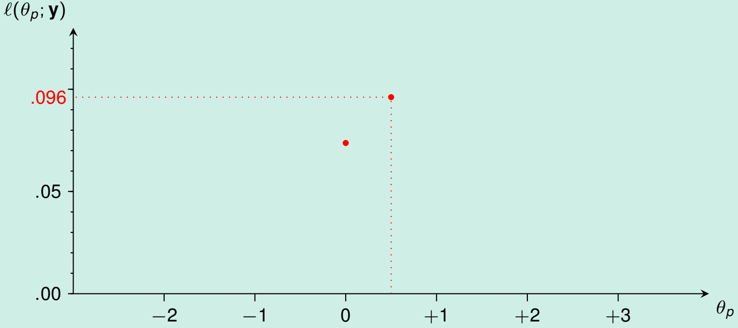
Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$



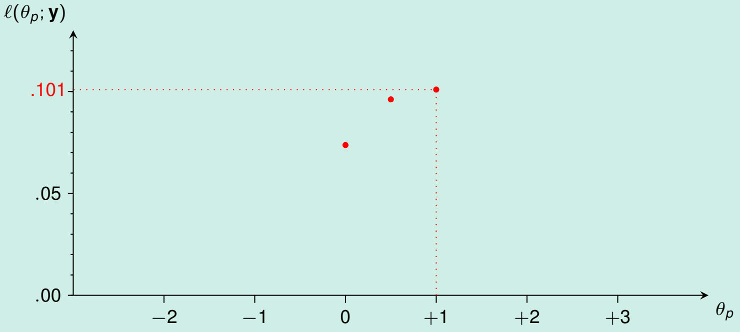
Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$
$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$
$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$



Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

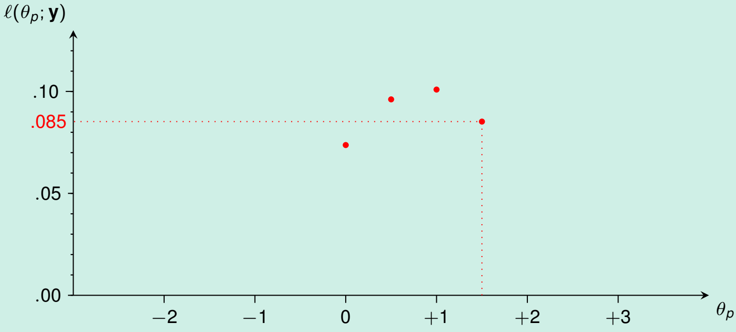
$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$

$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$

$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$



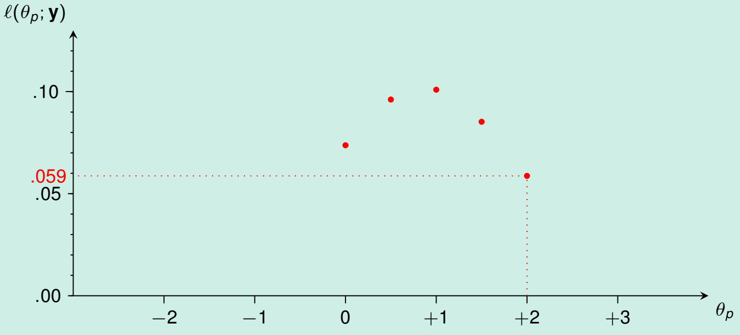
Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$
$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$
$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$
$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$
$\theta_p = 2.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .059$



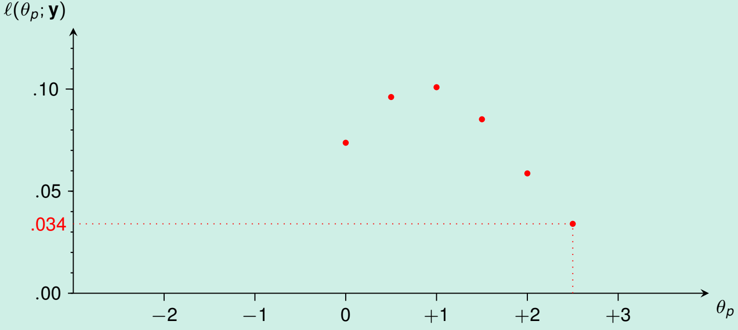
Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$$
$$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$$
$$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$$
$$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$$
$$\theta_p = 2.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .059$$
$$\theta_p = 2.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .034$$



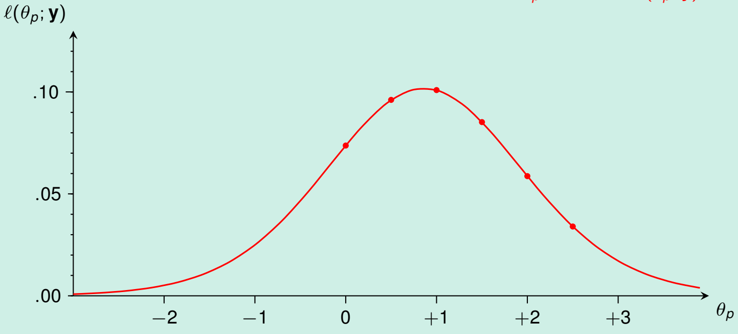
Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$
$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$
$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$
$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$
$\theta_p = 2.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .059$
$\theta_p = 2.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .034$



Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$

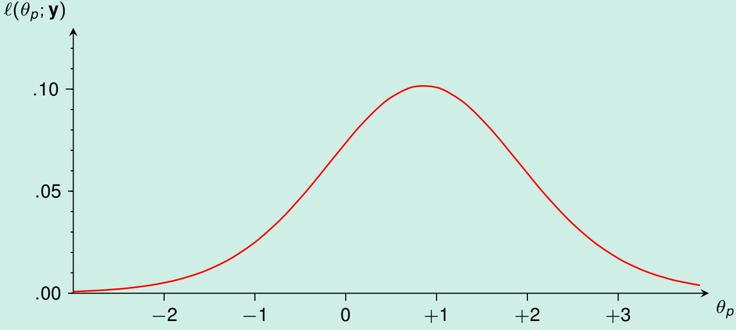
$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$

$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$

$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$

$\theta_p = 2.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .059$

$\theta_p = 2.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .034$



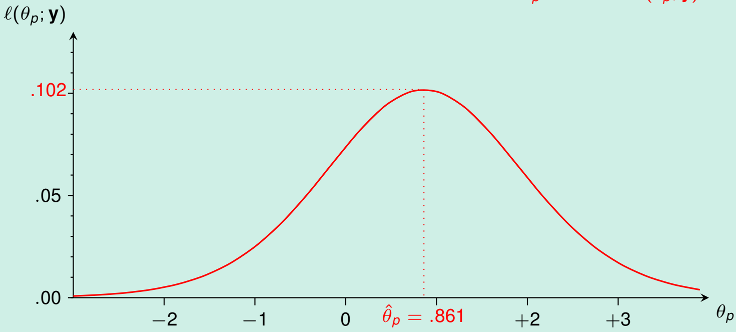
Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems			
	1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$
$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$
$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$
$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$
$\theta_p = 2.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .059$
$\theta_p = 2.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .034$



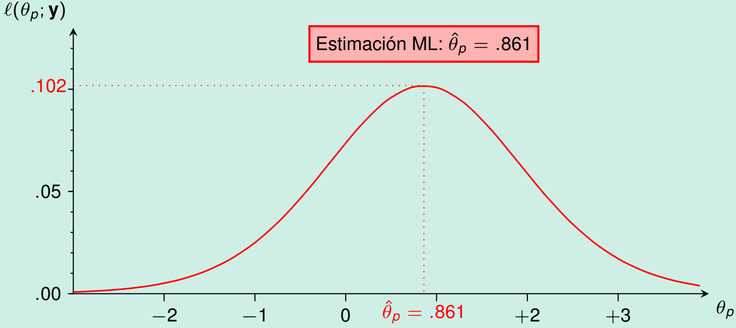
Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo

	Ítems 1	2	3	4
Persona p	0	1	1	0
	$\beta_1 = 2.0$	$\beta_2 = 0.0$	$\beta_3 = 1.0$	$\beta_4 = 0.5$

$\theta_p = ?$

$\theta_p = 0.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .074$
$\theta_p = 0.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .096$
$\theta_p = 1.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .101$
$\theta_p = 1.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .085$
$\theta_p = 2.0 \Rightarrow \ell(\theta_p; \mathbf{y}) = .059$
$\theta_p = 2.5 \Rightarrow \ell(\theta_p; \mathbf{y}) = .034$



Estimación por máxima verosimilitud del parámetro θ_p

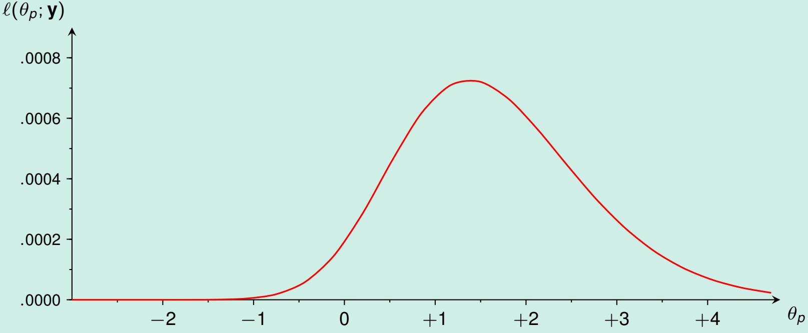
Ejemplo 2

	Ítems											
	1	2	3	4	5	6	7	8	9	10	11	12
Persona p	0	1	1	1	1	1	0	1	1	1	1	1
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}
	-0.5	-0.7	-1.0	$+3.7$	-1.5	-0.9	-0.4	-1.1	-1.4	-1.7	-0.8	-0.1

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo 2

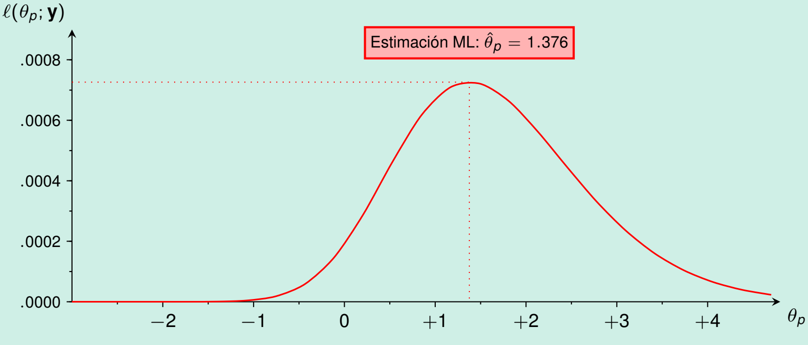
	Ítems											
	1	2	3	4	5	6	7	8	9	10	11	12
Persona p	0	1	1	1	1	1	0	1	1	1	1	1
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}
	-0.5	-0.7	-1.0	$+3.7$	-1.5	-0.9	-0.4	-1.1	-1.4	-1.7	-0.8	-0.1



Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo 2

	Ítems											
	1	2	3	4	5	6	7	8	9	10	11	12
Persona p	0	1	1	1	1	1	0	1	1	1	1	1
	β_1 -0.5	β_2 -0.7	β_3 -1.0	β_4 $+3.7$	β_5 -1.5	β_6 -0.9	β_7 -0.4	β_8 -1.1	β_9 -1.4	β_{10} -1.7	β_{11} -0.8	β_{12} -0.1



Estimación por máxima verosimilitud del parámetro θ_p

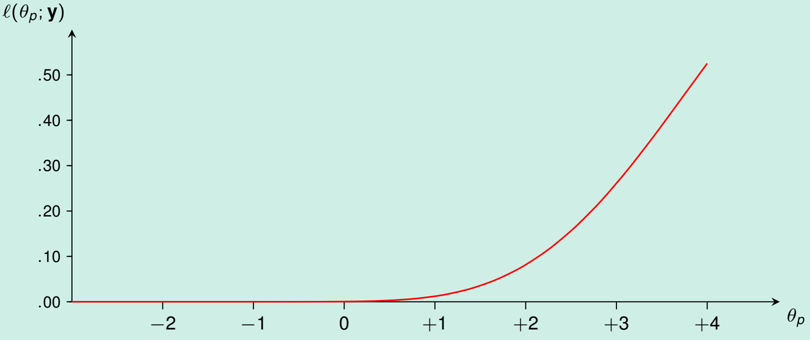
Ejemplo 3

	Ítems											
	1	2	3	4	5	6	7	8	9	10	11	12
Persona p	1	1	1	1	1	1	1	1	1	1	1	1
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}
	-0.5	-0.7	-1.0	$+3.7$	-1.5	-0.9	-0.4	-1.1	-1.4	-1.7	-0.8	-0.1

Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo 3

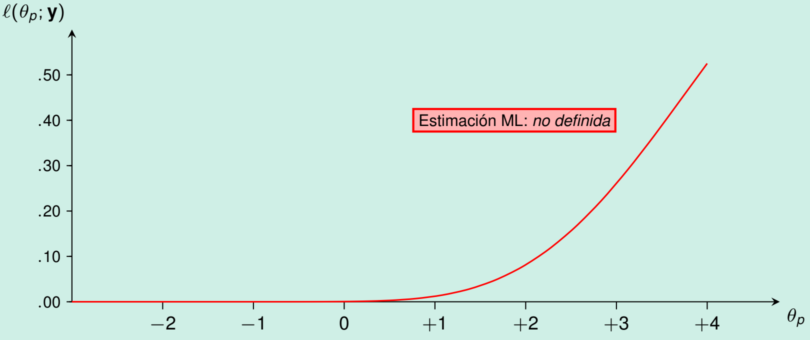
	Ítems											
	1	2	3	4	5	6	7	8	9	10	11	12
Persona p	1	1	1	1	1	1	1	1	1	1	1	1
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}
	-0.5	-0.7	-1.0	$+3.7$	-1.5	-0.9	-0.4	-1.1	-1.4	-1.7	-0.8	-0.1



Estimación por máxima verosimilitud del parámetro θ_p

Ejemplo 3

	Ítems											
	1	2	3	4	5	6	7	8	9	10	11	12
Persona p	1	1	1	1	1	1	1	1	1	1	1	1
	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}
	-0.5	-0.7	-1.0	$+3.7$	-1.5	-0.9	-0.4	-1.1	-1.4	-1.7	-0.8	-0.1



Estimación por máxima verosimilitud del parámetro θ_p

Encontrar el máximo de la función de verosimilitud $\ell(\theta_p; \mathbf{y})$: Conclusiones

- También en el caso de $\ell(\theta_p; \mathbf{y})$ se pueden utilizar las reglas de cálculo diferencial para encontrar el máximo de la función de verosimilitud $\ell(\theta_p; \mathbf{y})$.

- Sin embargo, no existe una fórmula “cerrada” que permite obtener directamente el valor de θ_p que maximice (el logaritmo de) la función de verosimilitud.

  $\longrightarrow$ Se utilizan procedimientos iterativos (por ejemplo, el método de Newton-Raphson).

- Para personas que aciertan todos los ítems o que fallan todos los ítems, no existe la estimación por máxima verosimilitud.

  $\longrightarrow$ Se utilizan variantes, por ejemplo:
  - maxima verosimilitud ponderada (WML);
  - *expected a posteriori* (EAP).

Estimación por máxima verosimilitud del parámetro θ_p

Encontrar el máximo de la función de verosimilitud $\ell(\theta_p; \mathbf{y})$: Conclusiones

- También en el caso de $\ell(\theta_p; \mathbf{y})$ se pueden utilizar las reglas de cálculo diferencial para encontrar el máximo de la función de verosimilitud $\ell(\theta_p; \mathbf{y})$.

- Sin embargo, no existe una fórmula "cerrada" que permite obtener directamente el valor de θ_p que maximice (el logaritmo de) la función de verosimilitud.

  $\longrightarrow$ Se utilizan procedimientos iterativos (por ejemplo, el método de Newton-Raphson).

- Para personas que aciertan todos los ítems o que fallan todos los ítems, no existe la estimación por máxima verosimilitud.

  $\longrightarrow$ Se utilizan variantes, por ejemplo:
  - maxima verosimilitud ponderada (WML);
  - *expected a posteriori* (EAP).

Estimación por máxima verosimilitud del parámetro θ_p

Encontrar el máximo de la función de verosimilitud $\ell(\theta_p; \mathbf{y})$: Conclusiones

- También en el caso de $\ell(\theta_p; \mathbf{y})$ se pueden utilizar las reglas de cálculo diferencial para encontrar el máximo de la función de verosimilitud $\ell(\theta_p; \mathbf{y})$.

- Sin embargo, no existe una fórmula "cerrada" que permite obtener directamente el valor de θ_p que maximice (el logaritmo de) la función de verosimilitud.

  $\longrightarrow$ Se utilizan procedimientos iterativos (por ejemplo, el método de Newton-Raphson).

- Para personas que aciertan todos los ítems o que fallan todos los ítems, no existe la estimación por máxima verosimilitud.

  $\longrightarrow$ Se utilizan variantes, por ejemplo:
  - maxima verosimilitud ponderada (WML);
  - *expected a posteriori* (EAP).

Índice

Estimar los parámetros de los ítems (β_i)

¿Cómo obtener valores para las β_i?

- En la práctica, no conocemos ni los valores de las θ_p de las distintas personas, ni los valores de las β_i de los distintos ítems.

  $\longrightarrow$ En vez de estimar solo un parámetro, hay que estimar simultáneamente múltiples parámetros.

- Se pueden utilizar los mismos principios que en los ejemplos anteriores. La única diferencia es que la función de verosimilitud ahora es una función de múltiples parámetros.

- Para el modelo de Rasch, se han propuesto 3 diferentes variantes de máxima verosimilitud:
  1. Máxima verosimilitud conjunta (JML: *Joint Maximum Likelihood*)
  2. Máxima verosimilitud marginal (MML: *Marginal Maximum Likelihood*)
  3. Máxima verosimilitud condicional (CML: *Conditional Maximum Likelihood*)

  A continuación, se introducen brevemente estas variantes y sus ventajas y desventajas.

Estimar los parámetros de los ítems (β_i)

¿Cómo obtener valores para las β_i?

- En la práctica, no conocemos ni los valores de las θ_p de las distintas personas,
  ni los valores de las β_i de los distintos ítems.

  $\longrightarrow$ En vez de estimar solo un parámetro, hay que estimar simultáneamente múltiples parámetros.

- Se pueden utilizar los mismos principios que en los ejemplos anteriores.
  La única diferencia es que la función de verosimilitud ahora es una función de múltiples parámetros.

- Para el modelo de Rasch, se han propuesto 3 diferentes variantes de máxima verosimilitud:
  1. Máxima verosimilitud conjunta (JML: *Joint Maximum Likelihood*)
  2. Máxima verosimilitud marginal (MML: *Marginal Maximum Likelihood*)
  3. Máxima verosimilitud condicional (CML: *Conditional Maximum Likelihood*)

  A continuación, se introducen brevemente estas variantes y sus ventajas y desventajas.

Estimar los parámetros de los ítems (β_i)

¿Cómo obtener valores para las β_i?

- En la práctica, no conocemos ni los valores de las θ_p de las distintas personas,
  ni los valores de las β_i de los distintos ítems.

  $\longrightarrow$ En vez de estimar solo un parámetro, hay que estimar simultáneamente múltiples parámetros.

- Se pueden utilizar los mismos principios que en los ejemplos anteriores.
  La única diferencia es que la función de verosimilitud ahora es una función de múltiples parámetros.

- Para el modelo de Rasch, se han propuesto 3 diferentes variantes de máxima verosimilitud:
  1. Máxima verosimilitud conjunta (JML: *Joint Maximum Likelihood*)
  2. Máxima verosimilitud marginal (MML: *Marginal Maximum Likelihood*)
  3. Máxima verosimilitud condicional (CML: *Conditional Maximum Likelihood*)

  A continuación, se introducen brevemente estas variantes y sus ventajas y desventajas.

JML: Estimar simultáneamente los parámetros de personas e ítems

Estimación por máxima verosimilitud conjunta (JML)

- Se considera la función de verosimilitud para todos los parámetros θ_p y β_i simultáneamente:

$$\ell(\theta_1, \theta_2, \ldots, \theta_n, \beta_1, \beta_2, \ldots, \beta_s;\ \mathbf{Y}) = \prod_{p=1}^{n} \prod_{i=1}^{s} [f_i(\theta_p)]^{y_{pi}}\ [1 - f_i(\theta_p)]^{1-y_{pi}},$$

  donde n y s denotan el número de personas e ítems, respectivamente;

  y $f_i(\theta_p) = \dfrac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$

- Sin embargo, el estimador en esta estrategia no es consistente.

  Consistencia quiere decir: si aumenta el tamaño de la muestra, entonces las estimaciones tienden a ser más precisas.

  El problema a fondo es: al aumentar el tamaño de la muestra, se añaden nuevos parámetros al modelo.

JML: Estimar simultáneamente los parámetros de personas e ítems

Estimación por máxima verosimilitud conjunta (JML)

- Se considera la función de verosimilitud para todos los parámetros θ_p y β_i simultáneamente:

$$\ell(\theta_1, \theta_2, \ldots, \theta_n, \beta_1, \beta_2, \ldots, \beta_s;\ \mathbf{Y}) \ = \prod_{p=1}^{n} \prod_{i=1}^{s} [f_i(\theta_p)]^{y_{pi}}\ [1 - f_i(\theta_p)]^{1-y_{pi}}\ ,$$

  donde n y s denotan el número de personas e ítems, respectivamente;

$$\text{y } f_i(\theta_p) = \frac{\mathrm{e}^{\theta_p - \beta_i}}{1 + \mathrm{e}^{\theta_p - \beta_i}}$$

- Sin embargo, el estimador en esta estrategia no es consistente.

  Consistencia quiere decir: si aumenta el tamaño de la muestra, entonces las estimaciones tienden a ser más precisas.

  El problema a fondo es: al aumentar el tamaño de la muestra, se añaden nuevos parámetros al modelo.

MML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud mariginal (MML)

- El método MML evita el problema del crecimiento del número de parámetros, añadiendo un supuesto al modelo sobre la distribución de los parámetros de las personas:

  En particular, para todas las personas ($p = 1, \dots, n$):

  $$\theta_p \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$$

  donde μ y σ^2 son los parámetros (la media y varianza) de la distribución normal.

- En la función de verosimilitud, se consideran únicamente los parámetros β_i de los ítems y los parámetros μ y σ^2 de la distribución de los parámetros de las personas.

  Es decir, se maximiza:

  $$\ell(\beta_1, \beta_2, \dots, \beta_s, \mu, \sigma^2; \mathbf{Y})$$

MML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud mariginal (MML)

- El método MML evita el problema del crecimiento del número de parámetros, añadiendo un supuesto al modelo sobre la distribución de los parámetros de las personas:

  En particular, para todas las personas ($p = 1, \ldots, n$):

  $$\theta_p \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$$

  donde μ y σ^2 son los parámetros (la media y varianza) de la distribución normal.

- En la función de verosimilitud, se consideran únicamente los parámetros β_i de los ítems y los parámetros μ y σ^2 de la distribución de los parámetros de las personas.

  Es decir, se maximiza:

  $$\ell(\beta_1, \beta_2, \ldots, \beta_s, \mu, \sigma^2;\ \mathbf{Y})$$

MML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud mariginal (MML)

- Entonces, el método de MML es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems (y los parámetros μ y σ^2);
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.
- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.
- La desventaja principal del método MML es que requiera:
  1. que θ efectivamente se distribuya como se supone (de forma normal).
  2. que la muestra de personas sea una muestra aleatoria de la población.

  $\Rightarrow$ Si no se cumplen estos requisitos, entonces las estimaciones de los parámetros β_i pueden tener sesgos.

MML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud mariginal (MML)

- Entonces, el método de MML es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems (y los parámetros μ y σ^2);
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.
- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.
- La desventaja principal del método MML es que requiera:
  1. que θ efectivamente se distribuya como se supone (de forma normal).
  2. que la muestra de personas sea una muestra aleatoria de la población.

  $\Rightarrow$ Si no se cumplen estos requisitos,
  entonces las estimaciones de los parámetros β_i pueden tener sesgos.

MML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud mariginal (MML)

- Entonces, el método de MML es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems (y los parámetros μ y σ^2);
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.
- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.
- La desventaja principal del método MML es que requiera:
  1. que θ efectivamente se distribuya como se supone (de forma normal).
  2. que la muestra de personas sea una muestra aleatoria de la población.

  $\Rightarrow$ Si no se cumplen estos requisitos, entonces las estimaciones de los parámetros β_i pueden tener sesgos.

MML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud mariginal (MML)

- Entonces, el método de MML es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems (y los parámetros μ y σ^2);
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.

- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.

- La desventaja principal del método MML es que requiera:
  1. que θ efectivamente se distribuya como se supone (de forma normal).
  2. que la muestra de personas sea una muestra aleatoria de la población.

  $\Rightarrow$ Si no se cumplen estos requisitos, entonces las estimaciones de los parámetros β_i pueden tener sesgos.

MML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud mariginal (MML)

- Entonces, el método de MML es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems (y los parámetros μ y σ^2);
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.
- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.
- La desventaja principal del método MML es que requiera:
  1. que θ efectivamente se distribuya como se supone (de forma normal).
  2. que la muestra de personas sea una muestra aleatoria de la población.

  $\Rightarrow$ Si no se cumplen estos requisitos, entonces las estimaciones de los parámetros β_i pueden tener sesgos.

CML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud condicional (CML)

- El método CML evita el problema del crecimiento del número de parámetros, aprovechando la propiedad en el modelo de Rasch de que el número de aciertos es un estadístico suficiente para estimar θ_p.

  Esto quiere decir que todas las personas con el mismo número de aciertos tendrán la misma estimación para θ_p.

- Como consecuencia, considerando la función de verosimilitud condicional al número de aciertos, se obtiene una expresión que es independiente de las θ_p de las personas:

  $$\ell(\beta_1, \beta_2, \ldots, \beta_s;\ \mathbf{Y} \mid \mathbf{a})$$

  (donde $\mathbf{a}$ es el vector con los posibles números de aciertos.)

CML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud condicional (CML)

- El método CML evita el problema del crecimiento del número de parámetros, aprovechando la propiedad en el modelo de Rasch de que el número de aciertos es un estadístico suficiente para estimar θ_p.

  Esto quiere decir que todas las personas con el mismo número de aciertos tendrán la misma estimación para θ_p.

- Como consecuencia, considerando la función de verosimilitud condicional al número de aciertos, se obtiene una expresión que es independiente de las θ_p de las personas:

$$\ell(\beta_1, \beta_2, \ldots, \beta_s;\ \mathbf{Y} \mid \mathbf{a})$$

  (donde **a** es el vector con los posibles números de aciertos.)

CML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud condicional (CML)

- El método de CML, igual a MML, es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems;
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.
- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.
- Ventaja del método CML sobre MML: No requiere
  1. el supuesto de la distribución de θ.
  2. que la muestra obtenida sea una muestra aleatoria de la población.
- Desventaja del método CML:

  No siempre es posible utilizarlo (por ejemplo, en el caso de diseños incompletos).

CML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud condicional (CML)

- El método de CML, igual a MML, es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems;
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.
- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.
- Ventaja del método CML sobre MML: No requiere
  1. el supuesto de la distribución de θ.
  2. que la muestra obtenida sea una muestra aleatoria de la población.
- Desventaja del método CML:

  No siempre es posible utilizarlo (por ejemplo, en el caso de diseños incompletos).

CML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud condicional (CML)

- El método de CML, igual a MML, es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems;
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.
- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.
- Ventaja del método CML sobre MML: No requiere
  1. el supuesto de la distribución de θ.
  2. que la muestra obtenida sea una muestra aleatoria de la población.
- Desventaja del método CML:

  No siempre es posible utilizarlo (por ejemplo, en el caso de diseños incompletos).

CML: Estimar los parámetros de los ítems

Estimación por máxima verosimilitud condicional (CML)

- El método de CML, igual a MML, es un procedimiento en dos pasos:
  1. Primero, se estiman los parámetros de los ítems;
  2. Segundo, se estiman los parámetros θ_p de las distintas personas, utilizando las estimaciones de los parámetros de los ítems en el primer paso.

- Se ha comprobado que de esta forma se obtienen estimaciones consistentes.

- Ventaja del método CML sobre MML: No requiere
  1. el supuesto de la distribución de θ.
  2. que la muestra obtenida sea una muestra aleatoria de la población.

- Desventaja del método CML:

  No siempre es posible utilizarlo (por ejemplo, en el caso de diseños incompletos).

Índice

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- Cada vez más, se utilizan métodos Bayesianos en la estadística.
  También en la psicometría.

- La estadística Bayesiana tiene varias ventajas sobre los métodos tradicionales, tanto teóricas como prácticas.

  Entre otros:

  1. Interpretaciones más intuitivas de los resultados obtenidos;
  2. Resultados más correctos sobre la precisión de las estimaciones de los parámetros;
  3. Es *relativamente fácil* utilizarlo con modelos más complejos

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- Cada vez más, se utilizan métodos Bayesianos en la estadística.
  También en la psicometría.

- La estadística Bayesiana tiene varias ventajas sobre los métodos tradicionales, tanto teóricas como prácticas.

  Entre otros:

  1. Interpretaciones más intuitivas de los resultados obtenidos;
  2. Resultados más correctos sobre la precisión de las estimaciones de los parámetros;
  3. Es *relativamente fácil* utilizarlo con modelos más complejos

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- Cada vez más, se utilizan métodos Bayesianos en la estadística.
  También en la psicometría.
- La estadística Bayesiana tiene varias ventajas sobre los métodos tradicionales, tanto teóricas como prácticas.

  Entre otros:

  1. Interpretaciones más intuitivas de los resultados obtenidos;
  2. Resultados más correctos sobre la precisión de las estimaciones de los parámetros;
  3. Es *relativamente fácil* utilizarlo con modelos más complejos

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- Cada vez más, se utilizan métodos Bayesianos en la estadística.
  También en la psicometría.

- La estadística Bayesiana tiene varias ventajas sobre los métodos tradicionales, tanto teóricas como prácticas.

  Entre otros:

  1. Interpretaciones más intuitivas de los resultados obtenidos;
  2. Resultados más correctos sobre la precisión de las estimaciones de los parámetros;
  3. Es *relativamente fácil* utilizarlo con modelos más complejos

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- Cada vez más, se utilizan métodos Bayesianos en la estadística.
  También en la psicometría.

- La estadística Bayesiana tiene varias ventajas sobre los métodos tradicionales, tanto teóricas como prácticas.

  Entre otros:

  1. Interpretaciones más intuitivas de los resultados obtenidos;
  2. Resultados más correctos sobre la precisión de las estimaciones de los parámetros;
  3. Es *relativamente fácil* utilizarlo con modelos más complejos

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- Cada vez más, se utilizan métodos Bayesianos en la estadística.
  También en la psicometría.

- La estadística Bayesiana tiene varias ventajas sobre los métodos tradicionales, tanto teóricas como prácticas.

  Entre otros:

  1. Interpretaciones más intuitivas de los resultados obtenidos;
  2. Resultados más correctos sobre la precisión de las estimaciones de los parámetros;
  3. Es *relativamente fácil* utilizarlo con modelos más complejos

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- La estimación Bayesiana difiere en tres puntos importantes de la estimación en el marco tradicional (que incluyue máxima verosimilitud):
  1. Se considera la distribución de los parámetros.
     ¡Los parámetros son variables!

     La distribución de los parámetros expresa nuestra incertidumbre sobre los parámetros.
  2. Se especifica una distribución a priori.
  3. Se explora la distribución a posteriori.

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- La estimación Bayesiana difiere en tres puntos importantes de la estimación en el marco tradicional (que incluyue máxima verosimilitud):
  1. Se considera la distribución de los parámetros.
     ¡Los parámetros son variables!

     La distribución de los parámetros expresa nuestra incertidumbre sobre los parámetros.
  2. Se especifica una distribución a priori.
  3. Se explora la distribución a posteriori.

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- La estimación Bayesiana difiere en tres puntos importantes de la estimación en el marco tradicional (que incluyue máxima verosimilitud):
  1. Se considera la distribución de los parámetros.
     ¡Los parámetros son variables!

     La distribución de los parámetros expresa nuestra incertidumbre sobre los parámetros.
  2. Se especifica una distribución a priori.
  3. Se explora la distribución a posteriori.

Estimación Bayesiana

Nota: Estimación Bayesiana de los parámetros en el modelo de Rasch

- La estimación Bayesiana difiere en tres puntos importantes de la estimación en el marco tradicional (que incluyue máxima verosimilitud):
  1. Se considera la distribución de los parámetros.
     ¡Los parámetros son variables!

     La distribución de los parámetros expresa nuestra incertidumbre sobre los parámetros.
  2. Se especifica una distribución a priori.
  3. Se explora la distribución a posteriori.

Índice

Índice

Precisión de las estimaciones

Precisión de las estimaciones: Introducción

- Después de haber estimado los parámetros de ítems y personas, queda la pregunta:

  ¿Qué tan precisas son estas estimaciones?

  Por ejemplo, para el parámetro de la persona p:

  ¿Qué tanto difiere la estimación $\hat{\theta}_p$ del valor verdadero θ_p?

- Si se ha utilizado el método de máxima verosimilitud, entonces se inspecciona la función de información.

Precisión de las estimaciones

Precisión de las estimaciones: Introducción

- Después de haber estimado los parámetros de ítems y personas, queda la pregunta:

  ¿Qué tan precisas son estas estimaciones?

  Por ejemplo, para el parámetro de la persona p:

  ¿Qué tanto difiere la estimación $\hat{\theta}_p$ del valor verdadero θ_p?

- Si se ha utilizado el método de máxima verosimilitud, entonces se inspecciona la función de información.

Precisión de las estimaciones

Precisión de las estimaciones: Introducción

- Después de haber estimado los parámetros de ítems y personas, queda la pregunta:

  ¿Qué tan precisas son estas estimaciones?

  Por ejemplo, para el parámetro de la persona p:

  ¿Qué tanto difiere la estimación $\hat{\theta}_p$ del valor verdadero θ_p?

- Si se ha utilizado el método de máxima verosimilitud,
  entonces se inspecciona la función de información.

La función de información en la estadística

En la estadística teórica se han examinado las propiedades de los estimadores obtenidos por máxima verosimilitud.

Uno de los resultados más importantes es el siguiente:

Distribución de estimadores por máxima verosimilitud

- Si θ es el valor verdadero de un parámetro de interés y $\hat{\theta}$ es el estimador de θ por máxima verosimilitud entonces (bajo condiciones regulares) se cumple que:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

- En la expresión anterior, $\mathcal{I}(\theta)$ es la función de información de Fisher, definida por:

$$\mathcal{I}(\theta) = -\mathscr{E}\left[\frac{\partial^2 \ln \ell(\theta \mid \mathbf{Y})}{\partial \theta^2} \,\middle|\, \theta\right]$$

- Nota: Se trata de una aproximación ($n \to \infty$).

La función de información en la estadística

En la estadística teórica se han examinado las propiedades de los estimadores obtenidos por máxima verosimilitud.

Uno de los resultados más importantes es el siguiente:

Distribución de estimadores por máxima verosimilitud

- Si θ es el valor verdadero de un parámetro de interés y $\hat{\theta}$ es el estimador de θ por máxima verosimilitud

  entonces (bajo condiciones regulares) se cumple que:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

- En la expresión anterior, $\mathcal{I}(\theta)$ es la función de información de Fisher, definida por:

$$\mathcal{I}(\theta) = -\mathscr{E}\left[\frac{\partial^2 \ln \ell(\theta \mid \mathbf{Y})}{\partial \theta^2} \,\middle|\, \theta\right]$$

- Nota: Se trata de una aproximación ($n \to \infty$).

La función de información en la estadística

En la estadística teórica se han examinado las propiedades de los estimadores obtenidos por máxima verosimilitud.

Uno de los resultados más importantes es el siguiente:

Distribución de estimadores por máxima verosimilitud

- Si θ es el valor verdadero de un parámetro de interés y $\hat{\theta}$ es el estimador de θ por máxima verosimilitud entonces (bajo condiciones regulares) se cumple que:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

- En la expresión anterior, $\mathcal{I}(\theta)$ es la función de información de Fisher, definida por:

$$\mathcal{I}(\theta) = -\mathscr{E}\left[\frac{\partial^2 \ln \ell(\theta \mid \mathbf{Y})}{\partial \theta^2} \;\middle|\; \theta\right]$$

- Nota: Se trata de una aproximación ($n \to \infty$).

La función de información en la estadística

En la estadística teórica se han examinado las propiedades de los estimadores obtenidos por máxima verosimilitud.

Uno de los resultados más importantes es el siguiente:

Distribución de estimadores por máxima verosimilitud

- Si θ es el valor verdadero de un parámetro de interés
  y $\hat{\theta}$ es el estimador de θ por máxima verosimilitud
  entonces (bajo condiciones regulares) se cumple que:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

- En la expresión anterior, $\mathcal{I}(\theta)$ es la función de información de Fisher, definida por:

$$\mathcal{I}(\theta) = -\mathscr{E}\left[\frac{\partial^2 \ln \ell(\theta \,|\, \mathbf{Y})}{\partial\theta^2} \,\middle|\, \theta\right]$$

- Nota: Se trata de una aproximación ($n \to \infty$).

La función de información en el modelo de Rasch

La función de información para los parámetros en el modelo de Rasch

- Se puede derivar la función de información para los parámetros en el modelo de Rasch:
  - Tanto $\mathcal{I}(\beta)$,
  - Como $\mathcal{I}(\theta_p)$.

- Aquí consideramos solo $\mathcal{I}(\theta_p)$.
  Es la función de información para el parámetro de una persona θ_p, si éste se estima considerando valores conocidos para las β_i.

- Se ha mostrado que:

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} f_i(\theta_p) \left[1 - f_i(\theta_p) \right]$$

donde

$$f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

La función de información en el modelo de Rasch

La función de información para los parámetros en el modelo de Rasch

- Se puede derivar la función de información para los parámetros en el modelo de Rasch:
  - Tanto $\mathcal{I}(\boldsymbol{\beta})$,
  - Como $\mathcal{I}(\theta_p)$.

- Aquí consideramos solo $\mathcal{I}(\theta_p)$.
  Es la función de información para el parámetro de una persona θ_p, si éste se estima considerando valores conocidos para las β_i.

- Se ha mostrado que:

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} f_i(\theta_p)\left[1 - f_i(\theta_p)\right]$$

donde

$$f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

La función de información en el modelo de Rasch

La función de información para los parámetros en el modelo de Rasch

- Se puede derivar la función de información para los parámetros en el modelo de Rasch:
  - Tanto $\mathcal{I}(\boldsymbol{\beta})$,
  - Como $\mathcal{I}(\theta_p)$.

- Aquí consideramos solo $\mathcal{I}(\theta_p)$.
  Es la función de información para el parámetro de una persona θ_p, si éste se estima considerando valores conocidos para las β_i.

- Se ha mostrado que:

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} f_i(\theta_p)\left[1 - f_i(\theta_p)\right]$$

donde

$$f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

La función de información en el modelo de Rasch

La función de información para los parámetros en el modelo de Rasch

- Se puede derivar la función de información para los parámetros en el modelo de Rasch:
  - Tanto $\mathcal{I}(\boldsymbol{\beta})$,
  - Como $\mathcal{I}(\theta_p)$.

- Aquí consideramos solo $\mathcal{I}(\theta_p)$.
  Es la función de información para el parámetro de una persona θ_p, si éste se estima considerando valores conocidos para las β_i.

- Se ha mostrado que:

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} f_i(\theta_p) \left[1 - f_i(\theta_p) \right]$$

donde

$$f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

La función de información en el modelo de Rasch

La función de información para los parámetros en el modelo de Rasch

- Se puede derivar la función de información para los parámetros en el modelo de Rasch:
  - Tanto $\mathcal{I}(\boldsymbol{\beta})$,
  - Como $\mathcal{I}(\theta_p)$.

- Aquí consideramos solo $\mathcal{I}(\theta_p)$.
  Es la función de información para el parámetro de una persona θ_p, si éste se estima considerando valores conocidos para las β_i.

- Se ha mostrado que:

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} f_i(\theta_p) \left[1 - f_i(\theta_p)\right]$$

donde

$$f_i(\theta_p) = \frac{e^{\theta_p - \beta_i}}{1 + e^{\theta_p - \beta_i}}$$

La función de información en el modelo de Rasch

La función de información para el parámetro θ_p

- Entonces, tenemos

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} f_i(\theta_p)\left[1 - f_i(\theta_p)\right]$$

- $\mathcal{I}_i(\theta_p)$ se llama la función de información del ítem i e indica cuánta información aporta el ítem i para estimar θ_p.
- $\mathcal{I}(\theta_p)$ [o $\mathcal{I}_{\text{test}}(\theta_p)$] se llama la función de información del test.

La función de información en el modelo de Rasch

La función de información para el parámetro θ_p

- Entonces, tenemos

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} \underbrace{f_i(\theta_p)\left[1 - f_i(\theta_p)\right]}_{\equiv \mathcal{I}_i(\theta_p)}$$

- $\mathcal{I}_i(\theta_p)$ se llama la función de información del ítem i
  e indica cuánta información aporta el ítem i para estimar θ_p.
- $\mathcal{I}(\theta_p)$ [o $\mathcal{I}_{\text{test}}(\theta_p)$] se llama la función de información del test.

La función de información en el modelo de Rasch

La función de información para el parámetro θ_p

- Entonces, tenemos

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} \underbrace{f_i(\theta_p)\left[1 - f_i(\theta_p)\right]}_{\equiv \mathcal{I}_i(\theta_p)}$$

  - $\mathcal{I}_i(\theta_p)$ se llama la función de información del ítem i e indica cuánta información aporta el ítem i para estimar θ_p.
  - $\mathcal{I}(\theta_p)$ [o $\mathcal{I}_{\text{test}}(\theta_p)$] se llama la función de información del test.

La función de información en el modelo de Rasch

La función de información para el parámetro θ_p

- Entonces, tenemos

$$\mathcal{I}(\theta_p) = \sum_{i=1}^{n} \underbrace{f_i(\theta_p)\left[1 - f_i(\theta_p)\right]}_{\equiv \mathcal{I}_i(\theta_p)}$$

  - $\mathcal{I}_i(\theta_p)$ se llama la función de información del ítem *i*
    e indica cuánta información aporta el ítem *i* para estimar θ_p.
  - $\mathcal{I}(\theta_p)$ [o $\mathcal{I}_{\text{test}}(\theta_p)$] se llama la función de información del test.

La función de información en el modelo de Rasch

Gráfica de la función de información de los ítems en el modelo de Rasch

Entonces, en el modelo Rasch, la función de información del ítem i se da por:

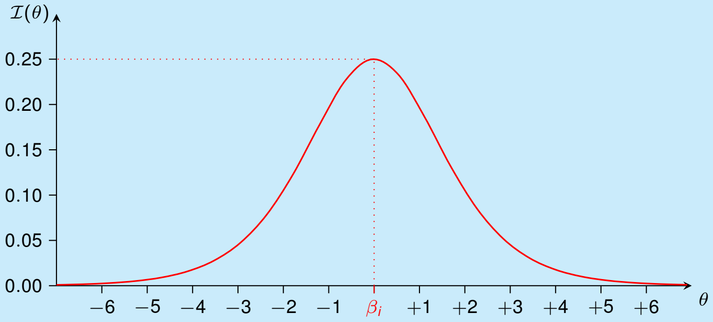
$$\mathcal{I}_i(\theta) \;=\; f_i(\theta) \;\times\; \left[1 \;-\; f_i(\theta)\right]$$

La función de información en el modelo de Rasch

Gráfica de la función de información de los ítems en el modelo de Rasch

Entonces, en el modelo Rasch, la función de información del ítem *i* se da por:

$$\mathcal{I}_i(\theta) \;=\; f_i(\theta) \;\times\; \left[1 \;-\; f_i(\theta)\right]$$

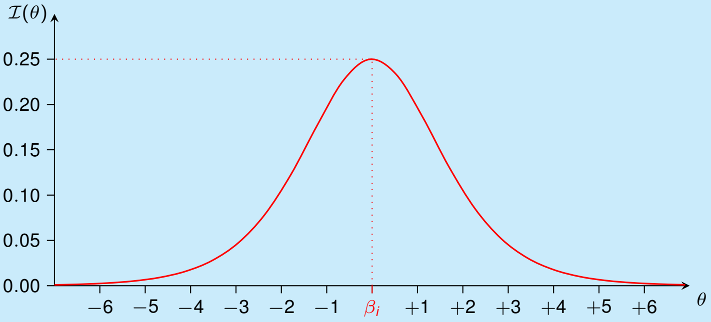


La función de información en el modelo de Rasch

Gráfica de la función de información de los ítems en el modelo de Rasch

Entonces, en el modelo Rasch, la función de información del ítem *i* se da por:

$$\mathcal{I}_i(\theta) \;=\; f_i(\theta) \;\times\; \left[1 \;-\; f_i(\theta)\right]$$

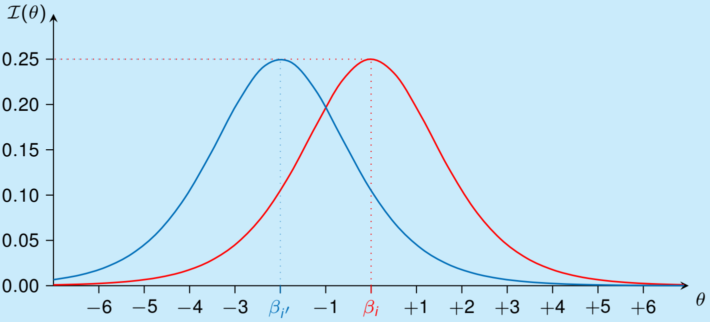


La función de información en el modelo de Rasch

Gráfica de la función de información de los ítems en el modelo de Rasch

Entonces, en el modelo Rasch, la función de información del ítem *i* se da por:

$$\mathcal{I}_i(\theta) \;=\; f_i(\theta) \;\times\; \left[1 \;-\; f_i(\theta)\right]$$

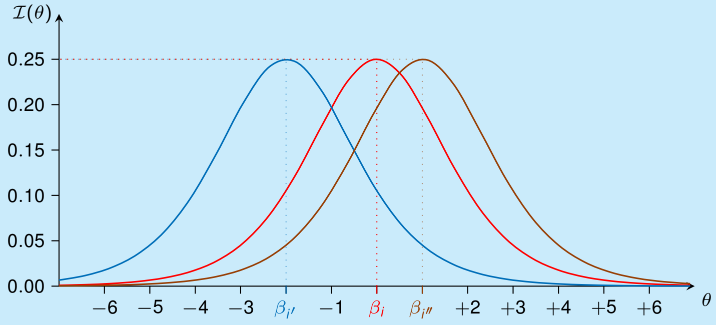


La función de información en el modelo de Rasch

Gráfica de la función de información de los ítems en el modelo de Rasch

Entonces, en el modelo Rasch, la función de información del ítem *i* se da por:

$$\mathcal{I}_i(\theta) \;=\; f_i(\theta) \;\times\; \left[1 \;-\; f_i(\theta)\right]$$



La función de información en el modelo de Rasch

Gráfica de la función de información del test en el modelo de Rasch

También la función de información del test se puede gráficar:

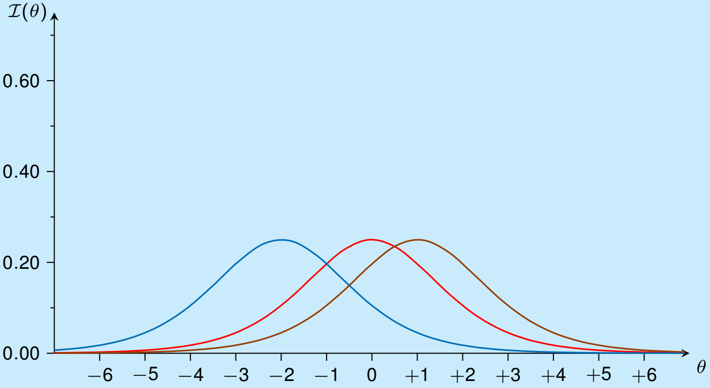
$$\mathcal{I}_{\text{test}}(\theta) \;=\; \sum_i \mathcal{I}_i(\theta)$$

La función de información en el modelo de Rasch

Gráfica de la función de información del test en el modelo de Rasch

También la función de información del test se puede gráficar:

$$\mathcal{I}_{\text{test}}(\theta) \; = \; \sum_i \mathcal{I}_i(\theta)$$

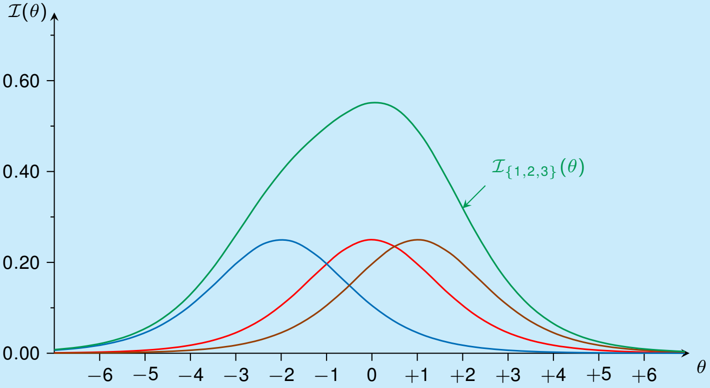


La función de información en el modelo de Rasch

Gráfica de la función de información del test en el modelo de Rasch

También la función de información del test se puede gráficar:

$$\mathcal{I}_{\text{test}}(\theta) = \sum_i \mathcal{I}_i(\theta)$$

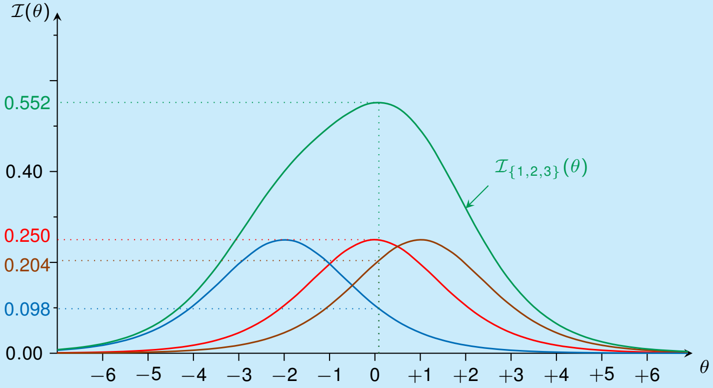


La función de información en el modelo de Rasch

Gráfica de la función de información del test en el modelo de Rasch

También la función de información del test se puede gráficar:

$$\mathcal{I}_{\text{test}}(\theta) \;=\; \sum_i \mathcal{I}_i(\theta)$$



La función de información en el modelo de Rasch

Algunas conclusiones

- El error estándar de medición asociado con una estimación $\hat{\theta}$ se da por:

$$\mathrm{SE}(\hat{\theta}) \;\approx\; \frac{1}{\sqrt{\mathcal{I}(\theta)}}$$

- La información de un test (o el error estándar de medición) no es la misma para todas las personas. Generalmente, la medición es más precisa para personas cuyo nivel de habilidad se encuentra por el promedio de los grados de dificultad de los ítems.

  Vice versa, si es importante medir con más precisión cierto (rango de) nivel de habilidad, se recomienda incluir ítems en el test cuyo grado de dificultad se encuentra cercano a dicho nivel.

- Al añadir ítems a un test, la información del test incrementa y el error estándar de medición se reduce.

La función de información en el modelo de Rasch

Algunas conclusiones

- El error estándar de medición asociado con una estimación $\hat{\theta}$ se da por:

$$\mathrm{SE}(\hat{\theta}) \approx \frac{1}{\sqrt{\mathcal{I}(\theta)}}$$

- La información de un test (o el error estándar de medición) no es la misma para todas las personas. Generalmente, la medición es más precisa para personas cuyo nivel de habilidad se encuentra por el promedio de los grados de dificultad de los ítems.

  Vice versa, si es importante medir con más precisión cierto (rango de) nivel de habilidad, se recomienda incluir ítems en el test cuyo grado de dificultad se encuentra cercano a dicho nivel.

- Al añadir ítems a un test, la información del test incrementa y el error estándar de medición se reduce.

La función de información en el modelo de Rasch

Algunas conclusiones

- El error estándar de medición asociado con una estimación $\hat{\theta}$ se da por:

$$\mathrm{SE}(\hat{\theta}) \approx \frac{1}{\sqrt{\mathcal{I}(\theta)}}$$

- La información de un test (o el error estándar de medición) no es la misma para todas las personas. Generalmente, la medición es más precisa para personas cuyo nivel de habilidad se encuentra por el promedio de los grados de dificultad de los ítems.

  Vice versa, si es importante medir con más precisión cierto (rango de) nivel de habilidad, se recomienda incluir ítems en el test cuyo grado de dificultad se encuentra cercano a dicho nivel.

- Al añadir ítems a un test, la información del test incrementa y el error estándar de medición se reduce.

Índice

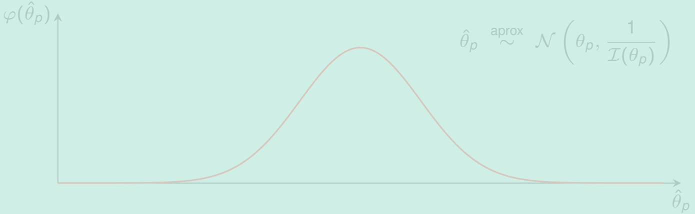
Un intervalo de confianza para θ_p

Ejemplo

- Como hemos mencionado anteriormente:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

Gráficamente:



Un intervalo de confianza para θ_p

Ejemplo

- Como hemos mencionado anteriormente:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

Gráficamente:

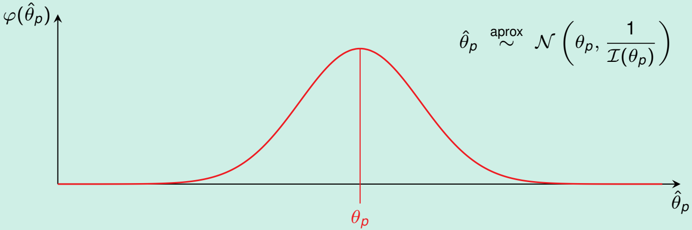
Un intervalo de confianza para θ_p

Ejemplo

- Como hemos mencionado anteriormente:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

Gráficamente:



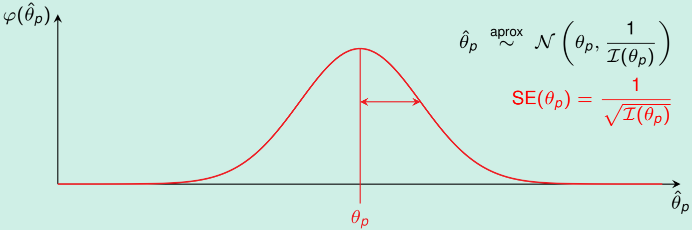
Un intervalo de confianza para θ_p

Ejemplo

- Como hemos mencionado anteriormente:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

Gráficamente:



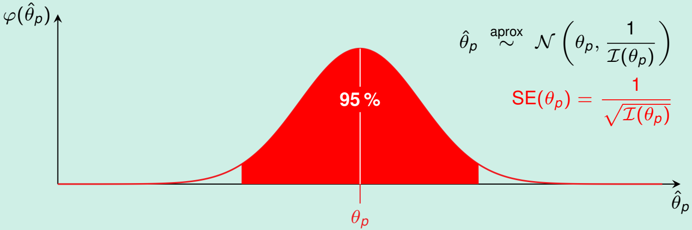
Un intervalo de confianza para θ_p

Ejemplo

- Como hemos mencionado anteriormente:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

Gráficamente:



- Entonces:

$$\Pr\left(\theta_p - 1.96\,\mathrm{SE}(\theta_p) \leqslant \hat{\theta}_p \leqslant \theta_p - 1.96\,\mathrm{SE}(\theta_p)\right) = 0.95$$

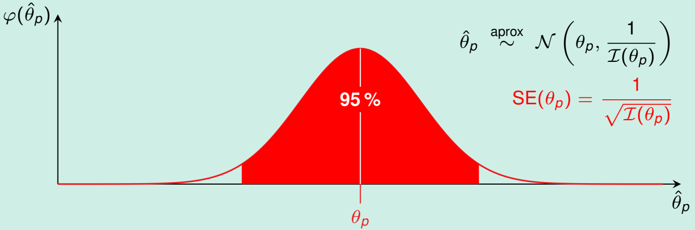
Un intervalo de confianza para θ_p

Ejemplo

- Como hemos mencionado anteriormente:

$$\hat{\theta} \overset{\text{aprox}}{\sim} \mathcal{N}\left(\theta, \frac{1}{\mathcal{I}(\theta)}\right).$$

Gráficamente:



- Entonces:

$$\Pr\left(\theta_p - 1.96\,\text{SE}(\theta_p) \leqslant \hat{\theta}_p \leqslant \theta_p - 1.96\,\text{SE}(\theta_p)\right) = 0.95$$

$$\iff \Pr\left(\hat{\theta}_p - 1.96\,\text{SE}(\theta_p) \leqslant \theta_p \leqslant \hat{\theta}_p - 1.96\,\text{SE}(\theta_p)\right) = 0.95$$

Un intervalo de confianza para θ_p

Derivar un intervalo de confianza para θ_p

- A partir de

$$\Pr\left(\hat{\theta}_p - 1.96\,\mathrm{SE}(\theta_p) \leqslant \theta_p \leqslant \hat{\theta}_p - 1.96\,\mathrm{SE}(\theta_p)\right) = 0.95$$

- y una estimación $\hat{\theta}_p$
- y sustituyendo $\hat{\theta}_p$ en la expresión $\mathrm{SE}(\theta_p)$

se construye un intervalo de confianza de 95 %:

$$\left[\hat{\theta}_p - 1.96\,\mathrm{SE}(\hat{\theta}_p), \hat{\theta}_p + 1.96\,\mathrm{SE}(\hat{\theta}_p)\right].$$

Un intervalo de confianza para θ_p

Derivar un intervalo de confianza para θ_p

- A partir de

$$\Pr\left(\hat{\theta}_p - 1.96\,\mathrm{SE}(\theta_p) \leqslant \theta_p \leqslant \hat{\theta}_p - 1.96\,\mathrm{SE}(\theta_p)\right) = 0.95$$

- y una estimación $\hat{\theta}_p$
- y sustituyendo $\hat{\theta}_p$ en la expresión $\mathrm{SE}(\theta_p)$

se construye un intervalo de confianza de 95 %:

$$\left[\hat{\theta}_p - 1.96\,\mathrm{SE}(\hat{\theta}_p), \hat{\theta}_p + 1.96\,\mathrm{SE}(\hat{\theta}_p)\right].$$

Un intervalo de confianza para θ_p

Derivar un intervalo de confianza para θ_p

- A partir de

$$\Pr\left(\hat{\theta}_p - 1.96\,\mathrm{SE}(\theta_p) \leqslant \theta_p \leqslant \hat{\theta}_p - 1.96\,\mathrm{SE}(\theta_p)\right) = 0.95$$

- y una estimación $\hat{\theta}_p$
- y sustituyendo $\hat{\theta}_p$ en la expresión $\mathrm{SE}(\theta_p)$

se construye un intervalo de confianza de 95 %:

$$\left[\hat{\theta}_p - 1.96\,\mathrm{SE}(\hat{\theta}_p), \hat{\theta}_p + 1.96\,\mathrm{SE}(\hat{\theta}_p)\right].$$

Un intervalo de confianza para θ_p

Derivar un intervalo de confianza para θ_p

- A partir de

$$\Pr\left(\hat{\theta}_p - 1.96\,\mathrm{SE}(\theta_p) \;\leqslant\; \theta_p \;\leqslant\; \hat{\theta}_p - 1.96\,\mathrm{SE}(\theta_p)\right) = 0.95$$

- y una estimación $\hat{\theta}_p$
- y sustituyendo $\hat{\theta}_p$ en la expresión $\mathrm{SE}(\theta_p)$

se construye un intervalo de confianza de 95 %:

$$\left[\hat{\theta}_p - 1.96\,\mathrm{SE}(\hat{\theta}_p), \hat{\theta}_p + 1.96\,\mathrm{SE}(\hat{\theta}_p)\right].$$

Un intervalo de confianza para θ_p

Ejemplo

- Supongamos que:

$$\hat{\theta}_p = 0.087$$

- y

$$\text{SE}(\hat{\theta}_p) = \frac{1}{\sqrt{\mathcal{I}(\hat{\theta}_p)}} = 0.552.$$

- Entonces, un intervalo de confianza de 95 % se obtiene por:

$$[0.552 - 1.96 \times 0.087, 0.552 + 1.96 \times 0.087]$$
$$[0.381, 0.723]$$

Un intervalo de confianza para θ_p

Ejemplo

- Supongamos que:

$$\hat{\theta}_p = 0.087$$

- y

$$\mathrm{SE}(\hat{\theta}_p) = \frac{1}{\sqrt{\mathcal{I}(\hat{\theta}_p)}} = 0.552.$$

- Entonces, un intervalo de confianza de 95 % se obtiene por:

$$[0.552 - 1.96 \times 0.087, 0.552 + 1.96 \times 0.087]$$

$$[0.381, 0.723]$$

Un intervalo de confianza para θ_p

Ejemplo

- Supongamos que:

$$\hat{\theta}_p = 0.087$$

- y

$$\mathrm{SE}(\hat{\theta}_p) = \frac{1}{\sqrt{\mathcal{I}(\hat{\theta}_p)}} = 0.552.$$

- Entonces, un intervalo de confianza de 95 % se obtiene por:

$$\left[0.552 - 1.96 \times 0.087, 0.552 + 1.96 \times 0.087\right]$$

$$\left[0.381, 0.723\right]$$

Un intervalo de confianza para θ_p

Ejemplo

- Supongamos que:

$$\hat{\theta}_p = 0.087$$

- y

$$\text{SE}(\hat{\theta}_p) = \frac{1}{\sqrt{\mathcal{I}(\hat{\theta}_p)}} = 0.552.$$

- Entonces, un intervalo de confianza de 95 % se obtiene por:

$$\left[0.552 - 1.96 \times 0.087, 0.552 + 1.96 \times 0.087\right]$$
$$\left[0.381, 0.723\right]$$

Índice

La bondad de ajuste del modelo

Después del proceso de estimar los parámetros del modelo, se pasa a la fase de evaluar la plausibilidad de los supuestos del modelo.

¿Es razonable creer que el modelo es “verdad” para los datos recopilados?

¿El modelo describe adecuadamente el proceso del cual se originaron los datos recopilados?

¡Ojo! Ningún modelo es perfecto.

Con muchos datos se rechazará cualquier modelo.

La bondad de ajuste del modelo

Después del proceso de estimar los parámetros del modelo, se pasa a la fase de evaluar la plausibilidad de los supuestos del modelo.

¿Es razonable creer que el modelo es “verdad” para los datos recopilados?

¿El modelo describe adecuadamente el proceso del cual se originaron los datos recopilados?

¡Ojo! Ningún modelo es perfecto.

Con muchos datos se rechazará cualquier modelo.

La bondad de ajuste del modelo

Después del proceso de estimar los parámetros del modelo, se pasa a la fase de evaluar la plausibilidad de los supuestos del modelo.

¿Es razonable creer que el modelo es “verdad” para los datos recopilados?

¿El modelo describe adecuadamente el proceso del cual se originaron los datos recopilados?

¡Ojo! Ningún modelo es perfecto.

Con muchos datos se rechazará cualquier modelo.

La bondad de ajuste del modelo

Ideas generales

- Se examinan aquellos aspectos de los datos que no se utilizaron para la estimación de los parámetros.

Ejemplo

- Para la estimación de los parámetros de las personas en el modelo Rasch solo importa el número de aciertos.
- Por otro lado, no todos los patrones de respuestas son igualmente probables bajo los supuestos del modelo.

$\Longrightarrow$ Investigar si no hay demasiadas excepciones.

La bondad de ajuste del modelo

Ideas generales

- Se examinan aquellos aspectos de los datos que no se utilizaron para la estimación de los parámetros.

Ejemplo

- Para la estimación de los parámetros de las personas en el modelo Rasch solo importa el número de aciertos.
- Por otro lado, no todos los patrones de respuestas son igualmente probables bajo los supuestos del modelo.

$\Longrightarrow$ Investigar si no hay demasiadas excepciones.

La bondad de ajuste del modelo

Ideas generales

- Se examinan aquellos aspectos de los datos que no se utilizaron para la estimación de los parámetros.

Ejemplo

- Para la estimación de los parámetros de las personas en el modelo Rasch solo importa el número de aciertos.
- Por otro lado, no todos los patrones de respuestas son igualmente probables bajo los supuestos del modelo.

$\Longrightarrow$ Investigar si no hay demasiadas excepciones.

La bondad de ajuste del modelo

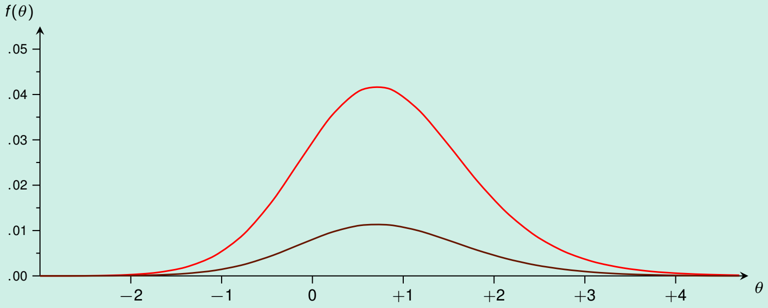
Ejemplo

	Preguntas					
	1	2	3	4	5	6
Persona 1	1	1	1	1	0	0
Persona 2	0	0	1	1	1	1
	β_1 -0.4	β_2 -0.2	β_3 -0.1	β_4 0.0	β_5 $+0.2$	β_6 $+0.5$

La bondad de ajuste del modelo

Ejemplo

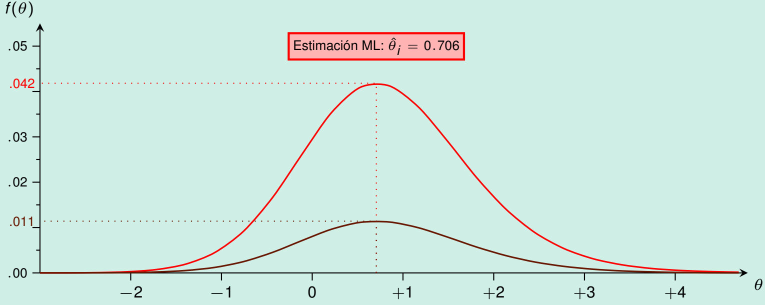
	Preguntas					
	1	2	3	4	5	6
Persona 1	1	1	1	1	0	0
Persona 2	0	0	1	1	1	1
	β_1 -0.4	β_2 -0.2	β_3 -0.1	β_4 0.0	β_5 $+0.2$	β_6 $+0.5$



La bondad de ajuste del modelo

Ejemplo

	Preguntas					
	1	2	3	4	5	6
Persona 1	1	1	1	1	0	0
Persona 2	0	0	1	1	1	1
	β_1	β_2	β_3	β_4	β_5	β_6
	-0.4	-0.2	-0.1	0.0	$+0.2$	$+0.5$



La bondad de ajuste del modelo

Ideas generales (continuación):

- Comúnmente, la evaluación de la bondad de ajuste es a través de contrastes de hipótesis (pruebas de significancia)

  Importante: En estos contrastes, el modelo es la hipótesis nula.
  Esperamos que **no** se rechace la hipótesis nula.
  ¡Esperamos que $p > .05$!

La bondad de ajuste del modelo

Ideas generales (continuación):

- Comúnmente, la evaluación de la bondad de ajuste es a través de contrastes de hipótesis (pruebas de significancia)

  Importante: En estos contrastes, el modelo es la hipótesis nula.
  Esperamos que **no** se rechace la hipótesis nula.
  ¡Esperamos que $p > .05$!

La bondad de ajuste del modelo

Tipos de pruebas de bondad de ajuste

Existen varios tipos de pruebas:

- Para evaluar el ajuste global
  ¿El modelo en su globalidad se ajusta?
  Si se rechaza el modelo (es decir, si la prueba resulta significativa), no sabemos por qué.
- Para evaluar el ajuste de cada ítem
  Por ejemplo, ítems que resuelvan las personas con niveles más bajas y no las personas con niveles altos tendrán mal ajuste.
- Para evaluar el ajuste de cada persona
  Por ejemplo, personas que resuelvan los ítems más difíciles y no los ítems fáciles tendrán mal ajuste
- Para evaluar el ajuste a supuestos específicos
  - ¿Es plausible es supuesto de unidimensionalidad?
  - ¿Independencia local?
  - . . .

La bondad de ajuste del modelo

Tipos de pruebas de bondad de ajuste

Existen varios tipos de pruebas:

- Para evaluar el ajuste global
  ¿El modelo en su globalidad se ajusta?
  Si se rechaza el modelo (es decir, si la prueba resulta significativa), no sabemos por qué.
- Para evaluar el ajuste de cada ítem
  Por ejemplo, ítems que resuelvan las personas con niveles más bajas y no las personas con niveles altos tendrán mal ajuste.
- Para evaluar el ajuste de cada persona
  Por ejemplo, personas que resuelvan los ítems más difíciles y no los ítems fáciles tendrán mal ajuste
- Para evaluar el ajuste a supuestos específicos
  - ¿Es plausible es supuesto de unidimensionalidad?
  - ¿Independencia local?
  - ...

La bondad de ajuste del modelo

Tipos de pruebas de bondad de ajuste

Existen varios tipos de pruebas:

- Para evaluar el ajuste global
  ¿El modelo en su globalidad se ajusta?
  Si se rechaza el modelo (es decir, si la prueba resulta significativa), no sabemos por qué.
- Para evaluar el ajuste de cada ítem
  Por ejemplo, ítems que resuelvan las personas con niveles más bajas y no las personas con niveles altos tendrán mal ajuste.
- Para evaluar el ajuste de cada persona
  Por ejemplo, personas que resuelvan los ítems más difíciles y no los ítems fáciles tendrán mal ajuste
- Para evaluar el ajuste a supuestos específicos
  - ¿Es plausible es supuesto de unidimensionalidad?
  - ¿Independencia local?
  - ...

La bondad de ajuste del modelo

Tipos de pruebas de bondad de ajuste

Existen varios tipos de pruebas:

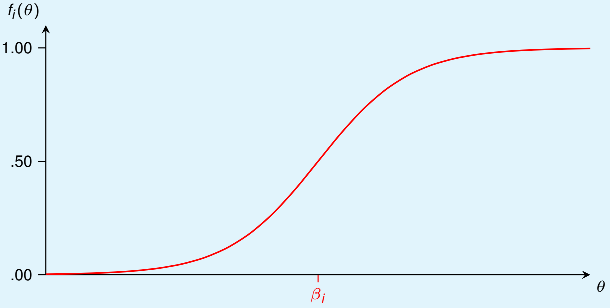
- Para evaluar el ajuste global
  ¿El modelo en su globalidad se ajusta?
  Si se rechaza el modelo (es decir, si la prueba resulta significativa), no sabemos por qué.
- Para evaluar el ajuste de cada ítem
  Por ejemplo, ítems que resuelvan las personas con niveles más bajas y no las personas con niveles altos tendrán mal ajuste.
- Para evaluar el ajuste de cada persona
  Por ejemplo, personas que resuelvan los ítems más difíciles y no los ítems fáciles tendrán mal ajuste
- Para evaluar el ajuste a supuestos específicos
  - ¿Es plausible es supuesto de unidimensionalidad?
  - ¿Independencia local?
  - ...

Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica

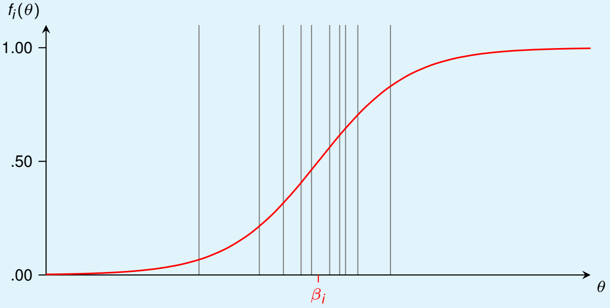
Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica



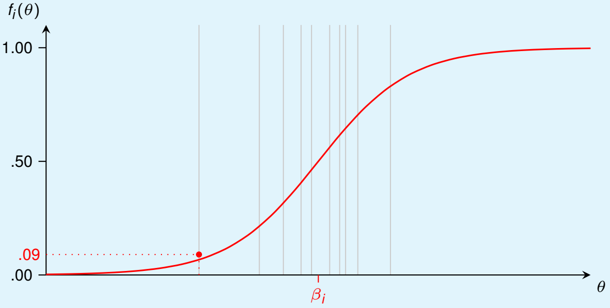
Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica



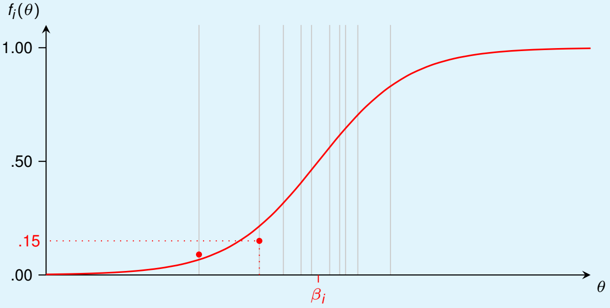
Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica



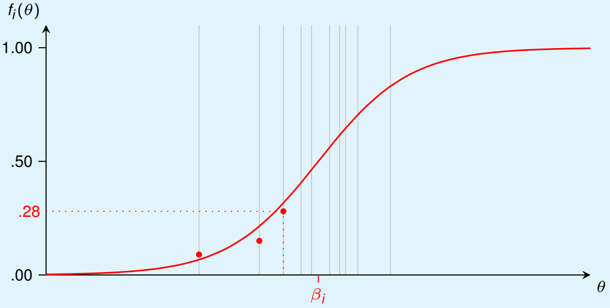
Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica



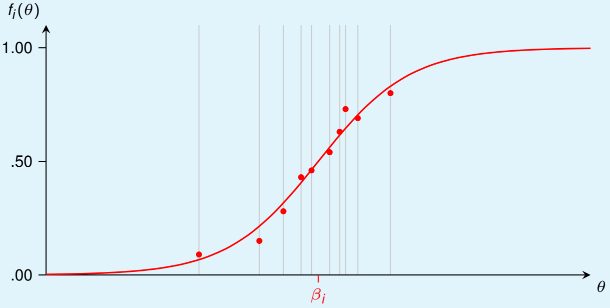
Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica



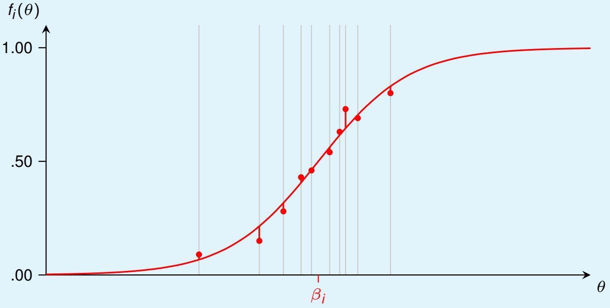
Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica



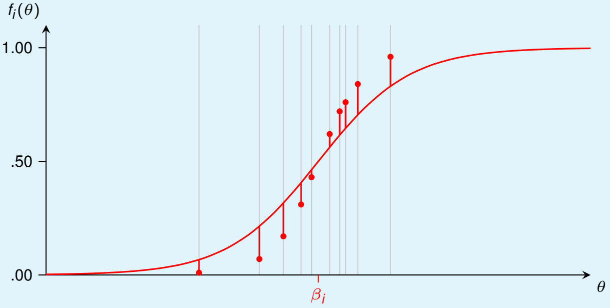
Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica



Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica



Evaluar la bondad de ajuste de ítems: Método gráfico

1. Graficar la curva característica del ítem
2. Dividir la muestra de personas en grupos según su valor en $\hat{\theta}$
   Calcular la media $\bar{\hat{\theta}}$ en cada grupo
3. Calcular y representar la proporción de aciertos en cada grupo
4. Apreciar la diferencia entre proporción observada y probabilidad teórica