El objetivo de esta tarea es comparar las estimaciones de los parámetros de los ítems en el modelo de Rasch que se obtienen a través de los métodos de máxima verosimilitud condicional (CML) y máxima verosimilitud marginal (MML) con los valores verdaderos de estos parámetros. Para esto se realizará un pequeño estudio de simulación que se describe en los siguientes pasos.

1. En al paso 1, se simularán datos bajo el modelo de Rasch para las respuestas de 1,000 personas en 7 ítems dicotómicos. "Simular" quiere decir que se construyan datos bajo los supuestos del modelo (de tal forma se sabe que los datos siguen totalmente el modelo de Rasch). Para este primer paso hay que llevar a cabo las siguientes tres tareas:

   (a) Construye un vector con valores (verdaderos) para los parámetros de dificultad (β) de los 7 ítems. En particular, construye el siguiente vector:

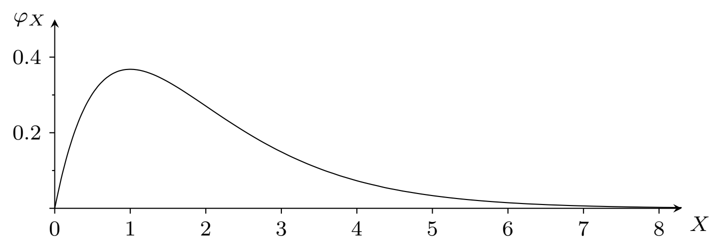
   $$\beta \;=\; (-3, -2, -1, 0, +1, +2, +3)$$

   (b) Construye un vector con valores (verdaderos) para los parámetros (θ) de las 1,000 personas. Se considerarán dos distribuciones teóricas distintas para extraer estos valores.
   Primero:

   - Una distribución normal con parámetros $\mu = 0$ y $\sigma^2 = 1$. Es decir, para todas las personas ($p = 1, ..., 1000$) se tiene
     $$\theta_p \;\sim\; \mathrm{N}(0, 1).$$
     Nótese que el valor esperado de los parámetros de las personas es 0.
     Para extraer de forma aleatoria valores de una distribución normal, puede utilizarse la función `rnorm` en R. En particular, la instrucción `rnorm(n = 1000, mean = 0, sd = 1)` extraerá 1,000 valores de la distribución normal anterior.

   Segundo:

   - Una distribución gamma con el "parámetro de forma" igual a 2. La siguiente figura gráfica esta función de densidad:

     El valor esperado de esta distribución gamma es igual a 2.
     Para extraer de forma aleatoria valores de una distribución gamma, puede utilizarse la función `rgamma` en R. En particular, la instrucción `rgamma(n = 1000, shape = 2)` extraerá 1,000 valores de la distribución gamma anterior.

   (c) Construye una matriz de datos para las 1,000 personas y los 7 ítems utilizando los valores asignados a los parámetros en los puntos anteriores. Para esto, se calcula la probabilidad de acertar—utilizando la ecuación básica del modelo de Rasch—para cada una de las 1,000 personas en cada uno de los 7 ítems y se asigna un valor de 1 en la celda correspondiente con dicha probabilidad (y un valor de 0 con la probabilidad complementaria).
   Para este procedimiento, se puede utilizar la función `sim.rasch` del paquete `eRm` en R. Esta función simplemente requiere como argumentos el vector de θs de las personas (en `persons =`) y el vector de βs de los ítems (en `items =`) para generar la matriz de datos.

2. En el Paso 2, se obtienen, para la matriz de datos simulados en el paso anterior, las estimaciones de los parámetros de los ítems tanto por CML como por MML. Se pueden utilizar las siguientes funciones en R:

   - `RM` del paquete `eRm`. Esta función ajusta el modelo de Rasch utilizando CML y requiere como único argumento (`X =`) la matriz de datos.
   - `tam.mml` del paquete `TAM`. Esta función ajuste el modelo de Rasch utilizando MML y requiere como argumento (`resp =`) la matriz de datos.

   Ambas funciones generan un objeto con contiene información del ajuste del modelo, incluyendo las estimaciones para los parámetros de los ítems.

3. En el Paso 3, se extraen las betas de los objetos con información de ajuste generados en el paso anterior:
   - Si el objeto generado por la función `RM` se llama `fit.cml`, entonces se saca el vector de las βs estimadas con la instrucción `fit.cml[["betapar"]]`.
   - Si el objeto generado por la función `tam.mml` se llama `fit.mml`, entonces se saca el vector de βs estimadas con la instrucción `fit.mml[["xsi"]][["xsi"]]`.

   Como cuestión técnica, considera para ambos casos si es necesario cambiar o no el signo de los valores en estos vectores para que representen el parámetro de dificultad β.

4. En el Paso 4, se toma en cuenta la indeterminación en el modelo para hacer comparables las estimaciones de los parámetros (β en este caso) obtenidas por ambos métodos de estimación y hacerlas comparables también con los valores verdaderos de los parámetros (véase el Paso 1(a)). Asegúrate que, tanto las estimaciones por el método CML como las del MML tengan el promedio igual a 0 (es decir, si es necesario, suma o resta una constante a todas las βs para que su media sea igual a 0).
   Nótese que por la forma en la que se construyó el vector con los valores verdaderos para β en el Paso 1(a), su media es igual a 0.

5. En el ultimo paso, se comparan las estimaciones de los parámetros (bajo ambos métodos de estimación) con sus valores verdaderos. Para esto:
   (a) Calcula para cada párametro, la diferencia *absoluta* (con la función `abs`) entre el valor verdadero y el valor estimado, tanto para las estimaciones CML como las MML.
   (b) Interpreta el resultado al que has llegado:
      - Para estos datos simulados, ¿son superiores las estimaciones por CML o las por MML?
      - ¿Tu conclusión depende de la distribución teórica de la cual se extrajeron los valores verdaderos de θ (en el Paso 1(b))?

6. Vuelve a ejecutar el código completo (Pasos 1 a 5); en cada ejecución, como el vector de θs y los datos simulados cambian (ya que se basan en procedimientos para extraer números aleatorios), el resultado final puede ser diferente. Intenta llegar a una conclusión general, por ejemplo, con base en 10 ejecuciones del código.