**Estrategia de análisis - propuesta**

**1. Generación de banco de ítems verdaderos**

Generación de un banco único con diferentes tipos de itemes. Esto se genera una vez y se guarda. Los tipos de itemes refieren a los siguientes modelos que se listan a continuación pero pueden surgir otros. Hacerlo lo suficientemente grande

En el proyecto se trabajará con los siguientes tipos de ítems en la parte unidimensional y multidimensional.

| **Dicotómicos** | **Prioridad** |
| --- | --- |
| Rasch | (1) |
| 1PL | (1) |
| 2PL | (1) |
| 3PL | (1) |
| 4PL | (2) |
| OPLM One-Parameter Logistic Model with Imputed Slopes | (2) |
| Logit cúbico | (1) |
| MHM-Modelo homogéneo monótono | A revisar |
| DMH- Modelo de Doble monotonicidad | A revisar |

| **Politómicos** | **Prioridad** |
| --- | --- |
| Rating Scale Model RSM | (1) |
| Partial Credit Model PCM | (1) |
| Generalized Partial Credit Model G-PCM | (1) |
| OPLM-po One-Parameter Logistic Model (polytomous items) with Imputed Slopes | (2) |
| Rating Scale version of Graded Response Model | (1) |
| Graded Response Mode | (1) |
| Monotone Homogeneity Model (polytomous items | A revisar |
| Weak Double Monotonicity Mode | A revisar |
| Strong Double Monotonicity Mode | A revisar |
| Isotonic Ordinal Probabilistic Mode | A revisar |

Se sugiere armar una tabla de Ítems verdaderos que contenga:

| Nombre item | Modelo | P1 | P2 | ….. | P10 |
| --- | --- | --- | --- | --- | --- |
| R1/It1 | Rasch | Valor parámetro (aleatorio-runif)) |  |  |  |
| M2P1 | 2PL | Valor parámetro (aleatorio-runif) | Valor parámetro (aleatorio-runif) |  |  |
| ……... |  |  |  |  |  |

* Nombre Item(puedes ser alfanumérico)
* Modelo (nombre del modelo)
* Cantidad de itemes (n1,ítems de un tipo) ( n2 items de otro tipo), etc
* Luego tantas columnas como sean necesarias para abarcar los parámetros de los modelos por ejemplo P1,P2,P3, etc (los parámetros tendrán significado en función del nombre del modelo y cuando sobran se dejan vacío). En el caso de los politómicos lleva más parámetros por las opciones. Definir itemes de 3 o 4 opciones pero no vamos a variar la cantidad de opciones, lo hacemos para un tipo solo.
* Los parámetros se generan aleatoriamente.
* Esto se arma una sola vez (por ej **función genbanco.R**) . Se guarda el objeto como Rdata o csv como se quiera. Se levanta al inicio entonces el banco y se parametriza que se va a usar pero ejemplo solo Rasch, o 2PL y 3PL. Esto nos permite luego generar mezcla de tipos.

**2. Generación de base de respuestas**

A partir del banco de itemes verdaderos se genera la base de respuestas para diferentes sujetos por ej n=1000 o n=750 para cada ítem de cada modelo (**función genresp.R**) Aca vamos tener mucho más de lo necesario se corre una sola vez y se guarda Rdata o csv.

**3. Diferentes experimentos**

**Objetivo 1.** Queremos observar como punto 1 el error cuadrático medio cometido en la estimación y el sesgo (puede agregarse alguna otra medida) según el procedimiento de selección de ítems MI,KL,SH,PI.Aleatorio.

Ante cada tipo de experimento hay que generar un banco calibrado **función caibrobanco.R**. Los parámetros serían:

a) vector con tipo de modelo,

b) tamaño de cada tipo de modelo,

c) modelo usado para calibrar

Ejemplo:

**calibrabanco( c(“Rasch”,”3pl”),(100,50),”2PL”)**

En el TAI la función calibrabanco se tiene que correr 3 veces:

calibrabanco( c(“Rasch”,”3pl”),(100,50),”2PL”)

calibrabanco( c(“Rasch”,”3pl”),(100,50),”NP”)

calibrabanco( c(“Rasch”,”3pl”),(100,50),”Isotono”)

Otra opción es programarlo para que estime siempre con NP y el isotono y que varie lo parámetrico asi se corre una sola vez y siempre genera la NP y la isotona (bajo cualquier experimento siempre quiero tener la NP y la isotona)

calibrabanco( c(“Rasch”,”3pl”),(100,50),”2PL”) (esta función va a generar una lista que tiene long 3 en [[1]] va la estimación del banco según 2PL que es lo que pedí, en [[2]] va la estimación nopar en una grilla, [[3]] la estimación isótona en una grilla.)

**Tipo Experimento 1**

Considerar un banco solo con modelos logísticos. Este tipo de experimento sería el que sale a partir de bancos con ítems bien generados. Acá queremos demostrar que nuestro método funciona tan bien como el que debería ser, estoy haciendo todo bien: genero un banco con 2PL (items 2PL), los estimo con 2PL, y aplico TAI.

**Ejemplo** tipo experimento 1

Por ejemplo un banco generado con modelo 2PL. A partir de “Respuestas” se toman los ítems 2PL que quiera (dejamos parametrizado el tamaño del banco pero inicialmente lo haremos igual para todos los experimentos).

Luego calibramos el banco usando el modelo adecuado en este caso 2PL, el modelo no paramétrico y el modelo isótono (#bancocalibrado). Luego aplicamos el TAI y seleccionamos los ítems de acuerdo a los diferentes métodos. Hay métodos que no interviene la calibración:

- En el método aleatorio no interviene calibración

- En Max. Info interviene el método de MaxVer para 2PL

- En Pseudo Info interviene el modelo isótono

- En KL interviene MV2PL, NP, INP

- En ESH ídem

Se simulan n sujetos y a cada uno de se aplica el TAI y los diferentes caminos dando en cada punto de break la estimación de theta mediante cada método. Para cada individuo tenemos 9 trayectorias (9 estimaciones):

|  |  | Método selección de itemes | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Aleatorio | MI | KL | ESH | Pseudo-I |
| Método Calibración |  | 1 |  |  |  |  |
| Paramétrica |  | 2 | 3 | 6 |  |
| No Paramétrica |  |  | 4 | 7 |  |
| NP Isotono |  |  | 5 | 8 | 9 |

Vamos a tener por ejemplo una vez aplicados 5 items, n estimaciones de theta siguiendo el camino de MI, n estimaciones por selección aleatoria, n estimación por KL etc.

(#puntos de corte) Los mismo a los 10 items, los 15, los 20, los 25, los 30, etc hasta 50.

En cada punto de break tenemos el sesgo, el error cuadrático medio y otras medidas de ajustes que definamos. Además podemos mirar la tasa de exposición de los itemes.

Propuesta Mario: Armar un cuerpo principal del TAI donde se definen ciertos parámetros. TAI {#sujetos, #banco calibrado, #puntos de corte} y se llaman a las diferentes funciones (por fuera) de selección de itemes. Modularizar el TAI. A futuro luego podemos incluir un módulo que sea llamado para control de exposición.

**Tipo experimento 2** El banco generado con ítems logísticos por ejemplo Rasch pero estimados con 3PL. Acá hacemos un error en el método de estimación. Acá esperamos que el método NP funcione mejor ya que siempre se ajusta.

**Tipo de experimento 3** Contamino el banco por ejemplo con ítems logit cúbicos. Idem esperamos que el NP funcione mejor.

Ejemplo:

**calibrabanco( c(“Rasch”, “Logit cúbico”),(100,50),”Rasch”)**

**4. Comentarios finales**

* El objetivo final es obtener las primeras medidas de ver como nos da el algoritmo con pseudo info.
* Luego que tengamos los posibles experimentos parametrizados vamos a seleccionar cuales vamos a realizar.
* En los experimentos el tamaño muestral en principio no lo variamos
* En estos experimentos no nos interesa el criterio de parada, analizamos la precisión. Ejercicio posterior: evaluar al setear un valor de precisión requerido el promedio de la longitud del TAI según los distintos métodos. En ese caso le imputaremos el control de precisión con un epsilon y no los puntos de corte (breaks) dentro del TAI.
* Luego que tengamos todo armado para dicotómico y politómico en lo unidimensional y en base al análisis de los resultados pasamos a modelos multidimensionales.