

Package ‘microEnem’

October 20, 2023

Title Manipulação dos Microdados do Enem

Version 0.0.0.9000

Author Alexandre Jaloto [aut, cre] (<<https://orcid.org/0000-0002-5291-1768>>)

Description Manipulação dos microdados do Exame Nacional do Ensino Médio (Enem). Calcula a nota na escala oficial do Enem.

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.2

Imports data.table,
mirt,
mirtCAT

URL <https://github.com/alexandrejaloto/microEnem>

BugReports <https://github.com/alexandrejaloto/microEnem/issues>

Depends R (>= 2.10)

R topics documented:

abre.resp	1
calc.nota	2
constantes	3
mod.caderno	5

Index	6
--------------	----------

abre.resp	<i>Abre a resposta</i>
-----------	------------------------

Description

Abre o vetor de resposta da variável TX_RESPOSTA

Usage

abre.resp(unico)

Arguments

`unico` vetor de respostas de um ou mais sujeitos no formato da variável TX_RESPOSTA dos microdados do Enem

Value

Um objeto de classe `matrix` em que as colunas são os itens e as linhas são os sujeitos

Examples

```
# importar primeiros 100 casos dos microdados de 2019
micro <- data.table::fread('MICRODADOS_ENEM_2019.csv', nrow = 100)

resp <- abre.resp(micro$TX_RESPOSTAS_LC)
```

calc.nota

Calcula nota do Enem

Description

Calcula a nota de um sujeito a escala do Enem (500, 100)

Usage

```
calc.nota(resps, codigo = NULL, lingua = NULL)
```

Arguments

`resps` vetor de respostas de um ou mais sujeitos. O vetor deve ser no mesmo formato da variável TX_RESPOSTA dos microdados do Enem. É importante destacar que desde 2010 nos cadernos de Linguagens e Códigos existem cinco itens de língua inglesa e cinco itens de língua espanhola, por isso em algumas edições o vetor de respostas nessa área possui 50 caracteres. A pessoa responde somente uma dessas línguas e nos microdados as respostas aos cinco itens da outra língua são marcados com "9". A função `calcula.nota` automaticamente transforma essa resposta "9" em NA e esses itens não considerados para o cálculo.

`codigo` código da prova disponível no dicionário dos microdados. Essa informação também está disponível no objeto `dic.cad` deste pacote

`lingua` vetor com indicação da língua estrangeira escolhida pela pessoa. Essa informação corresponde à variável TP_LINGUA nos microdados (0 para inglês, 1 para espanhol).

Details

A nota é calculada pelo método *expected a posteriori* (EAP), com 40 pontos de quadratura de -4 a 4. A média da distribuição priori é 0 e o desvio padrão, 1. No Enem as notas estão em uma escala com média 500 e desvio padrão 100. A referência dessa escala são os concluintes regulares de escola pública do Enem 2009. Ou seja, a média desses alunos no Enem 2009 foi 500 e o desvio padrão, 100.

Os parâmetros dos itens estão divulgados em uma escala com média 0 e desvio padrão 1. A referência dessa escala é a amostra utilizada na primeira calibração dos itens do Enem, em 2009. Para

posicionar os parâmetros na a escala oficial do Enem, aplicamos as seguintes equações de transformação:

$$a_{enem} = \frac{a_{01}}{s}$$

$$b_{enem} = b_{01} * s + m$$

$$c_{enem} = c_{01}$$

Onde a_{enem} , b_{enem} e c_{enem} são os parâmetros dos itens na escala oficial do Enem, a_{01} , b_{01} e c_{01} são os parâmetros dos itens divulgados nos microdados, e k e d são as constantes de transformação da escala dos parâmetros divulgados para a escala oficial. Essas constantes estão disponibilizadas no objeto `constantes` deste pacote.

Value

as notas na escala oficial do Enem

Examples

```
# importar primeiros 100 casos dos microdados de 2019
micro <- data.table::fread('MICRODADOS_ENEM_2019.csv', nrows = 100)

# selecionar quatro casos do cadernos 511 (LC, primeira aplicação)
resp <- micro$TX_RESPOSTAS_LC[c(84, 97:99)]

# calcular a nota
nota <- calc.nota(resp, codigo = 511)
nota

# comparar com a nota oficial
micro$NU_NOTA_LC[c(84, 97:99)]

# calcular a nota de um sujeito que não está nos microdados
# vetor de resposta de uma pessoa fictícia que respondeu o caderno de Linguagens e Códigos em 2019
resp <- c('99999BBDABBDBAADCBABBADAACBDDDDDEACBCACAAABBBECBEC')

# calcular a nota
nota <- calc.nota(resp, codigo = 511)
nota
```

constantes

Constantes de transformação

Description

Constantes de transformação da escala dos parâmetros divulgados para a escala oficial do Enem

Usage

`constantes`

Format

Um objeto do tipo `data.frame` com as constantes de cada área.

Temos duas escalas: a escala oficial do Enem, cuja referência são os concluintes regulares de escolas públicas do Enem 2009 (média 500, desvio padrão 100); a escala dos parâmetros dos microdados, cuja referência é a amostra de calibração do Enem 2009 (média 0, desvio padrão 1).

Para calcular as constantes de transformação de uma escala para a outra, aplicamos a equalização linear (Hambleton et al., 1991). Inicialmente estimamos a nota dos primeiros 300.000 sujeitos do banco dos microdados em cada área, de acordo com as especificações em `calcula.nota`. Em seguida, estabelecemos a igualdade entre a padronização das notas desses sujeitos:

$$\frac{Y_i - \bar{Y}}{DP_y} = \frac{X_i - \bar{X}}{DP_x}$$

Onde Y_i representa a nota do sujeito i na escala oficial do Enem e \bar{Y} e DP_y representam a média e o desvio padrão das notas dessa amostra de 300.000 sujeitos nessa mesma escala. X_i representa a nota do sujeito i na escala dos parâmetros divulgados. \bar{Y} e DP_y representam a média e o desvio padrão das notas da amostra nessa escala. Onde as notações com Y e y se referem às notas na escala oficial do Enem, e as notações com X e x se referem às notas na escala dos parâmetros divulgados. Y_i e X_i correspondem à nota do sujeito i , \bar{Y} e \bar{X} são a média desses 300.000 sujeitos e DP_y e DP_x são o desvio padrão dessa amostra.

Nessa equação, consideramos que as notas padronizadas das duas métricas são iguais, pois provêm da mesma amostra. Porém, na prática elas não são exatamente iguais, pois em uma calibração é pouco provável que a distribuição de habilidades da amostra selecionada tenha média exatamente zero e desvio padrão exatamente 1, apesar de os programas (como o pacote `mirt`) assumirem esses valores para identificação da métrica. Se isolarmos Y_i , teremos (Muñiz, 1997):

$$Y_i = \frac{DP_y}{DP_x} X_i + \bar{Y} - \frac{DP_y}{DP_x} \bar{X}$$

Dessa equação, extraímos as constantes k e d :

$$k = \frac{DP_y}{DP_x}$$

$$d = \bar{Y} - \frac{DP_y}{DP_x} \bar{X}$$

$$d = \bar{Y} - k\bar{X}$$

Portanto, para transformar a nota da escala dos microdados para a escala oficial do Enem, utilizamos a seguinte equação:

$$Y_i = kX_i + d$$

mod.caderno	<i>Modelo mirt do caderno</i>
-------------	-------------------------------

Description

Gera um modelo do pacote mirt do caderno do Enem

Usage

```
mod.caderno(codigo = NULL)
```

Arguments

codigo	código da prova disponível no dicionário dos microdados. Essa informação também está disponível no objeto dic.cad deste pacote
--------	--

Value

Um objeto mirt com informações sobre os parâmetros dos itens e da amostra de calibração. Para informações sobre esse objeto, consulte a documentação do pacote mirt.

Examples

```
mod <- mod.caderno(511)
mirt::coef(mod, IRTpars = TRUE, simplify = TRUE)
```

Index

* **datasets**

constantes, [3](#)

abre.resp, [1](#)

calc.nota, [2](#)

constantes, [3](#)

mod.caderno, [5](#)