# Package 'harmonious'

July 14, 2024

**Type** Package

**Title** Explanatory Item Response modeling in CmdStanR

**Version** 0.1.0

**Author** Nathan DePuy

**Maintainer** Nathan DePuy <depy@uiowa.edu>

**Description** A package for performing Bayesian inference on Explanatory Item Repsonse models in CmdStanR

**Depends** R (>= 4.0.0),
bayesplot,
Cairo,
ggplot2,
pander,
mathjaxr,
loo,
remotes,
instantiate,
cmdstanr

**Suggests** testthat (>= 3.0.0)

**Additional_repositories** https://mc-stan.org/r-packages/

**SystemRequirements** CmdStan (https://mc-stan.org/users/interfaces/cmdstan)

**Encoding** UTF-8

**LazyData** true

**Config/testthat/edition** 3

**RoxygenNote** 7.3.2

**Roxygen** list(markdown=TRUE)

**RdMacros** mathjaxr

**Remotes** stan-dev/cmdstanr

## Contents

---

`CreateMod` *Create Adjunct CmdStan Model Environment*

---

## Description

Creates an environment to estimate the $p \times i$ interaction model with CmdStan.

## Usage

```
CreateMod(
  coef_hyper,
  sd_hyper,
  nWarmup_init,
  nSamples_init,
  nWarmup_run,
  nSamples_run,
  aux_envir
)
```

## Arguments

| | |
|---|---|
| `coef_hyper` | Hyperparameter value for the standard deviation of normally distributed parameters |
| `sd_hyper` | Hyperparameter value for the shape parameter of gamma distributed parameters |
| `nWarmup_init` | Number of burn-in draws for the fixed $\theta$ model |
| `nWarmup_run` | Number of burn-in draws for the free $\theta$ model |
| `nSamples_run` | Number of sampled posterior values for the free $\theta$ model (after burn-in has completed) |
| `aux_envir` | An environment object that contains objects to be loaded into the `CreateMod` environment (such as is returned by the `genData` function) |
| `nSampels_init` | Number of posterior draws for the fixed $\theta$ model (after burn-in has completed) |

## Details

Creates an environment with methods for fitting the fixed $\theta$ and free $\theta$ models. Additionally, `CreateMod` includes a method for recursively checking $\hat{R}$ convergence.

## Value

an environment object containing fixed $\theta$ and free $\theta$ model results

## See Also

`initialize`, `sample`, `rhatCheck`

genData *Data Generating Function for the $p \times i$ Interaction Model*

## Description

Generates data for a $p \times i$ Interaction Model under the Explanatory Item Response theoretical framework.

## Usage

```
genData(P, I, J, K, seed = NULL, isCorrI = TRUE)
```

## Arguments

| | |
|---|---|
| P | Number of examinees |
| I | Number of items |
| J | Number of potential item feature categories |
| K | Number of potential person feature categories |
| seed | Integer seed for replication (if NULL, a random seed will be generated) |
| isCorrI | Draw Cholesky-factorized lower triangular matrix for modeling structured item residuals? |

## Details

Let a model of the log-odds transformed outcome $(\eta_{pi})$ be defined as:

$$\eta_{pi} = \theta_p \lambda_i + \tau_i + \sum_{k=1}^{K} \sum_{j=1}^{J} \beta_{jk_{(\eta)}} x_{ik} z_{pj}$$

where the linear predictor of latent trait measurements $\theta_p$is defined as:

$$\theta_p = \sum_{j=1}^{J} \beta_{j_{(\theta)}} z_{pj} + u_{p_{(\theta)}}$$

where the linear predictor of item slope/discrimination terms $(\lambda_i)$ is defined as:

$$\lambda_i = \sum_{k=1}^{K} \beta_{k_{(\lambda)}} x_{ik} + u_{i_{(\lambda)}}$$

and where the linear predictor of item intercept/easiness terms $(\tau_i)$ is defined as:

$$\tau_i = \sum_{k=1}^{K} \beta_{k_{(\tau)}} x_{ik} + u_{i_{(\tau)}}$$

For the linear predictor of latent trait measurements, the following assumption is made on its residual terms for the purposes of model identification:

$$u_{p_{(\theta)}} \sim \mathcal{N}(0, 1)$$

If an assumption is made that there is no strcutured dependency between the residual terms for the linear predictors of item intercepts and item slopes (`isCorrI = FALSE`), then:

$$u_{i_{(\lambda)}} \sim \mathcal{N}(0, \sigma_{\lambda}^2)$$
$$u_{i_{(\tau)}} \sim \mathcal{N}(0, \sigma_{\tau}^2)$$

Whereas if it is assumed that the residuals of the linear predictors of item intercepts and item slopes is assumed to have a predictable underlying structure (`isCorrI = TRUE`), then:

$$\begin{bmatrix} u_{i_{(\lambda)}} \\ u_{i_{(\tau)}} \end{bmatrix} \sim \mathcal{MVN}\left(\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{\Sigma} = \begin{bmatrix} \sigma_{\lambda}^2 & \sigma_{\lambda}\sigma_{\tau} \\ \sigma_{\tau}\sigma_{\lambda} & \sigma_{\tau}^2 \end{bmatrix}\right)$$

## Value

an environment containing simulated data

---

`genReport` *Generate Text Report of Free $\theta$ Model Results*

---

## Description

Helper function that generates a text document in the user-specified save directory

## Usage

```
genReport(saveDir, fileDetails, ...)
```

## Arguments

| | |
|---|---|
| `saveDir` | Path to directory where report is saved |
| `fileDetails` | string identifier for name of saved report |
| `...` | additional arguments passed from parent frame |

## Details

Model results include posterior descriptives and simulated data information

## Value

a `.txt` file of the free $\theta$ results

## initialize — *Run Fixed $\theta$ Model*

### Description

Runs the first step in the two-step estimation framework

### Usage

```
initialize(...)
```

### Arguments

| | |
|---|---|
| `...` | arguments passed from parent frame |

### Details

The first step in the two-step estimation framework for the $p \times i$ model

### Value

Model results for the standardized $\theta$ model added to the environment rendered by `CreateMod`

### See Also

`CreateMod`

## rCorr — *Randomly Sample a Cholesky Factorized, Lower Triangular Matrix from a Lewandowski-Kurowicka-Joe Distribution*

### Description

Generates a lower triangular, invertible, positive-definite matrix

### Usage

```
rCorr(nDim, eta = 1)
```

### Arguments

| | |
|---|---|
| `nDim` | desired dimension of the sampled $n \times n$ matrix |
| `eta` | concentration hyperparameter |

### Details

Using the onion method (Ghosh & Henderson, 2003), samples a Cholesky factorized lower-triangular matrix $L$ from a Lewandowski-Kurowicka-Joe (LKJ) distribution given the concentration hyperparameter $\eta$, such that:

$$A = LL^{\top}$$

where $A$ is a $n \times n$ matrix and is a member of the set of all symmetric, positive-definite matrices.

## Value

a lower triangular $n \times n$ matrix

## References

Ghosh, S. and Henderson, S. G. (2003). Behavior of the norta method for correlated random vector generation as the dimension increases. ACM Transactions on Modeling and Computer Simulation, 13(3), 276-294. https://doi.org/10.1145/937332.937336

---

`rhatCheck` *Recursive Validation for Rhat Convergence*

---

## Description

Helper function that validates $\hat{R}$ convergence.

## Usage

```
rhatCheck(...)
```

## Arguments

| | |
|---|---|
| ... | arguments passed from parent frame |

## Details

**Input:** `CmdStan` model, $N_{\max}$, $\hat{R}_1$
Sample `CmdStan` model
  **while** $\hat{R}_n \leq \hat{R}_{thresh}$ **or** $n \leq N_{\max}$ **do**
    Sample `CmdStan` model
    $n \mathrel{+}= 1$
  **end do**
**end while**

## Value

modified free $\theta$ model that is added to an environment created by `CreateMod`

## References

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P. (2021). Rank-normalization, folding, and localization: an improved r^ for assessing convergence of mcmc (with discussion). Bayesian Analysis, 16(2). https://doi.org/10.1214/20-ba1221

## See Also

`CreateMod`

`sample` *Run Free $\theta$ Model*

## Description

Runs the second step of the two-step estiamtion framework

## Usage

```
sample(...)
```

## Arguments

| | |
|---|---|
| `...` | arguments passed from parent frame |

## Details

The second step treating latent trait measurements ($\theta_p$) as freely estimated parameters in the two-step estimation framework for the $p \times i$ model

## Value

Model results for the full model added to the environment rendered by `CreateMod`

## See Also

`CreateMod`

# Index