Package: Bayesrel (via r-universe)

August 27, 2024

Type Package

Title Bayesian Reliability Estimation

Version 0.7.7

Date 2023-08-08

Description Functionality for reliability estimates. For 'unidimensional' tests: Coefficient alpha, 'Guttman's' lambda-2/-4/-6, the Greatest lower bound and coefficient omega_u ('unidimensional') in a Bayesian and a frequentist version. For multidimensional tests: omega_t (total) and omega_h (hierarchical). The results include confidence and credible intervals, the probability of a coefficient being larger than a cutoff, and a check for the factor models, necessary for the omega coefficients. The method for the Bayesian 'unidimensional' estimates, except for omega_u, is sampling from the posterior inverse 'Wishart' for the covariance matrix based measures (see 'Murphy', 2007, <https://groups.seas.harvard.edu/courses/cs281/papers/murphy-2007.pdf>. The Bayesian omegas (u, t, and h) are obtained by 'Gibbs' sampling from the conditional posterior distributions of (1) the single factor model, (2) the second-order factor model, (3) the bi-factor model, (4) the correlated factor model ('Lee', 2007, <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470024737>).

URL https://github.com/juliuspfadt/Bayesrel

BugReports https://github.com/juliuspfadt/Bayesrel/issues

License GPL-3

Encoding UTF-8

LazyData true

Imports LaplacesDemon, MASS, lavaan, coda, methods, stats, graphics, Rdpack, Rcpp (>= 1.0.4.6)

LinkingTo Rcpp, RcppArmadillo

RdMacros Rdpack

RoxygenNote 7.2.3

Depends R (>= 2.10)

Suggests knitr, rmarkdown, tinytest

Repository https://juliuspfadt.r-universe.dev

RemoteUrl https://github.com/juliuspfadt/bayesrel

RemoteRef HEAD

RemoteSha ce90973d5b79e3b1515ff77114e27e1de4c5126d

Contents

---

asrm   *5-Item questionnaire data from Nicolai (2018)*

Description

A dataset consisting of 78 participants who filled out the 5-item Altman Self-Rating Mania Scale, rating from 1 to 5 on a Likert scale

Usage

    asrm

Format

The format is a 5-column datamatrix containing 78 observations

Source

article

References

Nicolai, J., & Moshagen, M. (2018). Pathological buying symptoms are associated with distortions in judging elapsed time. Journal of Behavioral Addictions, 7(3), 752-759.

---

asrm_mis	*5-Item questionnaire data from Nicolai (2018) with 10 % missings*

Description

A dataset consisting of 78 participants who filled out the 5-item Altman Self-Rating Mania Scale, rating from 1 to 5 on a Likert scale, 10 % missings were inserted at random

Usage

```
asrm_mis
```

Format

The format is a 5-column data matrix containing 78 observations, missing are NA

Source

article

References

Nicolai, J., & Moshagen, M. (2018). Pathological buying symptoms are associated with distortions in judging elapsed time. Journal of Behavioral Addictions, 7(3), 752-759.

---

bomegas	*Estimate Bayesian reliability estimates for multidimensional scales following common factor models*

Description

When supplying a multidimensional data set the function estimates the reliability of the set by means of omega_total and the general factor saturation of the set by means of omega_hierarchical. The data may follow multiple multi-factorial factor models: a) the second-order factor model (crossloadings possible), b) the bi-factor model (no crossloadings), c) the correlated factor model (crossloadings possible, only omega_t) Error-covariances are not estimable.

The prior distributions of omega_t and omega_h are computed from the prior distributions of the respective factor model parameters. Specifically, normal distributions for the factor loadings and factor scores; an inverse gamma distribution for the manifest and latent residuals; an inverse Wishart distribution for the covariance matrix of the latent variables (correlated model). A Gibbs sampler

iteratively draws samples from the conditional posterior distributions of the factor model parameters. The posterior distributions of omega_t and omega_h are computed from the posterior samples of the factor model parameters.

The output contains the posterior distributions of omega_t and omega_h (only for second-order and bi-factor model), their mean, and credible intervals. If desired, one may also find the posterior implied covariance matrices, and the posterior factor model parameters in the output.

Usage

```
bomegas(
  data,
  n.factors = NULL,
  model = NULL,
  model.type = "second-order",
  n.iter = 2000,
  n.burnin = 200,
  n.chains = 3,
  thin = 1,
  interval = 0.95,
  missing = "impute",
  a0 = NA,
  b0 = NA,
  l0 = NA,
  A0 = NA,
  c0 = NA,
  d0 = NA,
  beta0 = NA,
  B0 = NA,
  p0 = NA,
  R0 = NA,
  param.out = FALSE,
  callback = function() {
  },
  disableMcmcCheck = FALSE
)
```

Arguments

data	A matrix or data.frame containing multivariate observations, rows = observations, columns = variables/items
n.factors	A number specifying the number of group factors that the items load on
model	A string that by default NULL (=balanced) distributes the items evenly among the number of group factors. This only works if the items are a multiple of the number of group factors and the items are already grouped in the data set, meaning, e.g., items 1-5 load on one factor, 6-10 on another, and so on. A model file can be specified in lavaan syntax style (f1=~.+.+.) to relate the items to the group factors. The items' names need to equal the column names in the data set, aka the variable names

model.type	A string indicating the factor model estimated, by default this is the "second-order" model. Another option is the "bi-factor" model; and eventually "correlated" for the correlated factor model
n.iter	A number for the iterations of the Gibbs Sampler
n.burnin	A number for the burnin in the Gibbs Sampler
n.chains	A number for the chains to run for the MCMC sampling
thin	A number for the thinning of the MCMC samples
interval	A number specifying the credible interval, the interval is the highest posterior density interval (HPD)
missing	A string denoting the missing data handling, can be "impute" or "listwise". With impute the missing data will be estimated during the MCMC sampling as further unknown parameters
a0	A number for the shape of the prior inverse gamma distribution for the manifest residual variances, when left as NA, the default value is set to 2
b0	A number for the scale of the prior inverse gamma distribution for the manifest residual variances, when left as NA, the default value is set to 1
l0	A number for the mean of the prior normal distribution for the manifest loadings, when left as NA, the default value is set to 0; can be a single value or a loading matrix
A0	A number for scaling the variance of the prior normal distribution for the manifest loadings, when left as NA, the default value is set to 1
c0	A number for the shape of the prior inverse gamma distribution for the latent residual variances, when left as NA, the default value is set to 2; necessary only for the second-order model
d0	A number for the scale of the prior inverse gamma distribution for the latent residual variances, when left as NA, the default value is set to 1, necessary only for the second-order model
beta0	A number for the mean of the prior normal distribution for the g-factor loadings, when left as NA, the default value is set to 0, can be a single value or a vector
B0	A number for scaling the variance of the prior normal distribution for the g-factor loadings, when left as NA, the default value is set to 2.5
p0	A number for the shape of the prior inverse gamma distribution for the variance of the g-factor, when left as NA, the default value is set to q^2-q when q are the number of group factors for the second-order and bi-factor model
R0	A number for the scale of the prior inverse gamma distribution for the variance of the g-factor, when left as NA, the default value is set to the number of items for the second-order and bi-factor model
param.out	A logical indicating if loadings and residual variances should be attached to the result, by default FALSE because it saves memory
callback	An empty function for implementing a progressbar call from a higher program (e.g., JASP)
disableMcmcCheck	A logical disabling the MCMC settings check for running tests and the likes

Value

The posterior means and the highest posterior density intervals for omega_t and omega_h (for the second-order and bi-factor model)

References

Lee S (2007). *Structural equation modeling: A Bayesian approach*. John Wiley & Sons.

Examples

```
# specify the model syntax relating the group factors to the items:
model <- "f1 =~ U17_r + U22_r + U29_r + U34_r
f2 =~ U4 + U14 + U19 + U27
f3 =~ U6 + U16 + U28 + U48
f4 =~ U23_r + U31_r + U36_r + U46_r
f5 =~ U10_r + U20_r + U35_r + U52_r"
# specifying the model can be omitted if the model structure is simple and balanced,
# meaning the an equal number of items load on each factor, and the items relate to the factors
# in the same order as they appear in the data.
# Note that the iterations are set very low for smoother running examples, you should use
# at least the defaults:
res <- bomegas(upps, n.factors = 5, model = NULL, n.iter = 200, n.burnin = 50,
n.chains = 2, missing = "listwise", model.type = "second-order")
```

---

cavalini   *8-Item Questionnaire Data from Cavalini (1992)*

Description

A dataset consisting of eight item questionnaire data. It's Likert scaled from 0-3. It is data measuring how annoyed people were by malodors

Usage

```
cavalini
```

Format

The format is a 8-column datamatrix containing 828 observations

Source

Doctoral Dissertation

References

Cavalini, P. M. (1992). It's an ill wind that brings no good: Studies on odour annoyance and the dispersion of odorant concentrations from industries. Rijksuniversiteit Groningen.

| multiFit | *model fit for the second-order and bi-factor model,* |

Description

Fit indices and posterior predictive check for the multidimensional model: comparison between posterior sample of model implied covariance matrices and sample covariance matrix. Gray bars should enclose the black dots for good fit. Also prints fit indices, LR (likelihood-ratio), RMSEA, SRMR. The RMSEA is from Garnier-Villareal & Jorgensen (2020)

Usage

```
multiFit(x, data, ppc = TRUE, cutoff = 0.08, ci = 0.9)
```

Arguments

x	A bomegas output object (list)
data	A matrix or data.frame containing the data set that produced x
ppc	A logical indicating if the PPC should be printed or not, the default is TRUE
cutoff	A value to compare the posterior sample of RMSEAs against. The result will contain the probability that the RMSEA is smaller than the cutoff value
ci	A value between 0 and 1 indicating the credible interval for the RMSEA

References

Garnier-Villarreal M, Jorgensen TD (2020). "Adapting Fit Indices for Bayesian Structural Equation Modeling: Comparison to Maximum Likelihood." *Psychological Methods*, **25**(1), 46–70.

Examples

```
multiFit(bomegas(upps, n.factors = 5, n.chains = 2, n.iter = 150,
n.burnin = 50, missing = "listwise"), upps)
```

| omegaFit | *graphical posterior predictive check for the 1-factor omega model, based on covariance matrix eigenvalues* |

Description

gives posterior predictive check for the 1-factor model: comparison between model implied covariance matrix and sample covariance matrix also displays frequentist fit indices

Usage

```
omegaFit(x, data, ppc = TRUE, cutoff = 0.08, ci = 0.9)
```

Arguments

x	A strel output object (list)
data	A matrix or data.frame containing the data set that produced x
ppc	A logical indicating if the PPC should be printed or not, the default is TRUE
cutoff	A value to compare the posterior sample of RMSEAs against. The result will contain the probability that the RMSEA is smaller than the cutoff value
ci	A value between 0 and 1 indicating the credible interval for the RMSEA

References

Garnier-Villarreal M, Jorgensen TD (2020). "Adapting Fit Indices for Bayesian Structural Equation Modeling: Comparison to Maximum Likelihood." *Psychological Methods*, **25**(1), 46–70.

Examples

```
omegaFit(strel(asrm, "omega", n.chains = 2, n.iter = 200), data = asrm)
```

---

omegasCFA	*Estimate reliability estimates for multidimensional scales in the frequentist framework*

Description

When supplying a data set that is multidimensional the function estimates the reliability of the set by means of omega_total and the general factor saturation of the set by means of omega_hierarchical The procedure entails fitting a hierarchical factor model using a CFA. The second-order (hierarchical, higher-order), the bi-factor, and the correlated factor model can be used in the CFA. The CFA is performed using lavaan 'Yves Rosseel', <https://CRAN.R-project.org/package=lavaan>. Coefficients omega_t and omega_h (only for second-order and bi-factor model) can be computed from the factor model parameters.

Usage

```
omegasCFA(
  data,
  n.factors = NULL,
  model = NULL,
  model.type = "second-order",
  interval = 0.95,
  missing = "fiml",
  fit.measures = FALSE
)
```

Arguments

`data`	A matrix or data.frame containing multivariate observations, rows = observations, columns = variables/items
`n.factors`	A number specifying the number of group factors that the items load on
`model`	A string that by default NULL (=balanced) distributes the items evenly among the number of group factors. This only works if the items are a multiple of the number of group factors and the items are already grouped in the data set, meaning, e.g., items 1-5 load on one factor, 6-10 on another, and so on. A model file can be specified in lavaan syntax style (f1=~.+.+.) to relate the items to the group factors. The items' names need to equal the column names in the data set, aka the variable names
`model.type`	A string denoting if the model that should be fit is the second-order or bi-factor model or the correlated factor model. This comes down to the researcher's theory about the measurement and the model fit.
`interval`	A number specifying the confidence interval, which is Wald-type
`missing`	A string denoting the missing data handling, can be "fiml" (full information ML) or "listwise". Specifying "pairwise" will defulat to "fiml"
`fit.measures`	A logical denoting if fit.measures from the CFA should be computed, the output then contains the chisq statistic, chisq df, chisq p-value, cfi, tli, rmsea, rmsea 90% ci lower, rmsea 90% ci upper, rmsea<.05 p-value, aic, bic, unbiased srmr, unbiased srmr 90% ci lower, unbiased srmr 90% ci upper, unbiased srmr<.05 p-value

Value

The point estimates and the Wald-type confidence intervals for omega_t and omega_h (for the second-order and bi-factor model)

Examples

```
res <- omegasCFA(upps, n.factors = 5, model = NULL, model.type = "bi-factor",
missing = "listwise")

# or with specified model syntax relating the group factors to the items:
model <- "f1 =~ U17_r + U22_r + U29_r + U34_r
f2 =~ U4 + U14 + U19 + U27
f3 =~ U6 + U16 + U28 + U48
f4 =~ U23_r + U31_r + U36_r + U46_r
f5 =~ U10_r + U20_r + U35_r + U52_r"
res <- omegasCFA(upps, n.factors = 5, model = model, model.type = "second-order",
missing = "listwise")
```

---

pOmegas	*prior and posterior probability of omega_t and omega_h being larger than thresholds*

Description

takes mcmc posterior samples of omega_t and omega_h and calculates the prior and posterior probability of the estimate being bigger or smaller than an arbitrary value

Usage

```
pOmegas(x, cutoff.t = 0.8, cutoff.h = 0.6)
```

Arguments

x	A bomegas output object (list)
cutoff.t	A number indicating the threshold for omega_t
cutoff.h	A number indicating the threshold for omega_h

Examples

```
pOmegas(bomegas(upps, n.factors = 5, n.chains = 2, n.iter = 150, n.burnin = 50,
disableMcmcCheck = TRUE, missing = "listwise"))
```

---

pStrel	*prior and posterior probability of estimate being bigger than threshold*

Description

takes a mcmc posterior sample of any of the single test reliability estimates and calculates the prior and posterior probability of the estimate being bigger or smaller than an arbitrary value (priors are stored in the package)

Usage

```
pStrel(x, estimate, low.bound)
```

Arguments

x	A strel output object (list)
estimate	A character string indicating what estimate to plot from the strel output object
low.bound	A number for the threshold to be tested against

Examples

```
pStrel(strel(asrm, "lambda2", n.chains = 2, n.iter = 150, freq = FALSE),
"lambda2", .80)
```

---

strel | *Estimate single test reliability coefficients for unidimensional scales*

Description

Reliability estimation of alpha, lambda2, the glb, and omega in a Bayesian an frequentist way. The results include posterior and bootstrapped distributions, point estimates, credible intervals, and confidence intervals.

Usage

```
strel(
  data = NULL,
  estimates = c("alpha", "lambda2", "glb", "omega"),
  interval = 0.95,
  n.iter = 1000,
  n.burnin = 50,
  thin = 1,
  n.chains = 3,
  n.boot = 1000,
  cov.mat = NULL,
  n.obs = NULL,
  freq = TRUE,
  Bayes = TRUE,
  para.boot = FALSE,
  item.dropped = FALSE,
  missing = "pairwise",
  omega.freq.method = "cfa",
  omega.int.analytic = TRUE,
  alpha.int.analytic = TRUE,
  callback = function() {
  },
  k0 = 1e-10,
  df0 = NULL,
  a0 = 2,
  b0 = 1,
  m0 = 0,
  disableMcmcCheck = FALSE
)
```

Arguments

data	The dataset to be analyzed, observations are rows, items are columns
estimates	A character vector containing the estimands, we recommend using lambda4 with only a few items due to the computation time
interval	A number specifying the uncertainty interval

n.iter	A number for the iterations of the Gibbs Sampler
n.burnin	A number for the burnin in the Gibbs Sampler
thin	A number for the thinning of the MCMC samples
n.chains	A number for the chains to run for the MCMC sampling
n.boot	A number for the bootstrap samples
cov.mat	A covariance matrix can be supplied instead of a dataset, but number of observations needs to be specified
n.obs	A number for the sample observations when a covariance matrix is supplied and the factor model is calculated
freq	A logical for calculating the frequentist estimates
Bayes	A logical for calculating the Bayesian estimates
para.boot	A logical for calculating the parametric bootstrap, the default is the non-parametric
item.dropped	A logical for calculating the if-item-dropped statistics
missing	A string specifying the way to handle missing data, 'listwise' is self-explanatory, 'pairwise' in the Bayesian paradigm means sampling the missing values as additional parameters from the joint conditional distribution, in the frequentist paradigm this means using the 'pairwise' covariance matrix, except the full information ML method for omega
omega.freq.method	
	A character string for the method of frequentist omega, either "cfa" (confirmatory factor analysis), or "pfa" (principal factor analysis), with "pfa" the interval is always bootstrapped
omega.int.analytic	
	A logical for calculating the omega confidence interval analytically, only works with cfa as the omega.freq.method
alpha.int.analytic	
	A logical for calculating the alpha confidence interval analytically
callback	Empty function call for external use
k0	A scalar multiplier for the diagonal of the scaling matrix of the inverse Wishart prior distribution for alpha, lambda2, and the glb
df0	The degrees of freedom of the inverse Wishart prior distribution for alpha, lambda2, and the glb, the default is NULL, which sets the df as the number of items
a0	The shape parameter of the inverse gamma prior distribution for the residual variances in the single factor model for omega
b0	The scale parameter of the inverse gamma prior distribution for the residual variances in the single factor model for omega
m0	The prior mean of the normal distribution on the factor loadings for omega
disableMcmcCheck	
	A logical disabling the MCMC settings check for running tests and the likes

Details

Reported are point estimates (posterior mean), Bayesian credible intervals (highest posterior density) and frequentist confidence intervals (non parametric or parametric bootstrap). The estimates supported are Cronbach alpha, Guttman's lambda2/4/6, the glb, and Mcdonald's omega_u (unidimensional). Beware of lambda4 with many indicators, the computational effort is considerable. The glb method uses adjusted code from the 'Rcsdp' package by 'Hector Corrada Bravo', <https://CRAN.R-project.org/package=Rcsdp>. This process applies a slightly adjusted solving algorithm from the 'CSDP' library by 'Brian Borchers' <https://github.com/coin-or/Csdp/wiki>, <doi:10.1080/10556789908805765>, but is wrapped in 'RcppArmadillo'. Guttman's Lambda-4 method is from 'Benton' (2015) <doi:10.1007/978-3-319-07503-7_19>. The principal factor analysis (pfa) for a version of frequentist omega_u can be found in 'Rencher' (2007) and is described in 'Schlegel' (2017) <https://www.r-bloggers.com/2017/03/iterated-principal-factor-method-of-factor-analysis-with-r/>. Coefficients alpha, lambda2/4, and the glb are estimated from the data covariance matrix. Coefficient omega is estimated from the centered data matrix. The analytic confidence interval of alpha is from 'Bonett' and 'Wright' (2015) <doi:10.1002/job.1960>.

The prior distribution on Cronbach's alpha (as well as lambda2 and the glb) is induced by the prior distribution on the covariance matrix, which is an inverse Wishart distribution with the identity matrix (multiplied by a scalar) as a scaling matrix and the number of items k as the degrees of freedom. The prior distribution on McDonald's omega is induced by the prior distributions on the single-factor model parameters, which are: a normal distribution centered on zero for the factor loadings and scores; an inverse gamma distribution with shape=2 and scale=1 for the residuals; and for the variance of the latent variables an inverse Wishart distribution with the number of items k as a scaling matrix (scalar, since it is of dimension one) and the sum k+2 as the degrees of freedom.

Value

The basic output displays the interval bounds of the coefficients, highest posterior density intervals for the Bayesian coefficients, and confidence intervals for the frequentist coefficients. The summary output shows the point estimates of the coefficients together with the interval bounds. The point estimates for the Bayesian coefficients are posterior means.

References

Bonett DG, Wright TA (2015). "Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning." *Journal of Organizational Behavior*, **36**(1), 3–15. doi:10.1002/job.1960.

Murphy KP (2007). "Conjugate Bayesian analysis of the Gaussian distribution." University of British Columbia.

Lee S (2007). *Structural equation modeling: A Bayesian approach*. John Wiley & Sons.

Pfadt JM, van den Bergh D, Sijtsma K, Moshagen M, Wagenmakers E (2021). "Bayesian estimation of single-test reliability coefficients." *Multivariate Behavioral Research*, 1–30. doi:10.1080/00273171.2021.1891855.

Rencher AC (2002). *Methods of multivariate analysis*. John Wiley & Sons, Inc. doi:10.1002/0471271357.

Examples

```
# note that these are very few iterations just for the example execution,
# you should use the defaults at least
summary(strel(asrm, estimates = "lambda2", n.chains = 2, n.iter = 200, n.boot = 200))
summary(strel(asrm, estimates = "lambda2", item.dropped = TRUE, n.chains = 2,
n.iter = 200, freq = FALSE))
```

---

upps   *20-item questionnaire data from Lozano et al. (2018)*

Description

A dataset consisting of 455 participants who filled out the 20-item short form of the UPPS-P, and impulsivity scale, rating from 0 to 4 on a Likert scale. The scale has five subscales measured by four items each: negative urgency (columns 1-4), perseverance (columns 5-8), premeditation (columns 9-12), sensation seeking (columns 13-16), positive urgency (columns 17-20). The data contain 13 missing values.

Usage

```
upps
```

Format

The format is a 20-column datamatrix containing 455 observations

Source

article

References

Lozano, Ó. M., Díaz-Batanero, C., Rojas, A. J., Pilatti, A., & Fernández-Calderón, F. (2018). Concordance between the original and short version of the Impulsive Behaviour Scale UPPS-P using an IRT model. PLOS ONE,13(3), 1–15. https://doi.org/10.1371/journal.pone.0194390

Index