# Package 'multid'

December 21, 2022

**Title** Multivariate Difference Between Two Groups

**Version** 0.7.1

**Description** Estimation of multivariate differences between two groups (e.g., multivariate sex differences) with regularized regression methods and predictive approach. See Lönnqvist & Ilmarinen (2021) <doi:10.1007/s11109-021-09681-2> and Ilmarinen et al. (2022) <doi:10.1177/08902070221088155>.
Includes tools that help in understanding difference score reliability, predictions of difference score variables, conditional intra-class correlations, and heterogeneity of variance estimates. Package development was supported by the Academy of Finland research grant 338891.

**License** GPL-3

**Encoding** UTF-8

**BugReports** https://github.com/vjilmari/multid/issues

**RoxygenNote** 7.1.2

**Imports** dplyr (>= 1.0.7), glmnet (>= 4.1.2), stats (>= 4.0.2), pROC (>= 1.18.0), lavaan (>= 0.6.9), emmeans (>= 1.6.3), lme4 (>= 1.1.27.1), quantreg (>= 5.88)

**Suggests** knitr (>= 1.39), rmarkdown (>= 2.14), overlapping (>= 1.7), rio (>= 0.5.29)

**VignetteBuilder** knitr

**NeedsCompilation** no

**Author** Ville-Juhani Ilmarinen [aut, cre] (<https://orcid.org/0000-0001-9493-379X>)

**Maintainer** Ville-Juhani Ilmarinen <vj.ilmarinen@gmail.com>

**Repository** CRAN

**Date/Publication** 2022-12-21 10:00:02 UTC

## R topics documented:

---

`colwise_pool` *Column-wise pooling of standard deviations*

---

## Description

Column-wise pooling of standard deviations

## Usage

```
colwise_pool(data, n1, n2, m1, m2, sd1, sd2)
```

## Arguments

| | |
|---|---|
| `data` | Data frame of d_pooled_sd output for multiple samples. |
| `n1` | Sample sizes of group1. |
| `n2` | Sample sizes of group2. |
| `m1` | Means of group1. |
| `m2` | Means of group2. |
| `sd1` | Standard deviations of group1. |
| `sd2` | Standard deviations of group2. |

## Value

pooled SDs for groups and across groups

cvv *Coefficient of variance variation*

## Description

Calculates three different indices for variation between two or more variance estimates. VR = Variance ratio between the largest and the smallest variance. CVV = Coefficient of variance variation (Box, 1954). SVH = Standardized variance heterogeneity (Ruscio & Roche, 2012).

## Usage

```
cvv(data)
```

## Arguments

| | |
|---|---|
| data | Data frame of two or more columns or list of two or more variables. |

## Value

A vector including VR, CVV, and SVH.

## References

Box, G. E. P. (1954). Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification. The Annals of Mathematical Statistics, 25(2), 290–302.

Ruscio, J., & Roche, B. (2012). Variance Heterogeneity in Published Psychological Research: A Review and a New Index. Methodology, 8(1), 1–11. https://doi.org/10.1027/1614-2241/a000034

## Examples

```
d <- list(
  X1 = rnorm(10, sd = 10),
  X2 = rnorm(100, sd = 7.34),
  X3 = rnorm(1000, sd = 6.02),
  X4 = rnorm(100, sd = 5.17),
  X5 = rnorm(10, sd = 4.56)
)
cvv(d)
```

## cvv_manual *Coefficient of variance variation from manual input sample sizes and variance estimates*

### Description

Calculates three different indices for variation between two or more variance estimates. VR = Variance ratio between the largest and the smallest variance. CVV = Coefficient of variance variation (Box, 1954). SVH = Standardized variance heterogeneity (Ruscio & Roche, 2012).

### Usage

```
cvv_manual(sample_sizes, variances)
```

### Arguments

| | |
|---|---|
| sample_sizes | Numeric vector of length > 1. Sample sizes used for each variance estimate. |
| variances | Numeric vector of length > 1. Variance estimates. |

### Value

A vector including VR, CVV, and SVH.

### References

Box, G. E. P. (1954). Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification. The Annals of Mathematical Statistics, 25(2), 290–302.

Ruscio, J., & Roche, B. (2012). Variance Heterogeneity in Published Psychological Research: A Review and a New Index. Methodology, 8(1), 1–11. https://doi.org/10.1027/1614-2241/a000034

### Examples

```
cvv_manual(sample_sizes=c(10,100,1000,75,3),
variances=c(1.5,2,2.5,3,3.5))
```

## d_pooled_sd *Standardized mean difference with pooled standard deviation*

### Description

Standardized mean difference with pooled standard deviation

## Usage

```
d_pooled_sd(
  data,
  var,
  group.var,
  group.values,
  rename.output = TRUE,
  infer = FALSE
)
```

## Arguments

| | |
|---|---|
| data | A data frame. |
| var | A continuous variable for which difference is estimated. |
| group.var | The name of the group variable. |
| group.values | Vector of length 2, group values (e.g. c("male", "female) or c(0,1)). |
| rename.output | Logical. Should the output values be renamed according to the group.values? Default TRUE. |
| infer | Logical. Statistical inference with Welch test? (default FALSE) |

## Value

Descriptive statistics and mean differences

## Examples

```
d_pooled_sd(iris[iris$Species == "setosa" | iris$Species == "versicolor", ],
  var = "Petal.Length", group.var = "Species",
  group.values = c("setosa", "versicolor"), infer = TRUE
)
```

---

# D_regularized *Multivariate group difference estimation with regularized binomial regression*

---

## Description

Multivariate group difference estimation with regularized binomial regression

## Usage

```
D_regularized(
  data,
  mv.vars,
  group.var,
  group.values,
```

```
  alpha = 0.5,
  nfolds = 10,
  s = "lambda.min",
  type.measure = "deviance",
  rename.output = TRUE,
  out = FALSE,
  size = NULL,
  fold = FALSE,
  fold.var = NULL,
  pcc = FALSE,
  auc = FALSE,
  pred.prob = FALSE,
  prob.cutoffs = seq(0, 1, 0.2),
  append.data = FALSE
)
```

## Arguments

| | |
|---|---|
| data | A data frame. |
| mv.vars | Character vector. Variable names in the multivariate variable set. |
| group.var | The name of the group variable. |
| group.values | Vector of length 2, group values (e.g. c("male", "female) or c(0,1)). |
| alpha | Alpha-value for penalizing function ranging from 0 to 1: 0 = ridge regression, 1 = lasso, 0.5 = elastic net (default). |
| nfolds | Number of folds used for obtaining lambda (range from 3 to n-1, default 10). |
| s | Which lambda value is used for predicted values? Either "lambda.min" (default) or "lambda.1se". |
| type.measure | Which measure is used during cross-validation. Default "deviance". |
| rename.output | Logical. Should the output values be renamed according to the group.values? Default TRUE. |
| out | Logical. Should results and predictions be calculated on out-of-bad data set? (Default FALSE) |
| size | Integer. Number of cases in regularization data per each group. Default 1/4 of cases. |
| fold | Logical. Is regularization applied across sample folds with separate predictions for each fold? (Default FALSE) |
| fold.var | Character string. Name of the fold variable. (default NULL) |
| pcc | Logical. Include probabilities of correct classification? Default FALSE. |
| auc | Logical. Include area under the receiver operating characteristics? Default FALSE. |
| pred.prob | Logical. Include table of predicted probabilities? Default FALSE. |
| prob.cutoffs | Vector. Cutoffs for table of predicted probabilities. Default seq(0,1,0.20). |
| append.data | Logical. If TRUE, the data is appended to the predicted variables. |

## Value

| | |
|---|---|
| D | Multivariate descriptive statistics and differences. |
| pred.dat | A data.frame with predicted values. |
| cv.mod | Regularized regression model from cv.glmnet. |
| P.table | Table of predicted probabilities by cutoffs. |

## References

Lönnqvist, J. E., & Ilmarinen, V. J. (2021). Using a continuous measure of genderedness to assess sex differences in the attitudes of the political elite. Political Behavior, 43, 1779–1800. https://doi.org/10.1007/s11109-021-09681-2

Ilmarinen, V. J., Vainikainen, M. P., & Lönnqvist, J. E. (2022). Is there a g-factor of genderedness? Using a continuous measure of genderedness to assess sex differences in personality, values, cognitive ability, school grades, and educational track. European Journal of Personality. https://doi.org/10.1177/089020702210881

## See Also

cv.glmnet

## Examples

```
D_regularized(
  data = iris[iris$Species == "setosa" | iris$Species == "versicolor", ],
  mv.vars = c("Sepal.Length", "Sepal.Width", "Petal.Length", "Petal.Width"),
  group.var = "Species", group.values = c("setosa", "versicolor")
)$D

# out-of-bag predictions
D_regularized(
  data = iris[iris$Species == "setosa" | iris$Species == "versicolor", ],
  mv.vars = c("Sepal.Length", "Sepal.Width", "Petal.Length", "Petal.Width"),
  group.var = "Species", group.values = c("setosa", "versicolor"),
  out = TRUE, size = 15, pcc = TRUE, auc = TRUE
)$D

# separate sample folds
# generate data for 10 groups
set.seed(34246)
n1 <- 100
n2 <- 10
d <-
  data.frame(
    sex = sample(c("male", "female"), n1 * n2, replace = TRUE),
    fold = sample(x = LETTERS[1:n2], size = n1 * n2, replace = TRUE),
    x1 = rnorm(n1 * n2),
    x2 = rnorm(n1 * n2),
    x3 = rnorm(n1 * n2)
  )

# Fit and predict with same data
```

```
D_regularized(
  data = d,
  mv.vars = c("x1", "x2", "x3"),
  group.var = "sex",
  group.values = c("female", "male"),
  fold.var = "fold",
  fold = TRUE,
  rename.output = TRUE
)$D

# Out-of-bag data for each fold
D_regularized(
  data = d,
  mv.vars = c("x1", "x2", "x3"),
  group.var = "sex",
  group.values = c("female", "male"),
  fold.var = "fold",
  size = 17,
  out = TRUE,
  fold = TRUE,
  rename.output = TRUE
)$D
```

---

`D_regularized_fold` *Use manually defined data folds for regularization and obtain estimates for each separately.*

---

## Description

Use manually defined data folds for regularization and obtain estimates for each separately.

## Usage

```
D_regularized_fold(
  data,
  mv.vars,
  group.var,
  group.values,
  alpha = 0.5,
  s = "lambda.min",
  type.measure = "deviance",
  rename.output = TRUE,
  fold.var,
  append.data = FALSE
)
```

## Arguments

| | |
|---|---|
| data | A data frame. |
| mv.vars | Character vector. Variable names in the multivariate variable set. |
| group.var | The name of the group variable. |
| group.values | Vector of length 2, group values (e.g. c("male", "female) or c(0,1)). |
| alpha | Alpha-value for penalizing function ranging from 0 to 1: 0 = ridge regression, 1 = lasso, 0.5 = elastic net (default). |
| s | Which lambda value is used for predicted values? Either "lambda.min" (default) or "lambda.1se". |
| type.measure | Which measure is used during cross-validation. Default "deviance". |
| rename.output | Logical. Should the output values be renamed according to the group.values? Default TRUE. |
| fold.var | Character string. Name of the fold variable. |
| append.data | Logical. If TRUE, the original data is appended to the predicted variables. |

## Value

| | |
|---|---|
| D | Multivariate descriptive statistics and differences. |
| pred.dat | A data.frame with predicted values. |
| cv.mod | Regularized regression model from cv.glmnet. |

## See Also

cv.glmnet

## Examples

```
set.seed(34246)
n1 <- 100
n2 <- 10
d <-
  data.frame(
    sex = sample(c("male", "female"), n1 * n2, replace = TRUE),
    fold = sample(x = LETTERS[1:n2], size = n1 * n2, replace = TRUE),
    x1 = rnorm(n1 * n2),
    x2 = rnorm(n1 * n2),
    x3 = rnorm(n1 * n2)
  )

D_regularized_fold(
  data = d,
  mv.vars = c("x1", "x2", "x3"),
  group.var = "sex",
  group.values = c("female", "male"),
  fold.var = "fold"
)$D
```

`D_regularized_fold_out` *Use separate data partitions for regularization and estimation across defined data folds.*

## Description

Use separate data partitions for regularization and estimation across defined data folds.

## Usage

```
D_regularized_fold_out(
  data,
  mv.vars,
  group.var,
  group.values,
  alpha = 0.5,
  s = "lambda.min",
  type.measure = "deviance",
  rename.output = TRUE,
  size = NULL,
  fold.var,
  pcc = FALSE,
  auc = FALSE,
  pred.prob = FALSE,
  prob.cutoffs = seq(from = 0, to = 1, by = 0.2),
  append.data = FALSE
)
```

## Arguments

| | |
|---|---|
| `data` | A data frame. |
| `mv.vars` | Character vector. Variable names in the multivariate variable set. |
| `group.var` | The name of the group variable. |
| `group.values` | Vector of length 2, group values (e.g. c("male", "female) or c(0,1)). |
| `alpha` | Alpha-value for penalizing function ranging from 0 to 1: 0 = ridge regression, 1 = lasso, 0.5 = elastic net (default). |
| `s` | Which lambda value is used for predicted values? Either "lambda.min" (default) or "lambda.1se". |
| `type.measure` | Which measure is used during cross-validation. Default "deviance". |
| `rename.output` | Logical. Should the output values be renamed according to the group.values? Default TRUE. |
| `size` | Integer. Size of regularization data per each group. Default 1/4 of cases. |
| `fold.var` | Name of the fold variable. |

| | |
|---|---|
| pcc | Logical. Include probabilities of correct classification? Default FALSE. |
| auc | Logical. Include area under the receiver operating characteristics? Default FALSE. |
| pred.prob | Logical. Include table of predicted probabilities? Default FALSE. |
| prob.cutoffs | Vector. Cutoffs for table of predicted probabilities. Default seq(0,1,0.20). |
| append.data | Logical. If TRUE, the testing data split is appended to the predicted variables. |

## Value

| | |
|---|---|
| D | Multivariate descriptive statistics and differences. |
| pred.dat | A data.frame with predicted values. |
| cv.mod | Regularized regression model from cv.glmnet. |
| P.table | Table of predicted probabilities by cutoffs. |

## Examples

```
set.seed(34246)
n1 <- 100
n2 <- 10
d <-
  data.frame(
    sex = sample(c("male", "female"), n1 * n2, replace = TRUE),
    fold = sample(x = LETTERS[1:n2], size = n1 * n2, replace = TRUE),
    x1 = rnorm(n1 * n2),
    x2 = rnorm(n1 * n2),
    x3 = rnorm(n1 * n2)
  )
D_regularized_fold_out(
  data = d,
  mv.vars = c("x1", "x2", "x3"),
  group.var = "sex",
  group.values = c("female", "male"),
  fold.var = "fold",
  size = 17,
  pcc = TRUE
)$D
```

---

# D_regularized_out *Use separate data partition for regularization and estimation.*

---

## Description

Use separate data partition for regularization and estimation.

## Usage

```
D_regularized_out(
  data,
  mv.vars,
  group.var,
  group.values,
  alpha = 0.5,
  nfolds = 10,
  s = "lambda.min",
  type.measure = "deviance",
  rename.output = TRUE,
  size = NULL,
  pcc = FALSE,
  auc = FALSE,
  pred.prob = FALSE,
  prob.cutoffs = seq(from = 0, to = 1, by = 0.2),
  append.data = FALSE
)
```

## Arguments

| | |
|---|---|
| data | A data frame. |
| mv.vars | Character vector. Variable names in the multivariate variable set. |
| group.var | The name of the group variable. |
| group.values | Vector of length 2, group values (e.g. c("male", "female) or c(0,1)). |
| alpha | Alpha-value for penalizing function ranging from 0 to 1: 0 = ridge regression, 1 = lasso, 0.5 = elastic net (default). |
| nfolds | Number of folds used for obtaining lambda (range from 3 to n-1, default 10). |
| s | Which lambda value is used for predicted values? Either "lambda.min" (default) or "lambda.1se". |
| type.measure | Which measure is used during cross-validation. Default "deviance". |
| rename.output | Logical. Should the output values be renamed according to the group.values? Default TRUE. |
| size | Integer. Size of regularization data per each group. Default 1/4 of cases. |
| pcc | Logical. Include probabilities of correct classification? Default FALSE. |
| auc | Logical. Include area under the receiver operating characteristics? Default FALSE. |
| pred.prob | Logical. Include table of predicted probabilities? Default FALSE. |
| prob.cutoffs | Vector. Cutoffs for table of predicted probabilities. Default seq(0,1,0.20). |
| append.data | Logical. If TRUE, the testing data split is appended to the predicted variables. |

## Value

| | |
|---|---|
| D | Multivariate descriptive statistics and differences. |
| pred.dat | A data.frame with predicted values. |
| cv.mod | Regularized regression model from cv.glmnet. |
| P.table | Table of predicted probabilities by cutoffs. |

## Examples

```
D_regularized_out(
  data = iris[iris$Species == "setosa" |
    iris$Species == "versicolor", ],
  mv.vars = c(
    "Sepal.Length", "Sepal.Width",
    "Petal.Length", "Petal.Width"
  ),
  group.var = "Species",
  group.values = c("setosa", "versicolor"),
  size = 40,
  pcc = TRUE
)$D
```

---

# D_regularized_vanilla *Use same data partition for regularization and estimation.*

---

## Description

Use same data partition for regularization and estimation.

## Usage

```
D_regularized_vanilla(
  data,
  mv.vars,
  group.var,
  group.values,
  alpha = 0.5,
  nfolds = 10,
  s = "lambda.min",
  type.measure = "deviance",
  rename.output = TRUE,
  append.data = FALSE
)
```

## Arguments

| | |
|---|---|
| data | A data frame. |
| mv.vars | Character vector. Variable names in the multivariate variable set. |
| group.var | The name of the group variable. |
| group.values | Vector of length 2, group values (e.g. c("male", "female) or c(0,1)). |
| alpha | Alpha-value for penalizing function ranging from 0 to 1: 0 = ridge regression, 1 = lasso, 0.5 = elastic net (default). |
| nfolds | Number of folds used for obtaining lambda (range from 3 to n-1, default 10). |
| s | Which lambda value is used for predicted values? Either "lambda.min" (default) or "lambda.1se". |
| type.measure | Which measure is used during cross-validation. Default "deviance". |
| rename.output | Logical. Should the output values be renamed according to the group.values? Default TRUE. |
| append.data | Logical. If TRUE, the original data is appended to the predicted variables. |

## Value

| | |
|---|---|
| D | Multivariate descriptive statistics and differences. |
| pred.dat | A data.frame with predicted values. |
| cv.mod | Regularized regression model from cv.glmnet. |

## See Also

cv.glmnet

## Examples

```
D_regularized(
  data = iris[iris$Species == "setosa" | iris$Species == "versicolor", ],
  mv.vars = c("Sepal.Length", "Sepal.Width", "Petal.Length", "Petal.Width"),
  group.var = "Species", group.values = c("setosa", "versicolor")
)$D
```

---

# ml_dadas *Predicting algebraic difference scores in multilevel model*

---

## Description

Decomposes difference score predictions to predictions of difference score components by probing simple effects at the levels of the binary moderator.

## Usage

```
ml_dadas(
  model,
  predictor,
  diff_var,
  diff_var_values,
  scaled_estimates = FALSE,
  re_cov_test = FALSE,
  var_boot_test = FALSE,
  nsim = NULL,
  level = 0.95,
  seed = NULL,
  abs_diff_test = 0
)
```

## Arguments

| | |
|---|---|
| model | Multilevel model fitted with lmerTest. |
| predictor | Character string. Variable name of independent variable predicting difference score. |
| diff_var | Character string. A variable indicative of difference score components (two groups). |
| diff_var_values | Vector. Values of the component score groups in diff_var. |
| scaled_estimates | Logical. Are scaled estimates obtained? Does fit a reduced model for correct standard deviations. (Default FALSE) |
| re_cov_test | Logical. Significance test for random effect covariation? Does fit a reduced model without the correlation. (Default FALSE) |
| var_boot_test | Logical. Compare variance by lower-level groups at the upper-level in a reduced model with bootstrap? (Default FALSE) |
| nsim | Numeric. Number of bootstrap simulations. |
| level | Numeric. The confidence level required for the var_boot_test output (Default .95) |
| seed | Numeric. Seed number for bootstrap simulations. |
| abs_diff_test | Numeric. A value against which absolute difference between component score predictions is tested (Default 0). |

## Value

| | |
|---|---|
| dadas | A data frame including main effect, interaction, regression coefficients for component scores, dadas, and comparison between interaction and main effect. |
| scaled_estimates | Scaled regression coefficients for difference score components and difference score. |

| | |
|---|---|
| vpc_at_reduced | Variance partition coefficients in the model without the predictor and interactions. |
| re_cov_test | Likelihood ratio significance test for random effect covariation. |
| boot_var_diffs | List of different variance bootstrap tests. |

## Examples

```
## Not run:
set.seed(95332)
n1 <- 10 # groups
n2 <- 10 # observations per group

dat <- data.frame(
  group = rep(c(LETTERS[1:n1]), each = n2),
  w = sample(c(-0.5, 0.5), n1 * n2, replace = TRUE),
  x = rep(sample(1:5, n1, replace = TRUE), each = n2),
  y = sample(1:5, n1 * n2, replace = TRUE)
)
library(lmerTest)
fit <- lmerTest::lmer(y ~ x * w + (w | group),
  data = dat
)

round(ml_dadas(fit,
  predictor = "x",
  diff_var = "w",
  diff_var_values = c(0.5, -0.5)
)$dadas, 3)

## End(Not run)
```

---

# pcc

*Returns probabilities of correct classification for both groups in independent data partition.*

---

## Description

Returns probabilities of correct classification for both groups in independent data partition.

## Usage

```
pcc(data, pred.var, group.var, group.values)
```

## Arguments

| | |
|---|---|
| data | Data frame including predicted values (e.g., pred.dat from D_regularized_out). |
| pred.var | Character string. Variable name for predicted values. |
| group.var | The name of the group variable. |
| group.values | Vector of length 2, group values (e.g. c("male", "female) or c(0,1)). |

## Value

Vector of length 2. Probabilities of correct classification.

## Examples

```
D_out <- D_regularized_out(
  data = iris[iris$Species == "versicolor" | iris$Species == "virginica", ],
  mv.vars = c("Sepal.Length", "Sepal.Width", "Petal.Length", "Petal.Width"),
  group.var = "Species", group.values = c("versicolor", "virginica"),
  size = 15
)

pcc(
  data = D_out$pred.dat,
  pred.var = "pred",
  group.var = "group",
  group.values = c("versicolor", "virginica")
)
```

---

# qcc *Quantile correlation coefficient*

---

## Description

For computation of tail dependence as correlations estimated at different variable quantiles (Choi & Shin, 2022; Lee et al., 2022) summarized across two quantile regression models where x and y switch roles as independent/dependent variables.

## Usage

```
qcc(
  x,
  y,
  tau = c(0.1, 0.5, 0.9),
  data,
  method = "br",
  boot_n = NULL,
  ci_level = 0.95
)
```

## Arguments

| | |
|---|---|
| x | Name of x variable. Character string. |
| y | Name of y variable. Character string. |
| tau | The quantile(s) to be estimated. A vector of values between 0 and 1, default c(.1,.5,.9). @seealso rq |
| data | Data frame. |

| | |
|---|---|
| method | The algorithmic method used to compute the fit (default "br"). @seealso rq |
| boot_n | Number of bootstrap redraws (default NULL = no bootstrap inference). |
| ci_level | Level for percentile bootstrap confidence interval. Numeric values between 0 and 1. Default .95. |

## Value

| | |
|---|---|
| r | Pearson's correlation estimate for comparison. |
| rho_tau | Correlations at different tau values (quantiles). |
| r_boot_est | Pearson's correlation bootstrap estimates. |
| rho_tau_boot_est | Bootstrap estimates for correlations at different tau values (quantiles). |

## References

Choi, J.-E., & Shin, D. W. (2022). Quantile correlation coefficient: A new tail dependence measure. Statistical Papers, 63(4), 1075–1104. https://doi.org/10.1007/s00362-021-01268-7

Lee, J. A., Bardi, A., Gerrans, P., Sneddon, J., van Herk, H., Evers, U., & Schwartz, S. (2022). Are value–behavior relations stronger than previously thought? It depends on value importance. European Journal of Personality, 36(2), 133–148. https://doi.org/10.1177/08902070211002965

## Examples

```
set.seed(2321)
d <- data.frame(x = rnorm(2000))
d$y <- 0.10 * d$x + (0.20) * d$x^2 + 0.40 * d$x^3 + (-0.20) * d$x^4 + rnorm(2000)
qcc_boot <- qcc(x = "x", y = "y", data = d, tau = 1:9 / 10, boot_n = 50)
qcc_boot$rho_tau
```

---

# reliability_dms

*Reliability calculation for difference score variable that is a difference between two mean variables calculated over upper-level units (e.g., sex differences across countries)*

---

## Description

Calculates reliability of difference score (Johns, 1981) based on two separate ICC2 values (Bliese, 2000), standard deviations of mean values over upper-level units, and correlations between the mean values across upper-level units.

## Usage

```
reliability_dms(
  model = NULL,
  data = NULL,
  diff_var,
  diff_var_values,
  var,
  group_var
)
```

## Arguments

| | |
|---|---|
| model | Multilevel model fitted with lmer (default NULL) |
| data | Long format data frame (default NULL) |
| diff_var | Character string. A variable indicative of difference score components (two groups). |
| diff_var_values | Vector. Values of the component score groups in diff_var. |
| var | Character string. Name of the dependent variable or variable of which mean values are calculated. |
| group_var | Character string. Upper-level clustering unit. |

## Value

A vector including ICC2s (r11 and r22), SDs (sd1, sd2, and sd_d12), means (m1, m2, and m_d12), correlation between means (r12), and reliability of the mean difference variable.

## References

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions (pp. 349–381). Jossey-Bass.

Johns, G. (1981). Difference score measures of organizational behavior variables: A critique. Organizational Behavior and Human Performance, 27(3), 443–463. https://doi.org/10.1016/0030-5073(81)90033-7

## Examples

```
set.seed(4317)
n2 <- 20
n1 <- 200
ri <- rnorm(n2, m = 0.5, sd = 0.2)
rs <- 0.5 * ri + rnorm(n2, m = 0.3, sd = 0.15)
d.list <- list()
for (i in 1:n2) {
  x <- rep(c(-0.5, 0.5), each = n1 / 2)
  y <- ri[i] + rs[i] * x + rnorm(n1)
```

```
  d.list[[i]] <- cbind(x, y, i)
}

d <- data.frame(do.call(rbind, d.list))
names(d) <- c("x", "y", "cntry")
reliability_dms(
  data = d, diff_var = "x",
  diff_var_values = c(-0.5, 0.5), var = "y", group_var = "cntry"
)
```

---

sem_dadas *Predicting algebraic difference scores in structural equation model*

---

## Description

Predicting algebraic difference scores in structural equation model

## Usage

```
sem_dadas(
  data,
  var1,
  var2,
  center = FALSE,
  scale = FALSE,
  predictor,
  covariates = NULL,
  estimator = "MLR",
  level = 0.95,
  sampling.weights = NULL,
  abs_coef_diff_test = 0
)
```

## Arguments

| | |
|---|---|
| data | A data frame. |
| var1 | Character string. Variable name of first component score of difference score (Y_1). |
| var2 | Character string. Variable name of second component score of difference score (Y_2). |
| center | Logical. Are var1 and var2 centered around their grand mean? (Default FALSE) |
| scale | Logical. Are var1 and var2 scaled with their pooled sd? (Default FALSE) |
| predictor | Character string. Variable name of independent variable predicting difference score. |
| covariates | Character string or vector. Variable names of covariates (Default NULL). |
| estimator | Character string. Estimator used in SEM (Default "MLR"). |

`level` Numeric. The confidence level required for the result output (Default .95)

`sampling.weights`
Character string. Name of sampling weights variable.

`abs_coef_diff_test`
Numeric. A value against which absolute difference between component score predictions is tested (Default 0).

## Value

`descriptives` Means, standard deviations, and intercorrelations.

`parameter_estimates`
Parameter estimates from the structural equation model.

`variance_test` Variances and covariances of component scores.

`transformed_data`
Data frame with variables used in SEM.

`dadas` One sided dadas-test for positivity of abs(b_11-b_21)-abs(b_11+b_21).

`results` Summary of key results.

## References

Edwards, J. R. (1995). Alternatives to Difference Scores as Dependent Variables in the Study of Congruence in Organizational Research. Organizational Behavior and Human Decision Processes, 64(3), 307–324.

## Examples

```
## Not run:
set.seed(342356)
d <- data.frame(
  var1 = rnorm(50),
  var2 = rnorm(50),
  x = rnorm(50)
)
sem_dadas(
  data = d, var1 = "var1", var2 = "var2",
  predictor = "x", center = TRUE, scale = TRUE,
  abs_coef_diff_test = 0.20
)$results

## End(Not run)
```

vpc_at *Variance partition coefficient calculated at different level-1 values*

## Description

Calculates variance estimates (level-2 Intercept variance) and variance partition coefficients (i.e., intra-class correlation) at selected values of predictor values in two-level linear models with random effects (intercept, slope, and their covariation).

## Usage

```
vpc_at(model, lvl1.var, lvl1.values)
```

## Arguments

| | |
|---|---|
| model | Two-level model fitted with lme4. Must include random intercept, slope, and their covariation. |
| lvl1.var | Character string. Level 1 variable name to which random slope is also estimated. |
| lvl1.values | Level 1 variable values. |

## Value

Data frame of level 2 variance and std.dev. estimates at level 1 variable values, respective VPCs (ICC1s) and group-mean reliabilities (ICC2s) (Bliese, 2000).

## References

Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning Variation in Multilevel Models. Understanding Statistics, 1(4), 223–231. https://doi.org/10.1207/S15328031US0104_02

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions (pp. 349–381). Jossey-Bass.

## Examples

```
fit <- lme4::lmer(Sepal.Length ~ Petal.Length +
  (Petal.Length | Species),
data = iris
)

lvl1.values <-
  c(
    mean(iris$Petal.Length) - stats::sd(iris$Petal.Length),
    mean(iris$Petal.Length),
    mean(iris$Petal.Length) + stats::sd(iris$Petal.Length)
  )
```

```
vpc_at(
  model = fit,
  lvl1.var = "Petal.Length",
  lvl1.values = lvl1.values
)
```

# Index