Package ‘diffcor’

May 3, 2024

Type Package

Title Fisher's z-Tests Concerning Differences Between Correlations

Version 0.8.3

Depends R (>= 4.3.0), MASS

Date 2024-05-02

Author Christian Blötner

Maintainer Christian Blötner <c.bloetner@gmail.com>

Description Computations of Fisher's z-tests concerning different kinds of correlation differences. Additionally, approaches to estimating statistical power via Monte Carlo simulations are implemented. Important to note, the Pearson correlation coefficient is sensitive to linear association, but also to a host of statistical issues such as univariate and bivariate outliers, range restrictions, and heteroscedasticity (e.g., Duncan & Layard, 1973 <doi:10.1093/BIOMET/60.3.551>; Wilcox, 2013 <doi:10.1016/C2010-0-67044-1>). Thus, every power analysis requires that specific statistical prerequisites are fulfilled and can be invalid if the prerequisites do not hold.

License GPL (>= 2)

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2024-05-03 12:30:13 UTC

R topics documented:

diffcor.dep 2
diffcor.one 3
diffcor.two 4
diffpwr.dep 6
diffpwr.one 8
diffpwr.two 10
visual_mc 12

Index **14**

diffcor.dep — *Fisher's z-Tests of dependent correlations*

Description

Tests if the correlation between two variables (r12) differs from the correlation between the first and a third one (r13), given the intercorrelation of the compared constructs (r23). All correlations are automatically transformed with the Fisher z-transformation prior to computations. The output provides the compared correlations, test statistic as z-score, and p-values.

Usage

```
diffcor.dep(r12, r13, r23, n, cor.names = NULL,
alternative = c("one.sided", "two.sided"), digit = 3)
```

Arguments

`r12`	Correlation between the criterion with which both competing variables are correlated and the first of the two competing variables.
`r13`	Correlation between the criterion with which both competing variables are correlated and the second of the two competing variables.
`r23`	Intercorrelation between the two competing variables.
`n`	Sample size in which the observed effect was found
`cor.names`	OPTIONAL, label for the correlation. DEFAULT is NULL
`alternative`	A character string specifying if you wish to test one-sided or two-sided differences
`digit`	Number of digits in the output for all parameters, DEFAULT = 3

Value

`r12`	Correlation between the criterion with which both competing variables are correlated and the first of the two competing variables.
`r13`	Correlation between the criterion with which both competing variables are correlated and the second of the two competing variables.
`r23`	Intercorrelation between the two competing variables.
`z`	Test statistic for correlation difference in units of z distribution
`p`	p value for one- or two-sided testing, depending on alternative = c("one.sided", "two.sided)

Author(s)

Christian Blötner <c.bloetner@gmail.com>

References

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.

Eid, M., Gollwitzer, M., & Schmitt, M. (2015). Statistik und Forschungsmethoden (4.Auflage) [Statistics and research methods (4th ed.)]. Beltz.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87, 245-251.

Examples

```
diffcor.dep(r12 = .76, r13 = .70, r23 = .50, n = 271, digit = 4,
cor.names = NULL, alternative = "two.sided")
```

---

diffcor.one — *Fisher's z-test of difference between an empirical and a hypothesized correlation*

Description

The function tests whether an observed correlation differs from an expected one, for example, in construct validation. All correlations are automatically transformed with the Fisher z-transformation prior to computations. The output provides the compared correlations, a z-score, a p-value, a confidence interval, and the effect size Cohens q. According to Cohen (1988), q = |.10|, |.30| and |.50| are considered small, moderate, and large differences, respectively.

Usage

```
diffcor.one(emp.r, hypo.r, n, alpha = .05, cor.names = NULL,
alternative = c("one.sided", "two.sided"), digit = 3)
```

Arguments

`emp.r`	Empirically observed correlation
`hypo.r`	Hypothesized correlation which shall be tested
`n`	Sample size in which the observed effect was found
`alpha`	Likelihood of Type I error, DEFAULT = .05
`cor.names`	OPTIONAL, label for the correlation (e.g., "IQ-performance"). DEFAULT is NULL
`digit`	Number of digits in the output for all parameters, DEFAULT = 3
`alternative`	A character string specifying if you wish to test one-sided or two-sided differences

Value

r_exp	Vector of the expected correlations
r_obs	Vector of the empirically observed correlations
LL	Lower limit of the confidence interval of the empirical correlation, given the specified alpha level, DEFAULT = 95 percent
UL	Upper limit of the confidence interval of the empirical correlation, given the specified alpha level, DEFAULT = 95 percent
z	Test statistic for correlation difference in units of z distribution
p	p value for one- or two-sided testing, depending on alternative = c("one.sided", "two.sided)
Cohen_q	Effect size measure for differences of independent correlations

Author(s)

Christian Blötner <c.bloetner@gmail.com>

References

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.

Eid, M., Gollwitzer, M., & Schmitt, M. (2015). Statistik und Forschungsmethoden (4.Auflage) [Statistics and research methods (4th ed.)]. Beltz.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87, 245-251.

Examples

```
diffcor.one(c(.76, .53, -.32), c(.70, .35, -.40),
  c(225, 250, 210),
  cor.names = c("a-b", "c-d", "e-f"), digit = 2, alternative = "one.sided")
```

---

diffcor.two — *Fisher's z-Tests for differences of correlations in two independent samples*

Description

Tests whether the correlation between two variables differs across two independent studies/samples. The correlations are automatically transformed with the Fisher z-transformation prior to computations. The output provides the compared correlations, test statistic as z-score, p-values, confidence intervals of the empirical correlations, and the effect size Cohens q. According to Cohen (1988), q = |.10|, |.30| and |.50| are considered small, moderate, and large differences, respectively.

Usage

```
diffcor.two(r1, r2, n1, n2, alpha = .05, cor.names = NULL,
alternative = c("one.sided", "two.sided"), digit = 3)
```

Arguments

`r1`	Correlation coefficient in first sample
`r2`	Correlation coefficient in second sample
`n1`	First sample size
`n2`	Second sample size
`alpha`	Likelihood of Type I error, DEFAULT = .05
`cor.names`	OPTIONAL, label for the correlation (e.g., "IQ-performance"). DEFAULT is NULL
`digit`	Number of digits in the output for all parameters, DEFAULT = 3
`alternative`	A character string specifying if you wish to test one-sided or two-sided differences

Value

`r1`	Vector of the empirically observed correlations in the first sample
`r2`	Vector of the empirically observed correlations in the second sample
`LL1`	Lower limit of the confidence interval of the first empirical correlation, given the specified alpha level, DEFAULT = 95 percent
`UL1`	Upper limit of the confidence interval of the first empirical correlation, given the specified alpha level, DEFAULT = 95 percent
`LL2`	Lower limit of the confidence interval of the second empirical correlation, given the specified alpha level, DEFAULT = 95 percent
`UL2`	Upper limit of the confidence interval of the second empirical correlation, given the specified alpha level, DEFAULT = 95 percent
`z`	Test statistic for correlation difference in units of z distribution
`p`	p value for one- or two-sided testing, depending on alternative = c("one.sided", "two.sided)
`Cohen_q`	Effect size measure for differences of independent correlations

Author(s)

Christian Blötner <c.bloetner@gmail.com>

References

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.

Eid, M., Gollwitzer, M., & Schmitt, M. (2015). Statistik und Forschungsmethoden (4.Auflage) [Statistics and research methods (4th ed.)]. Beltz.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87, 245-251.

Examples

```
diffcor.two(r1 = c(.39, .52, .22),
  r2 = c(.29, .44, .12),
  n1 = c(66, 66, 66), n2 = c(96, 96, 96), alpha = .01,
  cor.names = c("a-b", "c-d", "e-f"), alternative = "one.sided")
```

---

diffpwr.dep — *Monte Carlo Simulation for the correlation difference between dependent correlations*

Description

Computation of a Monte Carlo simulation to estimate the statistical power of the comparison between the correlations of a variable with two competing variables that are also correlated with each other.

Usage

```
diffpwr.dep(n,
            rho12,
            rho13,
            rho23,
            alpha = 0.05,
            n.samples = 1000,
            seed = 1234)
```

Arguments

`n`	Sample size to be tested in the Monte Carlo simulation.
`rho12`	Assumed population correlation between the criterion with which both competing variables are correlated and the first of the two competing variables.
`rho13`	Assumed population correlation between the criterion with which both competing variables are correlated and the second of the two competing variables.
`rho23`	Assumed population correlation between the two competing variables.
`alpha`	Type I error. Default is .05.
`n.samples`	Number of samples generated in the Monte Carlo simulation. The recommended minimum is 1,000 iterations, which is also the default.
`seed`	To make the results reproducible, it is recommended to set a random seed.

Details

Depending on the number of generated samples (n.samples), correlation coefficients simulated. For each simulated sample, it is checked whether the correlations r12 and r13 differ, given the correlation r23. The ratio of simulated z-tests of the correlation difference tests exceeding the critical z-value, given the intended alpha-level and sample size, equals the achieved statistical power(see Muthén & Muthén, 2002 <doi:10.1207/S15328007SEM0904_8>; Robert & Casella, 2010 <doi:10.1007/978-1-4419-1576-4>, for overviews of the Monte Carlo method).

It should be noted that the Pearson correlation coefficient is sensitive to linear association, but also to a host of statistical issues such as univariate and bivariate outliers, range restrictions, and heteroscedasticity (e.g., Duncan & Layard, 1973 <doi:10.1093/BIOMET/60.3.551>; Wilcox, 2013 <doi:10.1016/C2010-0-67044-1>). Thus, every power analysis requires that specific statistical prerequisites are fulfilled and can be invalid with regard to the actual data if the prerequisites do not hold, potentially biasing Type I error rates.

Value

As dataframe with the following parameters

`rho12`	Assumed population correlation between the criterion with which both competing variables are correlated and the first of the two competing variables.
`cov12`	Coverage. Indicates the ratio of simulated confidence intervals including the assumed effect size rho12.
`bias12_M`	Difference between the mean of the distribution of the simulated correlations and rho12, divided by rho12.
`bias12_Md`	Difference between the median of the distribution of the simulated correlations and rho12, divided by rho12.
`rho13`	Assumed population correlation between the criterion with which both competing variables are correlated and the second of the two competing variables.
`cov13`	Coverage. Indicates the ratio of simulated confidence intervals including the assumed effect size rho13.
`bias13_M`	Difference between the mean of the distribution of the simulated correlations and rho13, divided by rho13.
`bias13_Md`	Difference between the median of the distribution of the simulated correlations and rho13, divided by rho13.
`rho23`	Assumed population correlation between the two competing variables.
`cov23`	Coverage. Indicates the ratio of simulated confidence intervals including the assumed effect size rho23.
`bias23_M`	Difference between the mean of the distribution of the simulated correlations and rho23, divided by rho23.
`bias23_Md`	Difference between the median of the distribution of the simulated correlations and rho23, divided by rho23.
`n`	Sample size to be tested in the Monte Carlo simulation.
`pwr`	Statistical power as the ratio of simulated difference tests that yielded statistical significance.

Biases should be as close to zero as possible and coverage should be ideally between .91 and .98 (Muthén & Muthén, 2002 <doi:10.1207/S15328007SEM0904_8>).

Author(s)

Christian Blötner `<c.bloetner@gmail.com>`

References

Duncan, G. T., & Layard, M. W. (1973). A Monte-Carlo study of asymptotically robust tests for correlation coefficients. Biometrika, 60, 551–558. https://doi.org/10.1093/BIOMET/60.3.551

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. Structural Equation Modeling: A Multidisciplinary Journal, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

Robert, C., & Casella, G. (2010). Introducing Monte Carlo methods with R. Springer. https://doi.org/10.1007/978-1-4419-1576-4

Wilcox, R. (2013). Introduction to robust estimation and hypothesis testing. Elsevier. https://doi.org/10.1016/C2010-0-67044-1

Examples

```
diffpwr.dep(n.samples = 1000,
                    n = 250,
                    rho12 = .30,
                    rho13 = .45,
                    rho23 = .50,
                    alpha = .05,
                    seed = 1234)
```

`diffpwr.one` *Monte Carlo Simulation for the correlation difference between an expected and an observed correlation*

Description

Computation of a Monte Carlo simulation to estimate the statistical power the correlation difference between an observed correlation coefficient and an a fixed value against which the correlation should be tested.

Usage

```
diffpwr.one(n,
            r,
            rho,
            alpha = .05,
            n.samples = 1000,
            seed = 1234)
```

Arguments

`n`	Sample size to be tested in the Monte Carlo simulation.
`r`	Assumed observed correlation.
`rho`	Correlation coefficient against which to test (reflects the null hypothesis).

alpha	Type I error. Default is .05.
n.samples	Number of samples generated in the Monte Carlo simulation. The recommended minimum is 1,000 iterations, which is also the default.
seed	To make the results reproducible, it is recommended to set a random seed.

Details

Depending on the number of generated samples (n.samples), correlation coefficients of size r are simulated. Confidence intervals are constructed around the simulated correlation coefficients. For each simulated coefficient, it is then checked whether the hypothesized correlation cofficient (rho) falls within this interval. All correlations are automatically transformed with the Fisher z-transformation prior to computations. The ratio of simulated confidence intervals excluding the hypothesized coefficient equals the statistical power, given the intended alpha-level and sample size (see Robert & Casella, 2010 <doi:10.1007/978-1-4419-1576-4>, for an overview of the Monte Carlo method).

It should be noted that the Pearson correlation coefficient is sensitive to linear association, but also to a host of statistical issues such as univariate and bivariate outliers, range restrictions, and heteroscedasticity (e.g., Duncan & Layard, 1973 <doi:10.1093/BIOMET/60.3.551>; Wilcox, 2013 <doi:10.1016/C2010-0-67044-1>). Thus, every power analysis requires that specific statistical prerequisites are fulfilled and can be invalid with regard to the actual data if the prerequisites do not hold, potentially biasing Type I error rates.

Value

As dataframe with the following parameters

r	Empirically observed correlation.
rho	Correlation against which r should be tested.
n	The sample size entered in the function.
cov	Coverage. Indicates the ratio of simulated confidence intervals including the assumed correlation r. Should be between .91 and .98 (Muthén & Muthén, 2002 <doi:10.1207/S15328007SEM0904_8>).
bias_M	Difference between the mean of the distribution of the simulated correlations and rho, divided by rho.
bias_Md	Difference between the median of the distribution of the simulated correlations and rho, divided by rho.
pwr	Statistical power as the ratio of simulated confidence intervals excluding rho.

Author(s)

Christian Blötner <c.bloetner@gmail.com>

References

Duncan, G. T., & Layard, M. W. (1973). A Monte-Carlo study of asymptotically robust tests for correlation coefficients. Biometrika, 60, 551–558. https://doi.org/10.1093/BIOMET/60.3.551

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. Structural Equation Modeling: A Multidisciplinary Journal, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

Robert, C., & Casella, G. (2010). Introducing Monte Carlo methods with R. Springer. https://doi.org/10.1007/978-1-4419-1576-4

Wilcox, R. (2013). Introduction to robust estimation and hypothesis testing. Elsevier. https://doi.org/10.1016/C2010-0-67044-1

Examples

```
diffpwr.one(n = 500,
            r = .30,
            rho = .40,
            alpha = .05,
            n.samples = 1000,
            seed = 1234)
```

`diffpwr.two` *Monte Carlo Simulation for the correlation difference between two correlations that were observed in two independent samples*

Description

Computation of a Monte Carlo simulation to estimate the statistical power the correlation difference between the correlation coefficients detected in two indepdenent samples (e.g., original study and replication study).

Usage

```
diffpwr.two(n1,
            n2,
            rho1,
            rho2,
            alpha = .05,
            n.samples = 1000,
            seed = 1234)
```

Arguments

n1	Sample size to be tested in the Monte Carlo simulation for the first sample.
n2	Sample size to be tested in the Monte Carlo simulation for the second sample.
rho1	Assumed population correlation to be observed in the first sample.
rho2	Assumed population correlation to be observed in the second sample.
alpha	Type I error. Default is .05.
n.samples	Number of samples generated in the Monte Carlo simulation. The recommended minimum is 1,000 iterations, which is also the default.
seed	To make the results reproducible, a random seed is specified.

Details

Depending on the number of generated samples (n.samples), correlation coefficients are simulated. For each simulated pair of coefficients, it is then checked whether the confidence intervals (with given alpha level) of the correlations overlap. All correlations are automatically transformed with the Fisher z-transformation prior to computations. The ratio of simulated non- overlapping confidence intervals equals the statistical power, given the alpha-level and sample sizes (see Robert & Casella, 2010 <doi:10.1007/978-1-4419-1576-4>, for an overview of the Monte Carlo method).

It should be noted that the Pearson correlation coefficient is sensitive to linear association, but also to a host of statistical issues such as univariate and bivariate outliers, range restrictions, and heteroscedasticity (e.g., Duncan & Layard, 1973 <doi:10.1093/BIOMET/60.3.551>; Wilcox, 2013 <doi:10.1016/C2010-0-67044-1>). Thus, every power analysis requires that specific statistical prerequisites are fulfilled and can be invalid with regard to the actual data if the prerequisites do not hold, potentially biasing Type I error rates.

Value

As dataframe with the following parameters

`rho1`	Assumed population correlation to be observed in the first sample.
`n1`	Sample size of the first sample.
`cov1`	Coverage. Ratio of simulated confidence intervals including rho1.
`bias1_M`	Difference between the mean of the distribution of the simulated correlations and rho1, divided by rho1.
`bias1_Md`	Difference between the median of the distribution of the simulated correlations and rho1, divided by rho1.
`rho2`	Assumed population correlation to be observed in the second sample.
`n2`	The sample size of the second sample.
`cov2`	Coverage. Ratio of simulated confidence intervals including rho2.
`bias2_M`	Difference between the mean of the distribution of the simulated correlations and rho2, divided by rho2.
`bias2_Md`	Difference between the median of the distribution of the simulated correlations and rho2, divided by rho2.
`pwr`	Statistical power as the ratio of simulated non-verlapping confidence intervals.

Biases should be as close to zero as possible and coverage should be ideally between .91 and .98 (Muthén & Muthén, 2002 <doi:10.1207/S15328007SEM0904_8>).

Author(s)

Christian Blötner <c.bloetner@gmail.com>

References

Duncan, G. T., & Layard, M. W. (1973). A Monte-Carlo study of asymptotically robust tests for correlation coefficients. Biometrika, 60, 551–558. https://doi.org/10.1093/BIOMET/60.3.551

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. Structural Equation Modeling: A Multidisciplinary Journal, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

Robert, C., & Casella, G. (2010). Introducing Monte Carlo methods with R. Springer. https://doi.org/10.1007/978-1-4419-1576-4

Wilcox, R. (2013). Introduction to robust estimation and hypothesis testing. Elsevier. https://doi.org/10.1016/C2010-0-67044-1

Examples

```
diffpwr.two(n1 = 1000,
            n2 = 594,
            rho1 = .45,
            rho2 = .39,
            alpha = .05,
            n.samples = 1000,
            seed = 1234)
```

`visual_mc` *Visualization of the simulated parameters*

Description

To evaluate the quality of the Monte Carlo simulation beyond bias and coverage parameters (Muthén & Muthén, 2002), it can be helpful to also inspect the simulated parameters visually. To this end, visual_mc() can be used to visualize the simulated parameters (including corresponding confidence intervals) in relation to the targeted parameter.

Usage

```
visual_mc(rho,
          n,
          alpha = .05,
          n.intervals = 100,
          seed = 1234)
```

Arguments

`rho`	Targeted correlation coefficient of the simulation.
`n`	An integer reflecting the sample size.
`alpha`	Type I error. Default is .05.
`n.intervals`	An integer reflecting the number of simulated parameters that should be visualized in the graphic. Default is 100.
`seed`	To make the results reproducible, a random seed is specified.

Value

A plot in which the targeted correlation coefficient is visualized with a dashed red line and the simulated correlation coefficients are visualized by black squares and confidence intervals (level depending on the specification made in the argument alpha).

Author(s)

Christian Blötner <c.bloetner@gmail.com>

References

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. Structural Equation Modeling: A Multidisciplinary Journal, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

Examples

```
visual_mc(rho = .25,
                    n = 300,
                    alpha = .05,
                    n.intervals = 100,
                    seed = 1234)
```

Index

∗ **&htest**
    diffcor.dep, 2
    diffcor.one, 3
    diffcor.two, 4
    diffpwr.dep, 6
    diffpwr.one, 8
    diffpwr.two, 10
∗ **graphs**
    visual_mc, 12
∗ **utilities**
    visual_mc, 12

diffcor.dep, 2
diffcor.one, 3
diffcor.two, 4
diffpwr.dep, 6
diffpwr.one, 8
diffpwr.two, 10

visual_mc, 12