Package ‘reproducer’

February 23, 2021

Title Reproduce Statistical Analyses and Meta-Analyses
Version 0.4.2
Date 2021-02-23
Encoding UTF-8
Maintainer Lech Madeyski <lech.madeyski@gmail.com>

Description Includes data analysis functions (e.g., to calculate effect sizes and 95% Confidence Intervals (CI) on Standardised Effect Sizes (d) for AB/BA cross-over repeated-measures experimental designs), data presentation functions (e.g., density curve overlaid on histogram), and the data sets analyzed in different research papers in software engineering (e.g., related to software defect prediction or multi-site experiment concerning the extent to which structured abstracts were clearer and more complete than conventional abstracts) to streamline reproducible research in software engineering.

Depends R (>= 3.5.0)
License GPL (>= 2)
LazyData true

URL https://madeyski.e-informatyka.pl/reproducible-research/
Imports openxlsx (>= 2.4.0), ggplot2 (>= 2.0.0), gridExtra(>= 0.9.1), xtable(>= 1.7-4), metafor(>= 1.9-2), lme4(>= 1.1-10), MASS(>= 7.3-45), stats(>= 3.5.2), reshape(>= 0.8.8), GetoptLong(>= 0.1.7), dplyr(>= 0.8.0.1), httr(>= 1.4.0), jsonlite(>= 1.6), tidyr(>= 0.8.3), readr(>= 1.3.1), stringr(>= 1.4.0), tibble(>= 2.1.1)

RoxygenNote 7.1.1
Suggests testthat, assertthat

NeedsCompilation no

Author Lech Madeyski [cre, aut, ctb] (The main contributor and maintainer), Barbara Kitchenham [ctb] (Data and code contributor), Tomasz Lewowski [ctb] (Data and code contributor), Marian Jureczko [ctb] (Data contributor), David Budgen [ctb] (Data contributor), Pearl Brereton [ctb] (Data contributor),
Jacky Keung [ctb] (Data contributor),
Stuart Charters [ctb] (Data contributor),
Shirley Gibbs [ctb] (Data contributor),
Amnart Pohthong [ctb] (Data contributor),
Giuseppe Scanniello [ctb] (Data contributor),
Carmine Gravino [ctb] (Data contributor)

Repository CRAN
Date/Publication 2021-02-23 18:40:02 UTC

R topics documented:

aggregateIndividualDocumentStatistics ........................................... 4
boxplotAndDensityCurveOnHistogram .............................................. 5
boxplotHV ................................................................. 6
Calc4GroupNPStats ........................................................... 6
calculateBasicStatistics .......................................................... 7
calculateGroupSummaryStatistics ................................................... 8
calculateHg ........................................................................ 9
CalculateLevel2ExperimentRData ..................................................... 10
calculatePhat .................................................................. 12
CalculateRLevel1 ................................................................. 13
calculateSmallSampleSizeAdjustment ............................................. 15
CalculateTheoreticalEffectSizes .................................................... 16
Ciolkowski09ESEM.MetaAnalysis.PBRvsCBRolAR ................................ 17
Cliffd ................................................................. 18
constructEffectSizes ............................................................... 19
ConstructLevel1ExperimentRData ................................................... 20
densityCurveOnHistogram ........................................................... 22
effectSizeCI ................................................................. 23
ExtractExperimentData ............................................................... 24
ExtractGroupSizeData ............................................................... 25
ExtractMAStatistics ................................................................. 27
ExtractSummaryStatisticsRandomizedExp ........................................ 29
fmt ........................................................................ 30
getEffectSizesABBA ............................................................... 31
getEffectSizesABBAIgnoringPeriodEffect ........................................ 32
getSimulationData ................................................................. 33
getTheoreticalEffectSizeVariancesABBA ........................................ 34
Kendallaupb ................................................................. 35
KitchenhamEtAl.CorrelationsAmongParticipants.Abrahao13TSE ............. 36
KitchenhamEtAl.CorrelationsAmongParticipants.Gravino15JVLC ........... 38
KitchenhamEtAl.CorrelationsAmongParticipants.Madeyski10 .............. 39
KitchenhamEtAl.CorrelationsAmongParticipants.Reggio15SSM ............. 41
KitchenhamEtAl.CorrelationsAmongParticipants.Ricca10TSE ............... 42
KitchenhamEtAl.CorrelationsAmongParticipants.Ricca14TOSEM ........... 44
KitchenhamEtAl.CorrelationsAmongParticipants.Romano18ESEM .......... 45
KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14EASE ....... 47
R topics documented:

- KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14JVLC
- KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM
- KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello15EMSE
- KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello17TOSEM
- KitchenhamEtAl.CorrelationsAmongParticipants.Torchiano17JVLC
- KitchenhamMadeyski.SimulatedCrossoverDataSets
- KitchenhamMadeyskiABBAReportedResults
- KitchenhamMadeyskiABBAReportedEffectSizes
- KitchenhamMadeyskiDocData
- KitchenhamMadeyskiExpData
- KitchenhamMadeyskiMetaAnalysisReportedResults
- KitchenhamMadeyskiReportedEffectSizes
- KitchenhamMadeyskiBudgen16.COCOMO
- KitchenhamMadeyskiBudgen16.DiffInDiffData
- KitchenhamMadeyskiBudgen16.FINNISH
- KitchenhamMadeyskiBudgen16.PolishData
- KitchenhamMadeyskiBudgen16.PolishSubjects
- KitchenhamMadeyskiBudgen16.SubjectData
- LaplaceDist
- Madeyski15EISEJ.OpenProjects
- Madeyski15EISEJ.PropProjects
- Madeyski15EISEJ.StudProjects
- Madeyski15SQJ.NDC
- MadeyskiKitchenham.EUBASData
- MadeyskiKitchenham.MetaAnalysis.PBRvsCBRorAR
- MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20190324
- MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20191022
- MetaAnalysisSimulations
- NP2GroupMetaAnalysisSimulation
- NP4GroupMetaAnalysisSimulation
- percentageInaccuracyOfLargeSampleVarianceApproximation
- plotOutcomesForIndividualsInEachSequenceGroup
- PrepareForMetaAnalysisGtoR
- printXTable
- proportionOfSignificantTValuesUsingCorrectAnalysis
- proportionOfSignificantTValuesUsingIncorrectAnalysis
- RandomExperimentSimulations
- RandomizedBlockDesignEffectSizes
- RandomizedBlocksAnalysis
- RandomizedBlocksExperimentSimulations
- RandomizedDesignEffectSizes
- readExcelSheet
- reproduceForestPlotRandomEffects
- reproduceMixedEffectsAnalysisWithEstimatedVarianceAndExperimentalDesignModerator
- reproduceMixedEffectsAnalysisWithExperimentalDesignModerator
- reproduceMixedEffectsForestPlotWithExperimentalDesignModerator
- reproduceSimulationResultsBasedOn500Reps1000Obs
- reproduceTablesOfPaperMetaAnalysisForFamiliesOfExperiments
This function assumes an ABBA crossover experiment has reported means and variances for each technique in each time period. We calculate the weighted mean and pooled within group variance for the observations arising from the two different sets of materials for a specific technique.

**Usage**

```
```

**Arguments**

- **D1.M**
  - is a vector of mean values from a set of experiments in a family reporting observations from participants using a specific document in the first time period with either the control or the treatment technique.

- **D1.SD**
  - is a vector of results from the set of experiment in a family reporting the standard deviations of observations from participants using the same document in the first time period with the same technique.

- **D1.N**
  - is a vector of the numbers of participants in each experiment in a family, using the same document for participants using either the same technique.

- **D2.M**
  - is a vector of mean values of observations from participants using the alternative document in the second time period, but using the same technique.

- **D2.SD**
  - is a vector of the standard deviations of observations from participants using the alternative document in the second time period with the same technique.

- **D2.N**
  - is a vector of the numbers of participants using the same document in the second time period for participants using the same technique.
Value
data frame incl. the overall weighted mean and pooled standard deviation

Author(s)
Barbara Kitchenham and Lech Madeyski

Examples
aggregateIndividualDocumentStatistics(10, 2, 20, 15, 2, 20)
# M SD
#1 12.5 2

Description
Boxplot and density curve overlaid on histogram

Usage
boxplotAndDensityCurveOnHistogram(df, colName, limLow, limHigh)

Arguments
df          Data frame with data to be displayed
colName     Name of the selected column in a given data frame
limLow      the limit on the lower side of the displayed range
limHigh     the limit on the higher side of the displayed range

Value
A figure being a density curve overlaid on histogram

Author(s)
Lech Madeyski

Examples
library(ggplot2)
library(grid)
library(gridExtra)
boxplotAndDensityCurveOnHistogram(Madeyski15EISEJ.PropProjects, "STUD", 0, 100)
boxplotAndDensityCurveOnHistogram(Madeyski15SQJ.NDC, "simple", 0, 100)
Calc4GroupNPStats

boxplotHV

Description

Box plot

Usage

boxplotHV(df, colName, limLow, limHigh, isHorizontal)

Arguments

der    Data frame with data to be displayed
colName    Name of the selected column in a given data frame
limLow    the limit on the lower side of the displayed range
limHigh    the limit on the higher side of the displayed range
isHorizontal    Boolean value to control whether the box plot should be horizontal or not (i.e., vertical)

Value

A box plot

Author(s)

Lech Madeyski

Examples

boxplotHV(Madeyski15EISEJ.PropProjects, "STUD", 0, 100, TRUE)
boxplotHV(Madeyski15EISEJ.PropProjects, "STUD", 0, 100, FALSE)
boxplotHV(Madeyski15SQJ.NDC, "simple", 0, 100, FALSE)
boxplotHV(Madeyski15SQJ.NDC, "simple", 0, 100, TRUE)

Calc4GroupNPStats

Calc4GroupNPStats

Description

This function does a non-parametric analysis of a randomized blocks experiment assuming 2 blocks and 2 treatment conditions.

Usage

Calc4GroupNPStats(x1, x2, x3, x4, sigfig = -1, alpha = 0.05)
Arguments

- `x1` is the data associated with treatment A in one block 1
- `x2` is the data associated with treatment B in block 1
- `x3` is the data associated with treatment A in block 2
- `x4` is the data associated with treatment B in block 2
- `sigfig` is the number of significant digits in the data. If >0 the data will be appropriately truncated
- `alpha` is the significance level for all statistical tests

Value

The function returns the point biserial version of Kendall’s tau and its variance, Cliff’s d and its variance, the probability of superiority, phat, and its variance, for the 4 group experiment experiment.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

```r
set.seed(123)
x=list()
x[[1]]=stats::rnorm(10,0,1)
x[[2]]=stats::rnorm(10,0.8,1)
x[[3]]=stats::rnorm(10,0.5,1)
x[[4]]=stats::rnorm(10,1.3,1)
Calc4GroupNPStats(x[[1]],x[[2]],x[[3]],x[[4]],sigfig=-1,alpha=0.05)
# A tibble: 1 x 17
#  N  phat phat.var phat.df phat.test phat.pvalue phat.sig d  vard d.sig cor sqse
#1 40 0.17 0.00497 31.0 -4.68 0.0000532 TRUE -0.660 0.0206 TRUE -0.347 0.0132
```

Description

This function calculates the following statistics for a set of data: length, mean, median, variance, standard error of the mean, and confidence interval bounds. The input data must be a vector of 2 or more numerical values.

Usage

```r
calculateBasicStatistics(x, alpha = 0.05)
```
calculateGroupSummaryStatistics

Arguments

x The data to be summarized

alpha The probability level to be used when constructing the confidence interval bounds.

Value

A dataframe comprising the length, mean, variance, standard error and confidence limit bounds of the input data x.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

ShortExperimentNames=c("E1","E2","E3","E4")
FullExperimentNames=c("EUBAS","R1UCLM","R2UCLM","R3UCLM")
Metrics=c("Comprehension","Modification")
Groups=c("A","B","C","D")
Type=c(rep("4G",4))
StudyID="S2"
Control="SC"
ReshapedData=ExtractExperimentData(KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM ,ExperimentNames=FullExperimentNames, idvar="ParticipantID",timevar="Period",ConvertToWide=TRUE)
NewTable= ConstructLevel1ExperimentRData(ReshapedData,StudyID,ShortExperimentNames,Groups, Metrics,Type,Control)
calculateBasicStatistics(NewTable$r)

# N Mean Median Variance SE LowerBound UpperBound
# 1 32 0.06175 0.1688 0.2482 0.08808 -0.1109 0.2344

calculateGroupSummaryStatistics

description

This function calculates the following statistics data within groups: length, mean, median, variance, standard error of the mean, and confidence interval bounds.

Usage

calculateGroupSummaryStatistics(x, Group)

Arguments

x The data to be summarized. This must be a vector of 2 or more numerical values

Group The categorical data data defining the groups. This must vector of the same length as x containing factors specifying the data groups
Value

A dataframe comprising the number, mean, variance, standard error and confidence limit bounds of the data in each category

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

```r
ShortExperimentNames=c("E1","E2","E3","E4")
FullExperimentNames=c("EUAS","RIUCML","R2UCLM","R3UCLM")
Metrics=c("Comprehension","Modification")
Groups=c("A","B","C","D")
Type=c(rep("4G",4))
StudyID="S2"
Control="SC"

ReshapedData= ExtractExperimentData(
    KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM,
    ExperimentNames=FullExperimentNames, idvar="ParticipantID",timevar="Period",
    ConvertToWide=TRUE
)

NewTable= ConstructLevel1ExperimentRData(ReshapedData, StudyID,
    ShortExperimentNames, Groups, Metrics, Type, Control
)

SeqGroupLev=NULL
N.NT=length(NewTable$r)
for (i in 1:N.NT) {
    if (NewTable$n[i]<=8) SeqGroupLev[i]=as.character(NewTable$n[i])
    if (NewTable$n[i]>8) SeqGroupLev[i]=as.character(9)
}
calculateGroupSummaryStatistics(NewTable$r,Group=SeqGroupLev)
```

```markdown
<table>
<thead>
<tr>
<th>#</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Variance</th>
<th>StDev</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>-0.0833</td>
<td>-0.1699</td>
<td>0.2314</td>
<td>0.4810</td>
<td>0.2405</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>0.3658</td>
<td>0.4477</td>
<td>0.2109</td>
<td>0.4592</td>
<td>0.1326</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>-0.1300</td>
<td>-0.2214</td>
<td>0.1933</td>
<td>0.4397</td>
<td>0.1099</td>
</tr>
</tbody>
</table>
```

Description

This function calculates Hedges g and Hedges g adjusted given the basic experimental statistics - the mean values for participants, number of observations (participants), and standard deviation in both the control group and the treatment group. Hence, the function assumes the data is held as summary statistics including the control group mean, standard deviation and sample size and equivalent values for treatment group.
Usage

`calculateHg(Mc, Mt, Nc, Nt, SDc, SDt)`

Arguments

- `Mc` is a vector containing the mean value of the control group for each experiment.
- `Mt` is a vector containing the mean value of the treatment group for each experiment.
- `Nc` is a vector containing the number of observations (participants) in the control group for each experiment.
- `Nt` is a vector of the number of observations (participants) in the treatment group for each experiment.
- `SDc` is a vector of the standard deviations of the control group for each experiment.
- `SDt` is a vector of the standard deviations of the treatment group for each experiment.

Value
data frame composed of Hedges’ g and Hedges’ g adjusted effect sizes

Author(s)
Barbara Kitchenham and Lech Madeyski

Examples

```r
calculateHg(10, 15, 20, 20, 2, 2)
#   Hg   HgAdjusted
# 1 2.5  2.450276
```

Description

This function analyses data on r values obtained in the format obtained from the `ConstructLevel1ExperimentRData` function and finds the r-value for each metric for each experiment.

Usage

```r
CalculateLevel2ExperimentRData(
    Level1Data,
    Groups,
    StudyID,
    ExperimentNames,
    Metrics,
    Type
)
```
CalculateLevel2ExperimentRData

Arguments

Level1Data: a tibble in the format produced by the ConstructLevel1ExperimentRData function which has r-values for each sequence group in a crossover experiment.

Groups: This is a list that defines the sequence group labels used in the dataset.

StudyID: This holds an identifier used to identify the origin of the experimental data in the output from this function.

ExperimentNames: This a list of identifiers used to define each experiment in the output from this function.

Metrics: This is a list of of character strings identifying each outcome metric reported in each of the experiments in the set of replicated experiments.

Type: this is a list of character strings specifying for each experiment whether the experiment is a two sequence group "2G" or four sequence group "4G" experiment. return RExp.Table This is a table containing the pooled data variance and the pooled difference variance for the experiment and the value r for the experiment for each metric.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

ShortExperimentNames=c("E1","E2","E3","E4")
FullExperimentNames=c("EUBAS","R1UCLM","R2UCLM","R3UCLM")
Metrics=c("Comprehension","Modification")
Groups=c("A","B","C","D")
Type=c(rep("4G",4))
StudyID="S2"
Control="SC"

# Obtain experimental data from each file and put in wide format
ReshapedData = ExtractExperimentData(
  KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM,
  ExperimentNames=FullExperimentNames, idvar="ParticipantID",timevar="Period",ConvertToWide=TRUE)
Lev1Data= ConstructLevel1ExperimentRData(ReshapedData, StudyID, ShortExperimentNames, Groups,
  Metrics, Type, Control)
CalculateLevel2ExperimentRData(Lev1Data,Groups=Groups,StudyID=StudyID,
  ExperimentNames=ShortExperimentNames,Metrics=Metrics,Type=Type)

# A tibble: 8 x 10
#  StudyID ExpID N Metric PooledVar1 PooledVar2 VarProp PooledVar PooledDiffVar r.Exp
#<chr> <chr> <int> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
# 1 S2 SZE1 24 Comprehension 0.0148 0.0212 0.412 0.0180 0.0248 0.311
# 3 S2 SZE2 22 Comprehension 0.0487 0.0224 0.684 0.0356 0.0534 0.250
# 4 S2 SZE2 22 Modification 0.0445 0.0266 0.626 0.0356 0.0628 0.117
# 5 S2 SZE3 22 Comprehension 0.0353 0.0402 0.467 0.0377 0.105 -0.391
# 6 S2 SZE3 22 Modification 0.0433 0.0414 0.511 0.0424 0.0997 -0.176
# 7 S2 SZE4 18 Comprehension 0.0439 0.0237 0.649 0.0338 0.0894 0.0222
**Description**

This function extracts the probability of superiority (i.e., Phat) and its confidence interval based on Brunner and Munzel (2000) heteroscedastic analog of WMW test. It is based on Wilcox’s `bmp` function with some amendments. It does not include a plotit facility. It uses the smallest non-zero variance to identify confidence intervals and statistical significance for values of Phat=0 and Phat=1. It ensures that confidence intervals do not take on invalid values such as values <0 or >1.

**Usage**

```r
calculatePhat(x, y, alpha = 0.05, sigfig = -1)
```

**Arguments**

- `x` is a vector of values from group 1
- `y` is a vector of values from group 2
- `alpha` is the Type 1 error level for statistical tests
- `sigfig` If `sigfig>0` the data in `x` and `y` is truncated to the specified number of significant digits.

**Value**

A list including the value of the t-test for PHat, the estimate of PHat and Cliff’s d, and the confidence intervals for PHat.

**Author(s)**

Rand Wilcox amendments by Barbara Kitchenham and Lech Madeyski

**Examples**

```r
x=c(1.2, 3.0, 2.2, 4.0, 2.5, 3.0)
y=c(3.4.2,4,6,7,5.9)
calculatePhat(x,y)
# $test.stat
# [1] 6.381249
# $phat
# [1] 0.9305556
# $dhat
# [1] 0.8611111
# $sig.level
# [1] 0.0001191725
# $s.e.
# [1] 0.06747199
# $ci.p
```
Description
This function calculates the r value for a 2-group (2G) or 4-Group (4G) Crossover experiment for each sequence group and each outcome metric. The function returns both the exact r value and the r value based on pooled variances for each sequence group and outcome metric.

Usage
CalculateRLevel1(
  Dataset, StudyID, Groups = c("A", "B", "C", "D"),
  ExperimentName, Metrics, Type, Control
)

Arguments
Dataset This holds the data for each participant in a 2-group or 4-group crossover experiment in the "wide" format. I.e., there is only one entry per participant. The data set should have been generated from a long version of the data based on a variable labelled "Period" which is used to define which participant data was collected in the first period of the experiment - see function ExtractLevel1ExperimentRData.
CalculateRLevel1

**StudyID**
This holds an identifier used to identify the origin of the experimental data in the output from this function.

**Groups**
This is a list that defined the sequence group identifiers used in the dataset.

**ExperimentName**
This an identifiers used to define the specific experiment in the output from this function.

**Metrics**
This is a list of metrics, e.g., ("Correctness","Time","Efficiency").

**Type**
This is a character string specifying whether the experiment is a two sequence group or four sequence group experiment.

**Control**
This is a character string that defines the control treatment in the experiment.

**Details**

script to obtain correlation coefficients

**Value**

This table is a tibble holding information identifying for each metric and sequence group the first time period and second time period variance, the pooled variance, the variance of the difference values and the exact r and pooled r.

```r
# importFrom stats
# importFrom var
# importFrom tibble

Author(s)
Barbara Kitchenham and Lech Madeyski

Examples

```r
ExperimentNames=c("EUBAS","R1UCLM","R2UCLM","R3UCLM")
ShortExperimentNames=c("E1","E2","E3","E4")
Metrics=c("Comprehension","Modification")
Type=c("4G", "4G", "4G", "4G")
Groups=c("A", "B", "C", "D")
StudyID="S2"
Control="SC"

# Obtain experimental data from a file and put in wide format
dataset2= KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM
ReshapedData=ExtractExperimentData(dataset2, ExperimentNames=ExperimentNames, idvar="ParticipantID",timevar="Period",ConvertToWide=TRUE)

# Calculate the correlations for each sequence group and each metric.
CalculateRLevel1(Dataset=ReshapedData[[1]], StudyID, Groups=c("A","B","C","D"), ExperimentName=ShortExperimentNames[1],Metrics,Type=Type[1],Control)
```

A tibble: 8 x 15

# A tibble: 8 x 15
# # A tibble: 8 x 15
# Study Exp Group Metric Id n ControlFirst var1 var2
# <chr> <chr> <chr> <chr> <chr> <int> <lgl> <dbl> <dbl>
# 1 S2 E1 A Compr. S2E1A 6 FALSE 0.0183 0.0163
# 2 S2 E1 B Compr. S2E1B 6 TRUE 0.0201 0.0326
# 3 S2 E1 C Compr. S2E1C 6 FALSE 0.00370 0.0155
# 4 S2 E1 D Compr. S2E1D 6 TRUE 0.0173 0.0201
# 5 S2 E1 A Modif. S2E1A 6 FALSE 0.0527 0.0383
# 6 S2 E1 B Modif. S2E1B 6 TRUE 0.0185 0.0482
```
Function calculates the Hedges small sample size adjustment for standardized mean effect sizes. It calculates the exact value unless the caller sets the parameter exact to FALSE, or the degrees of freedom is too large.

Function calculates the small sample size adjustment for standardized mean effect sizes

Usage

calculateSmallSampleSizeAdjustment(df, exact = TRUE)

calculateSmallSampleSizeAdjustment(df, exact = TRUE)

Arguments

df
A vector of degrees of freedom

exact
Default value=TRUE, if exact==TRUE the function returns the exact value of the adjustment(s) which is suitable for small values of df, if exact==FALSE the function returns the approximate version of the adjustment(s). See Hedges and Olkin 'Statistical methods for Meta-Analysis' Academic Press 1985.

Value

small sample size adjustment value

Author(s)

Barbara Kitchenham and Lech Madeyski
Examples

df <- 2
C <- calculateSmallSampleSizeAdjustment(df)

df = c(5, 10, 17)
adjexact = calculateSmallSampleSizeAdjustment(df)
# adjexact = 0.8407487 0.9227456 0.9551115
# Hedges and Olkin values 0.8408, 0.9228, 0.9551
adjapprox = calculateSmallSampleSizeAdjustment(df, FALSE)
# adjapprox = 0.8421053 0.9230769 0.9552239
df <- 2
a <- calculateSmallSampleSizeAdjustment(df)
# > a
# [1] 0.5641896

df = c(5, 10, 17)
adjexact = calculateSmallSampleSizeAdjustment(df)
# > adjexact
# [1] 0.8407487 0.9227456 0.9551115
# Hedges and Olkin values 0.8408, 0.9228, 0.9551
adjapprox = calculateSmallSampleSizeAdjustment(df, FALSE)
# > adjapprox
# [1] 0.8421053 0.9230769 0.9552239
# Another example:
df = c(10, 25, 50)
calculateSmallSampleSizeAdjustment(df, exact = TRUE)
# [1] 0.9227456 0.9696456 0.9849119
calculateSmallSampleSizeAdjustment(df, exact = FALSE)
# [1] 0.9230769 0.9696970 0.9849246

---

CalculateTheoreticalEffectSizes

**CalculateTheoreticalEffectSizes**

---

Description

This function constructs the theoretical effect sizes and distribution statistics four (normal, lognormal, Laplace & gamma) given specific parameter values for the distributions

Usage

`CalculateTheoreticalEffectSizes(mean, std, type = "n")`

Arguments

- **mean**: The theoretical central location parameter for the distribution specified by the type parameter.
- **std**: The theoretical spread parameter for the distribution specified by the type parameter.
type String identifying the distribution, "n" for normal, "ln" for lognormal, "lap" for Laplace, "g" for Gamma

Value
dataframe containing the expected standardized effect size, mean, variance, skewness and kurtosis statistics for samples from the specified distribution

Author(s)
Barbara Kitchenham and Lech Madeyski

Examples

CalculateTheoreticalEffectSizes(mean=0, std=1, type="l")
# A tibble: 1 x 5
# RawMean RawVariance RawEffectSize RawSkewness RawKurtosis
# <dbl> <dbl> <dbl> <dbl> <dbl>
# 1 1.65 4.67 0.763 6.18 88.5
CalculateTheoreticalEffectSizes(mean=0, std=1, type="n")
# A tibble: 1 x 5
# RawMean RawVariance RawEffectSize RawSkewness RawKurtosis
# <dbl> <dbl> <dbl> <dbl> <dbl>
# 1 0 1 0 0 3

Ciolkowski09ESEM.MetaAnalysis.PBRvsCBRorAR

Ciolkowski09ESEM.MetaAnalysis.PBRvsCBRorAR data form a set of primary studies on reading methods for software inspections. They were reported and analysed by M. Ciolkowski ("What do we know about perspective-based reading? an approach for quantitative aggregation in software engineering", in Proceedings of the 3rd International Symposium on Empirical Software Engineering and Measurement, ESEM'09, pp. 133-144, IEEE Computer Society, 2009), corrected and re-analysed by Madeyski and Kitchenham ("How variations in experimental designs impact the construction of comparable effect sizes for meta-analysis" (to be submitted)).

Description

If you use this data set please cite: Lech Madeyski and Barbara Kitchenham, "How variations in experimental designs impact the construction of comparable effect sizes for meta-analysis", 2015.

Usage

Ciolkowski09ESEM.MetaAnalysis.PBRvsCBRorAR
**Format**

A data frame with 21 rows and 7 variables:

- **Study**: Name of empirical study
- **Ref.**: Reference to the paper reporting primary study or experimental run where data were originally reported
- **Control**: Control treatment: Check-Based Reading (CBR) or Ad-hoc Reading (AR)
- **Within-subjects**: Yes - if the primary study used the within-subjects experimental design, No - if the primary study did not use the within-subjects experimental design
- **Cross-over**: Yes - if the primary study used the cross-over experimental design, No - if the primary study did not use the cross-over experimental design
- **d_ByCiolkowski**: d effect size calculated by Ciolkowski
- **d_ByOriginalAuthors**: d effect size as reported by the original authors

**Source**

[https://madeyski.e-informatyka.pl/reproducible-research/](https://madeyski.e-informatyka.pl/reproducible-research/)

**Examples**

Ciolkowski09ESEM.MetaAnalysis.PBRvsCBRorAR

---

**Cliffd**

**Cliffd**

**Description**

This function implements finds Cliff’s d and its confidence intervals. The null hypothesis is that for two independent group, P(X<Y)=P(X>Y). The function reports a 1-alpha confidence interval for P(X>Y)-P(X<Y). The algorithm computes a confidence interval for Cliff’s d using the method in Cliff, 1996, p. 140, eq 5.12. The function is based on code produce by Rand Wilcox but has been amended. The plotting function has been removed and the dependency on Wilcox’s binomci function has been removed. Construction of confidence intervals if values in one group are all larger than values in the other group has been amended to use the smallest non-zero variance method. Upper and lower confidence interval bounds cannot assume invalid values, i.e. values <-1 or >1.

**Usage**

Cliffd(x, y, alpha = 0.05, sigfig = -1)

**Arguments**

- **x** is a vector of values from group 1
- **y** is a vector of values from group 2
- **alpha** is the Type 1 error level for statistical tests
- **sigfig** is the number of significant digit. If sigfig>0 the data in x and y is truncated to the specified value.
Value

list including the value of Cliff's $d$ and its consistent variance and confidence intervals and the equivalent probability of superiority value and its confidence intervals. \[ x=c(1.2,3,2.2,4,2.5,3) \]
\[ y=c(3,4.2,4,6,7,5.9) \]

```
\[ \text{Cliff}(x,y) \]

\[ n1 \]
\[ [1] 6 \]

\[ n2 \]
\[ [1] 6 \]

\[ cl \]
\[ [1] -0.9772519 \]

\[ cu \]
\[ [1] -0.3476461 \]

\[ d \]
\[ [1] -0.8611111 \]

\[ sqse.d \]
\[ [1] 0.02017931 \]

\[ phat \]
\[ [1] 0.06944444 \]

\[ summary.dvals \]
\[ \text{P}(X<Y) \]
\[ [1,] 0.8888889 \]

\[ \text{P}(X=Y) \]
\[ [1] 0.08333333 \]

\[ \text{P}(X>Y) \]
\[ [1] 0.02777778 \]

\[ p.cl \]
\[ [1] 0.01137405 \]

\[ p.cu \]
\[ [1] 0.06944444 \]

```

Author(s)

Rand Wilcox, amendments Barbara Kitchenham and Lech Madeyski

Description

The function constructs various different $d$-style effect sizes for a set of different experiments given basic statistics from each experiment (the mean value of the control group $M_c$, the mean value of the treatment group $M_t$, the standard deviation of the control group $S_{Dc}$, standard deviation of the the treatment group $S_{Dt}$, the number of observations (participants) in the control group $N_c$, and the number of observations (participants) in the treatment group $N_t$). The input variables can be vectors or individual numbers but all input vectors must be of the same length. The function returns Glass's Delta, Cohen's $D$, point bi-serial $r$ (based on Hedges' $g$ unadjusted), Hedges' $g$ and Hegdes' $g$ adjusted for small sample size.

Usage

```
constructEffectSizes(Mc, Mt, SDc, SDt, Nc, Nt)
```

Arguments

- **Mc** is a vector containing the mean value of the control group for each experiment.
- **Mt** is a vector containing the mean value of the treatment group for each experiment.
- **SDc** is a vector of the standard deviations of the control group for each experiment.
- **SDt** is a vector of the standard deviations of the the treatment group for each experiment.
- **Nc** is a vector containing the the number of observations (participants) in the control group for each experiment.
- **Nt** is a vector of the number of observations (participants) in the treatment group for each experiment.
Value

data frame composed of five effect sizes (Glass delta, Cohen’s d, Hedges’ g, r, Hedges’ g adjusted)

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

constructEffectSizes(10, 15, 0.3, 0.2, 15, 15)

Mt = c(0.633, 0.673, 0.423, 0.727, 0.631)
Mc = c(0.612, 0.526, 0.356, 0.618, 0.534)
SDt = c(0.198, 0.115, 0.172, 0.088, 0.122)
SDc = c(0.159, 0.089, 0.111, 0.166, 0.119)
Nt = c(12, 12, 14, 10, 8)
Nc = c(12, 12, 14, 10, 8)
EffectSizes=constructEffectSizes(Mc, Mt, SDc, SDt, Nt, Nc)

<table>
<thead>
<tr>
<th>#</th>
<th>GlassDelta</th>
<th>Cohend</th>
<th>Hedgesg</th>
<th>r</th>
<th>HedgesgAdjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1320755</td>
<td>0.1221516</td>
<td>0.1169513</td>
<td>0.05837591</td>
<td>0.1129107</td>
</tr>
<tr>
<td>2</td>
<td>1.6516854</td>
<td>1.4931812</td>
<td>1.4296121</td>
<td>0.58151846</td>
<td>1.3802200</td>
</tr>
<tr>
<td>3</td>
<td>0.6036036</td>
<td>0.4803405</td>
<td>0.4628677</td>
<td>0.22547423</td>
<td>0.4493641</td>
</tr>
<tr>
<td>4</td>
<td>0.6566265</td>
<td>0.8648343</td>
<td>0.8204538</td>
<td>0.37953300</td>
<td>0.7857047</td>
</tr>
<tr>
<td>5</td>
<td>0.8151261</td>
<td>0.8604924</td>
<td>0.8049169</td>
<td>0.37335594</td>
<td>0.7608781</td>
</tr>
</tbody>
</table>

Description

This function returns the r value for a 2-group (2G) or 4-Group (4G) Crossover experiment for a group of 1 or more experiments for each sequence group and each outcome metric. For sets of 2 or more experiments, the experiments are assumed to be replicates and to report the same sets of Metrics and have the same Control treatment and use the same sequence Group identifiers, but are not necessarily the same Type. We return both the exact r value and the r value based on pooled variances for each sequence group and outcome metric.

Usage

ConstructLevel1ExperimentRData(
  Data,
  StudyID,
  ExperimentNames,
  Groups,
  Metrics,
  Type,
ConstructLevel1ExperimentRData

Control
)

Arguments

Data This is a list parameter each entry in the list holds the data for each participant in a 2-group or 4-group crossover experiment in the "wide" format. I.e., there is only one entry per participant. The data should have been generated from a long version of the data based on a variable labelled "Period" which is used to define which participant data was collected in the first period of the experiment - see function ExtractLevel1ExperimentRData.

StudyID This holds an identifier used to identify the origin of the experimental data in the output from this function.

ExperimentNames This a list of identifiers used to define each experiment in the output from this function.

Groups This is a list that defined the sequence group identifiers used in the dataset.

Metrics This is a list of of character strings identifying each outcome metric reported in each of the experiments in the set of replicated experiments.

Type this is a list of character strings specifying for each experiment whether the experiment is a 2-group or 4-group experiment

Control this is a character string that defines the control treatment in the experiment.

Value

R.Data.Table this is a tibble holding information identifying for each metric and sequence group the first time period and second time period variance, the pooled variance, the variance of the difference values and the exact r and pooled r.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

# ShortExperimentNames=c("E1","E2","E3","E4")
FullExperimentNames=c("EUBAS","R1UCLM","R2UCLM","R3UCLM")
Metrics=c("Comprehension","Modification")
Groups=Groups=c("A","B","C","D")
Type=c(rep("4G",4))
StudyID="S2"
Control="SC"

# Obtain experimental data from each file and put in wide format
dataset2=KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM
ReshapedData=ExtractExperimentData(dataset2, ExperimentNames=FullExperimentNames, idvar="ParticipantID",timevar="Period",ConvertToWide=TRUE)

# Calculate the correlations for each sequence group and each metric in each experiment
ConstructLevel1ExperimentRData(Data=ReshapedData, StudyID=StudyID,"
densityCurveOnHistogram

Description

Density curve overlaid on histogram

Usage

densityCurveOnHistogram(df, colName, limLow, limHigh)

Arguments

df
  Data frame with data to be displayed

colName
  Name of the selected column in a given data frame

limLow
  the limit on the lower side of the displayed range

limHigh
  the limit on the higher side of the displayed range

Value

A figure being a density curve overlaid on histogram

Author(s)

Lech Madeyski

Examples

densityCurveOnHistogram(Madeyski15EISEJ.PropProjects, "STUD", 0, 100)
  # densityCurveOnHistogram(data.frame(x<-rnorm(50, mean=50, sd=5)), "x", 0, 100)
Description

The procedure is based on finding the upper and lower 0.025 bounds for the related t-variable. The t-variable needs to be adjusted for bias by multiplying by c. The upper and lower bounds on the t-variable are then used to calculate upper and lower bounds on the repeated measures effect size (d_RM) by multiplying the upper and lower bound of the t-variable by sqrt((n1+n2)/(2*(n1*n2))). Upper and lower bounds on the equivalent independent groups effect size (d_IG) are found by multiplying the upper and lower bounds on d_RM by sqrt(1-r).

Usage

```r
effectSizeCI(
  expDesign,
  t,
  n1,
  n2,
  r = 0,
  epsilon = 1e-10,
  maxsteps = 1000,
  stepsize = 3
)
```

Arguments

- **expDesign**: Experimental design: 1) crossover repeated measures ("CrossOverRM"), 2) before-after repeated measures (expDesign="BeforeAfterRM"), 3) independent groups ("IG")
- **t**: t-statistics (t must be less than or equal to 37.62, the limit from the R function documentation)
- **n1**: The number of observations in sequence group 1 (expDesign="CrossOverRM"), the number of observations in group 1 (expDesign="IG"), or the total number of observations (expDesign="BeforeAfterRM")
- **n2**: The number of observations in sequence group 2 (expDesign="CrossOverRM") or the number of observations in group 2 (expDesign="IG")
- **r**: The correlation between outcomes for individual subject (the within subject correlation)
- **epsilon**: The precision of the iterative procedure
- **maxsteps**: The maximum number of steps of the iterative procedure (the procedure terminates at maxsteps or earlier if CI with enough precision have been calculated)
- **stepsize**: The size of steps (influences the convergence of the calculations, i.e., the number of steps required to obtain the final result of precision defined by the epsilon)
ExtractExperimentData

Value

A list of Confidence Intervals for: t-statistic (t_LB and t_UB), repeated-measures effect size d_RM (d_RM_LB, d_RM_UB), independent groups effect size (d_JG_LB, d_JG_UB)

Author(s)

Lech Madeyski and Barbara Kitchenham

Examples

```r
effectSizeCI(expDesign="CrossOverRM", t=14.4, n1=15, n2=15, r=0.6401)
effectSizeCI(expDesign = "BeforeAfterRM", t=14.16536, n1=15, n2=0, r=0.6146771)
effectSizeCI(expDesign = "IG", t=-6.344175, n1=15, n2=15)
effectSizeCI(expDesign="CrossOverRM", t=0.5581, n1=6, n2=6, r=0.36135)
effectSizeCI(expDesign = "CrossOverRM", r=0.855, t=4.33, n1=7, n2=6)
```

Description

This function reads datasets from a defined directory in the reproducer package that hold the results of a family crossover experiments in the long format. It converts the data to the wide format if required.

Usage

```r
ExtractExperimentData(
  DataSet,
  ExperimentNames,
  idvar = "ParticipantID",
  timevar = "Period",
  ConvertToWide = TRUE
)
```

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DataSet</td>
<td>This is a tibble holding the data for each crossover experiment in a family (a family can include only one experiment).</td>
</tr>
<tr>
<td>ExperimentNames</td>
<td>This is a list with the full names of each experiment.</td>
</tr>
<tr>
<td>idvar</td>
<td>This is the name of the column that contains the data for specific participants. It is only assumed to be unique within an experiment (default idvar=&quot;ParticipantID&quot;).</td>
</tr>
<tr>
<td>timevar</td>
<td>This is the name of the table column that defines which data was collected in a specific time period. This function assumes that there are only two time periods (default timevar=&quot;Period&quot;).</td>
</tr>
<tr>
<td>ConvertToWide</td>
<td>This determine whether the function converts the data to the wide format (default ConvertToWide=TRUE).</td>
</tr>
</tbody>
</table>
Value

A list with an entry for the data for each experiment. If ConvertToWide is TRUE, it returns the data in the wide format otherwise it returns the data as it was read. Within each list item the data is returned as a tibble #importFrom stats # importFrom tibble # importFrom base

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

ExperimentNames=c("EUBAS","R1UCLM","R2UCLM","R3UCLM")
Metrics=c("Comprehension","Modification")
Groups=c("A","B","C","D")
Type=c(rep("4G",4))
StudyID="S2"
Control="SC"

# Obtain experimental data from each file and put in wide format
dataset2= KitchenhamEtAl.CorrrelationsAmongParticipants.Scanniello14TOSEM
ReshapedData = ExtractExperimentData(dataset2, ExperimentNames=ExperimentNames, 
idvar="ParticipantID",timevar="Period",ConvertToWide=TRUE)
ReshapedData[[1]]

# A tibble: 24 x 15
#  ParticipantID ExperimentID.1 SequenceGroup.1 System.1 Treatment.1 Comprehension.1
#   <fct> <fct> <fct> <fct> <fct> <dbl>
# 1      1  EUBAS   A    S1    AM  0.77
# 2      5  EUBAS   A    S1    AM  0.61
# 3      9  EUBAS   A    S1    AM  0.61
# 4     13  EUBAS   A    S1    AM  0.52
# 5     17  EUBAS   A    S1    AM  0.43
# 6     21  EUBAS   A    S1    AM  0.77
# 7      2  EUBAS   B    S1    SC  0.92
# 8      6  EUBAS   B    S1    SC  0.63
# 9     10  EUBAS   B    S1    SC  0.51
#10     14  EUBAS   B    S1    SC  0.64
# ... with 14 more rows, and 9 more variables: Modification.1 <dbl>, CrossOverID.1 <fct>, 
# ExperimentID.2 <fct>, SequenceGroup.2 <fct>, System.2 <fct>, Treatment.2 <fct>, 
# Comprehension.2 <dbl>, Modification.2 <dbl>, CrossOverID.2 <fct>

Description

This function constructs a table identifying the number of participants in each sequence group for a set of experiments each of which used a crossover design.
Usage

```r
ExtractGroupSizeData(
  ExpDataWide,
  StudyID,
  ShortExperimentNames,
  Type,
  Groups = c("A", "B", "C", "D")
)
```

Arguments

- **ExpDataWide**: this is a list of tibbles each comprising data from one experiment in its wide format.
- **StudyID**: an identifier for the group of related experiments (i.e., a family).
- **ShortExperimentNames**: a list of character strings identifying each experiment.
- **Type**: A list identifying the type of crossover "2G" or "4G" for each experiment in the family.
- **Groups**: a list of the terms used to specify sequence groups in the experiments.

Value

A tibble containing the number of participants in each sequence group in each experiment.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

```r
ExperimentNames=c("EUBAS", "R1UCLM", "R2UCLM", "R3UCLM")
ShortExperimentNames=c("E1", "E2", "E3", "E4")
Metrics=c("Comprehension", "Modification")
Type=c("4G", "4G", "4G", "4G")
Groups=c("A", "B", "C", "D")
StudyID="S2"
Control="SC"
# Obtain experimental data from a file and put in wide format
dataset2= KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM
ReshapedData=ExtractExperimentData(dataset2, ExperimentNames=ExperimentNames, idvar="ParticipantID", timevar="Period", ConvertToWide=TRUE)
ExtractGroupSizeData(ReshapedData, StudyID, ShortExperimentNames, Type, Groups=Groups)
# A tibble: 16 x 4
# Study Exp Group n
# <chr> <chr> <chr> <int>
#1 S2 Exp1 A 6
#2 S2 Exp1 B 6
#3 S2 Exp1 C 6
```
## Description

This function extracts summary statistics from meta-analysis results obtained from the `rma` function of the `metafor` R package. If required the function transform back to standardized mean difference (effect size type "d" i.e. Hg) or point biserial correlations (effect size type "r"). Warning: the ‘ExtractMAStatistics’ function works with ‘metafor’ version 2.0-0, but changes to metafor’s method of providing access to its individual results may introduce errors into the function.

This function extracts summary statistics from meta-analysis results obtained from the `rma` function of the `metafor` R package. If required the function transform back to standardized mean difference (effect size type "d" i.e. Hg) or point biserial correlations (effect size type "r"). Warning: the ‘ExtractMAStatistics’ function works with ‘metafor’ version 2.0-0, but changes to metafor’s method of providing access to its individual results may introduce errors into the function.

## Usage

```r
ExtractMAStatistics(
  maresults,  # maresults
  Nc,         # Nc
  Nt,         # Nt
  Transform = TRUE,  # Transform = TRUE
  type = "d",    # type = "d"
  sig = 4,       # sig = 4
  returnse = FALSE  # returnse = FALSE
)

ExtractMAStatistics(
  maresults,  # maresults
  Nc,         # Nc
  Nt,         # Nt
  Transform = TRUE,  # Transform = TRUE
  type = "d",    # type = "d"
)```

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#4</td>
<td>S2</td>
<td>Exp1</td>
<td>D 6</td>
</tr>
<tr>
<td>#5</td>
<td>S2</td>
<td>Exp2</td>
<td>A 6</td>
</tr>
<tr>
<td>#6</td>
<td>S2</td>
<td>Exp2</td>
<td>B 6</td>
</tr>
<tr>
<td>#7</td>
<td>S2</td>
<td>Exp2</td>
<td>C 5</td>
</tr>
<tr>
<td>#8</td>
<td>S2</td>
<td>Exp2</td>
<td>D 5</td>
</tr>
<tr>
<td>#9</td>
<td>S2</td>
<td>Exp3</td>
<td>A 5</td>
</tr>
<tr>
<td>#10</td>
<td>S2</td>
<td>Exp3</td>
<td>B 5</td>
</tr>
<tr>
<td>#11</td>
<td>S2</td>
<td>Exp3</td>
<td>C 6</td>
</tr>
<tr>
<td>#12</td>
<td>S2</td>
<td>Exp3</td>
<td>D 6</td>
</tr>
<tr>
<td>#13</td>
<td>S2</td>
<td>Exp4</td>
<td>A 5</td>
</tr>
<tr>
<td>#14</td>
<td>S2</td>
<td>Exp4</td>
<td>B 5</td>
</tr>
<tr>
<td>#15</td>
<td>S2</td>
<td>Exp4</td>
<td>C 4</td>
</tr>
<tr>
<td>#16</td>
<td>S2</td>
<td>Exp4</td>
<td>D 4</td>
</tr>
</tbody>
</table>
sig = 4,
returnse = FALSE
)

Arguments

maresults is the output from the rma function.
Nc is the number of participants in the control condition group.
Nt is the number of participants in the treatment condition group.
Transform is a boolean value indicating whether the outcome values need to be transformed back to standardized mean difference ("d" i.e. Hg or d) or point biserial correlations ("r"). It is defaulted to TRUE. If this parameter is set to FALSE, no transformation will be applied.
type this indicates the type of transformation required - it defaults to 'd' which requests transformation from Zr to Hg, using "r" requests transformation from Zr to r.
sig indicates the number of significant digits requested in the output, the default is 4; it rounds the values of mean, pvalue, upper and lower bound to the specified number of significant digits.
returnse default to FALSE, if set to TRUE returns the standard error of the effect size

Value

data frame incl. summary statistics from meta-analysis results: overall mean value for the effect sizes, the p-value of the mean, the upper and lower confidence interval bounds (UB and LB), QE which is the heterogeneity test statistic and QEp which the p-value of the heterogeneity statistic

data frame incl. summary statistics from meta-analysis results: overall mean value for the effect sizes, the p-value of the mean, the upper and lower confidence interval bounds (UB and LB), QE which is the heterogeneity test statistic and QEp which the p-value of the heterogeneity statistic

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

ExpData=reproducer::KitchenhamMadeyskiBrereton.ExpData
#Extract the experiment basic statics
S1data=subset(ExpData,ExpData=="S1")
#Use the descriptive data to construct effect size
S1EffectSizes = reproducer::PrepareForMetaAnalysisGtoR(S1data$Mc,S1data$Mt,S1data$SDc,S1data$SDt,S1data$Nc,S1data$Nt)
# Do a random effect meta-analysis of the transformed r_pbs effect size
S1MA = metafor::rma(S1EffectSizes$zr, S1EffectSizes$vi)
# Extract summary statistics from meta-analysis results and transform back to Hg scale
S1MAStats=reproducer::ExtractMAStatistics(S1MA, sum(S1data$Nc),sum(S1data$Nt), TRUE, "d", 4)
# mean pvalue UB LB QE QEp
# Extract the experiment basic statistics
S1data = subset(ExpData, ExpData == "S1")
# Use the descriptive data to construct effect size
S1EffectSizes = reproduce::PrepareForMetaAnalysisGtoR(
    S1data$Mc, S1data$Mt, S1data$SDc, S1data$SDt, S1data$Nc, S1data$Nt)
# Do a random effect meta-analysis of the transformed r_pbs effect size
S1MA = metafor::rma(S1EffectSizes$zr, S1EffectSizes$vi)
# Extract summary statistics from meta-analysis results and transform back to Hg scale
ExtractMAStatistics(S1MA, sum(S1data$Nc), sum(S1data$Nt), TRUE, "d", 4)

# A tibble: 1 x 6
#  mean  pvalue    UB    LB   QE   QEp
#1 0.666 0.00207 1.12 0.238 4 0.41

---

**Description**

This function extracts data obtained from the lme4 package lmer function. It assumes a simple randomized experiment with each element having one or more repeated measures. It outputs the mean together with its standard error and confidence interval bounds.

**Usage**

ExtractSummaryStatisticsRandomizedExp(lmeRA, N, alpha = 0.05)

**Arguments**

- `lmeRA`  
  The output from the lmer function
- `N`  
  The total number of observations
- `alpha`  
  the probability level to be used when constructing the confidence interval bounds.

**Value**

`REASummary` A dataframe holding the number of observations `N`, the overall mean value as its standard error reported as by the lmer function, and its confidence interval bounds.

**Author(s)**

Barbara Kitchenham and Lech Madeyski
Examples

```r
ShortExperimentNames <- c("E1","E2","E3","E4")
FullExperimentNames <- c("EUBAS","R1UCLM","R2UCLM","R3UCLM")
Metrics <- c("Comprehension","Modification")
Groups <- c("A","B","C","D")
Type <- rep("4G",4)
StudyID <- "S2"
Control <- "SC"
ReshapedData <- ExtractExperimentData(
  KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM, 
  ExperimentNames=FullExperimentNames, idvar="ParticipantID", timevar="Period", 
  ConvertToWide=TRUE)
NewTable <- ConstructLevel1ExperimentRData(ReshapedData, StudyID, ShortExperimentNames, Groups,
  Metrics, Type, Control)
resRe <- lme4::lmer(r~(1|Id), data=NewTable)
summary(resRe)
```

```
# Linear mixed model fit by REML ['lmerMod']
# Formula: r ~ (1 | Id)
# REML criterion at convergence: 47.8
# Scaled residuals:
#   Min 1Q Median 3Q Max
# -1.4382 -0.9691 0.2190 0.8649 1.4761
# Random effects:
#   Groups Name Variance Std.Dev.
#     Id (Intercept) 0.03978 0.1994
#   Residual 0.20974 0.4580
# Number of obs: 32, groups: Id, 16
# Fixed effects:
#   Estimate Std. Error t value
# (Intercept) 0.06175 0.09508 0.649
```

```r
ExtractSummaryStatisticsRandomizedExp(lmeRA=resRe, N=32, alpha=0.05)
```

```
#  N  Mean     SE LowerBound UpperBound
#  1  32 0.06175 0.09508 -0.1319 0.2554
```

```
fmt
fmt
```

Description

Formatting function to set decimal precision in labels

Usage

```r
fmt()
```
getEffectSizesABBA

Author(s)
Lech Madeyski

Description
Function to calculate both effect sizes (dIG, dRM), i.e., independent groups and repeated measures standardized effect sizes and variances, for AB/BA crossover design studies. Function is used in a paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham.

Usage
getEffectSizesABBA(simulationData)

Arguments
simulationData - data set in a form required to calculate effect sizes in AB/BA crossover experimental designs

Value
data frame incl. calculated effect sizes and variances: # dIG - independent groups standardized effect size # var.dIG - variance of independent groups standardized effect size # dRM - repeated measures (within-subjects) standardized effect size # var.dRM - variance of repeated measures (within-subjects) standardized effect size calculated from t: var.dRM.Fromt=var.t*(1-r)*((N1+N2)/(2*N1*N2)) # dRM.Fromt - dRM calculated from t: var.dRM.Fromt=var.t*((N1+N2)/(2*N1*N2))+var.t*(1-r)*((N1+N2)/(2*N1*N2)) # var.dRM.Fromt2 - var.dRM calculated from t or rather dRM.Fromt: var.dRM.Fromt2=(df/(df-2))*((N1+N2)/(2*N1*N2))+dRM.Fromt^2/c^2 # var.dRM.Approx - var.dRM calculated on a basis of Johnson and Welch (1940) report an approximate formula for the variance of a t variable: var.dRM.Approx=((N1+N2)/(2*N1*N2)) + (dRM^2)/(2*(N1+N2-2)) # see paper and Equation 49 # var.dIG - independent groups standardized effect size calculated from t: var.dIG.Fromt=t*sqrt(1-r)*sqrt((N1+N2)/(2*N1*N2)) # var.dIG.Fromt - variance of independent groups standardized effect size calculated from t: var.dIG.Fromt=var.t*(1-r)*((N1+N2)/(2*N1*N2)) # dIG - dIG calculated from t: dIG.Fromt=t*sqrt((N1+N2)/(2*N1*N2)) # var.dIG - variance of independent groups standardized effect size calculated from t: var.dIG=var.t*(1-r)*((N1+N2)/(2*N1*N2)) # gRM - Hedges and Olkin (1985) unbiased estimator of the repeated measures effect size gRM=dRM*c # var.gRM - variance of gRM calculated as follows: var.gRM=(df/(df-2))*((N1+N2)/(2*N1*N2))*c^2+var.gRM^2 gRM^2/c^2 # Equation 56 # var.gRM2 - variance of gRM calculated as follows: var.gRM2=var.dRM*c^2 # gIG - Hedges and Olkin (1985) unbiased estimator of the independent groups effect size gIG=dIG*c # var.gIG - variance of gIG calculated as follows: var.gIG=(df/(df-2))*((N1+N2)/(2*N1*N2))*c^2+var.gIG^2 gIG^2/c^2 # Equation 57 #
getEffectSizesABBAIgnoringPeriodEffect

var.gIG2 - variance of gRM calculated as follows: var.gIG2=var.dIG*c^2 # r - the correlation between the values observed for the same subject

Author(s)
Lech Madeyski and Barbara Kitchenham

Examples

simulationData<-getSimulationData(25, 18.75, 50, 10, 5, 500) #generate simulated data set
es<-getEffectSizesABBA(simulationData) #return effect sizes and variances
#OR
simulationData<-getSimulationData(25, 18.75,50,10,5,15)
es<-getEffectSizesABBA(simulationData) #return effect sizes and variances

description
Function to calculate both effect sizes (dIG.ipe, dRM.ipe), i.e., independent groups and repeated measures standardized effect sizes and variances, for AB/BA crossover design studies ignoring period effect (thus wrong). Function was removed in the revision of the paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham.

Usage

gEffectSizesABBAIgnoringPeriodEffect(simulationData)

Arguments

simulationData - data set in a form required to calculate effect sizes in AB/BA crossover experimental designs

Value
data frame incl. calculated effect sizes and variances: # dIG.ipe - independent groups standardized effect size # var.dIG.ipe - variance of independent groups standardized effect size # dRM.ipe - repeated measures (within-subjects) standardized effect size # var.dRM.ipe - variance of repeated measures (within-subjects) standardized effect size # dIG.Fromt.ipe - independent groups standardized effect size calculated from t: dIG.Fromt=t*sqrt(1-r)*sqrt((N1+N2)/(2*N1*N2)) # var.dIG.Fromt.ipe - variance of independent groups standardized effect size calculated from t: var.dIG.Fromt=var.t*(1-r)*((N1+N2)/(2*N1*N2)) # dRM.ipe - dRM calculated from t: dRM.Fromt=t*sqrt((N1+N2)/(2*N1*N2)) # var.dRM.ipe - variance of dRM calculated from t: var.dRM.Fromt=var.t*((N1+N2)/(2*N1*N2)) # var.dRM.Fromt2.ipe - var.dRM calculated from t or rather dRM.Fromt: var.dRM.Fromt2=(df/(df-2))*((N1+N2)/(2*N1*N2)+dRM.Fromt^2)- dRM.Fromt^2/c^2 # unstandardized technique effect size # var.sig.ipe - sum of within-subjects variance and between-subjects variance # var.within.ipe - within-subjects variance # var.between.ipe - between-subjects variance # t.ipe - t-value # var.t.ipe - variance of t-variable
**getSimulationData**

**Author(s)**

Lech Madeyski and Barbara Kitchenham

**Examples**

```r
simulationData<-getSimulationData(25, 18.75, 50, 10, 5, 500) #generate simulated data set
es.ipe<-getEffectSizesABBAIgnoringPeriodEffect(simulationData) #return effect sizes and variances
```

**Description**

Function to generate the simulated data set used in a paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham

**Usage**

```r
getSimulationData(
  var,
  covar,
  meanA1,
  treatmentDiff,
  periodEffect,
  numOfSamples
)
```

**Arguments**

- `var`: Variance among subjects is a sum of the between subjects variance and the within subjects variance
- `covar`: Covariance equal to the between subjects variance
- `meanA1`: Mean for treatment sequence A1
- `treatmentDiff`: technique effect which is the difference between the effect of technique A and technique B
- `periodEffect`: Period effect which is the difference between period 1 and period 2
- `numOfSamples`: Number of samples ("rows" of data) required for each technique and period

**Details**

Functions related to a paper "Effect sizes and their variance for AB/BA crossover design studies" by Lech Madeyski and Barbara Kitchenham
Value

Data frame: 'data.frame': 4*numOfSamples obs. of 5 variables: $ pid : int 1 2 3 4 5 6 7 8 9 10 ... $
   technique: Factor w/ 2 levels "T1","T2": ... $ period : Factor w/ 2 levels "P1","P2": ... $ sequence : Factor w/ 2 levels "S1","S2": ... $ result : num ...

Author(s)

Lech Madeyski and Barbara Kitchenham

Examples

data<-getSimulationData(25, 18.75, 50, 10, 5, 500)  # generate the simulated data set from the paper
data<-getSimulationData(25, 18.75, 50, 10, 5, 15)

getTheoreticalEffectSizeVariancesABBA

getTheoreticalEffectSizeVariancesABBA

Description

Function provides the theoretical value of the t-statistic, variance of t, and variance of the effect sizes based on the parameters built into crossover model data simulated by the getSimulationData() function. Function is used in a paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham.

Usage

getTheoreticalEffectSizeVariancesABBA(
    theoreticalvarW,
    theoreticalTechniqueEffect,
    theoreticalrho,
    N1,
    N2
)

Arguments

theoreticalvarW - The within subject variance used to construct the simulation, i.e., the built-in Variance - the built-in Covariance
theoreticalTechniqueEffect - The technique effect built into the crossover model data
theoreticalrho - The between subject correlation built into the crossover model simulation data
N1 - The number of subjects in sequence group 1 in the crossover model simulation
N2 - The number of subjects in sequence group 2 in the crossover model simulation
Kendalltaupb

Value

data frame incl. calculated: theoretical - the theoretical value of the t-statistic theoretical
variance of t theoreticalvardIG - variance of the effect size dIG based on the parameters built into
crossover model data simulated by the getSimulationData function theoreticalvardRM - variance
of the effect size dRM based on the parameters built into crossover model data simulated by the
getSimulationData function

Author(s)

Lech Madeyski and Barbara Kitchenham

Examples

# Generates data used in Table 15 of the paper
theoreticalEffectSizeVariances <- getTheoreticalEffectSizeVariancesABBA(6.25,-10,0.75,15,15)

Description

Computes point bi-serial version of Kendall’s tau plus a 1-alpha confidence interval using the
method recommended by Long and Cliff (1997). The algorithm is based on Wilcox’s code but was
extended to return the consistent variance and the confidence intervals based on the t-distribution.
Also added a Diagnostic parameter to output internal calculations.

Usage

Kendalltaupb(x, y = NULL, alpha = 0.05)

Arguments

x 
either a matrix with two columns containg two correlated variables or a vector
of variables

y 
if y=NULL, assume x is a matrix with two columns, otherwise y is a vector of
variables with x[i] and x[i] being from the same experimental unit

alpha 
= 0.05, the Type 1 error level used for statistical tests

Value

list containing the estimate of Kendall’s tau, it hypothesis testing variance, and the t-test value
obtained from it, the significance of the t-test, the consistent variance of tau and its confidence
intervals based on both the normal distribution and the t-test (recommended by Long and Cliff)

Author(s)

Rand Wilcox, Barbara Kitchenham and Lech Madeyski
Examples

```r
x <- c(1.2, 3, 1.8, 2, 0.5, 0.5, 1, 3, 1)
y <- c(1, 1, 1, 1, 0, 0, 0, 0, 0)
Kendalltau(x, y, alpha = .05)
$cor
# [1] 0.3555556
$ci
# [1] -0.04198026 0.55555556
$cit
# [1] -0.1240567 0.55555556
$test
# [1] 1.431084
$sqse
# [1] 0.0617284
$consistentvar
# [1] 0.04113925
$siglevel
# [1] 0.1524063
```


Description


Usage

KitchenhamEtAl.CorrelationsAmongParticipants.Abrahao13TSE

Format

A data frame with 224 rows and 8 variables:

- **ExperimentID**<fct> | ExperimentID: A unique identifier for each of the five experiments in the data set.
- **ParticipantID**<fct> | Participant ID: An identifier for each participant, unique for a specific experiment.
- **SequenceGroup**<fct> | Experimental Sequence Group: A (DM-NODM,ECP-EPlat or MShop-Theatre), B (NODM-DM,ECP-EPlat or MShop-Theatre), C(DM-NODM,EPlat-ECP or Theatre-MShop), D(NODM-DM,EPlat-ECP or Theatre-MShop)
- **System**<fct> | Software systems used in the experiment: ECP an e-commerce platform from which CDs and books can be bought, EPlat a system for the management of courses, lectures and students of a university, M-Shop a system for managing sales at a music shop, Theatre a system for managing bookings for a theatre.
- **Period**<fct> | Time period of the cross-over experiment: 1 or 2
- **Treatment**<fct> | Experimental Treatment: A Dynamic Model (DM) vs No Dynamic Model (NODM)
- **Comprehension**<dbl> | Dependent variable: The comprehension level the software engineer achieved based on the F-measure
- **CrossOverID**<fct> | CrossOver category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For a 2 group crossover, the category is set to CO1 only
- **Ability**<fct> | Ability: An assessment of the ability of participants: Low, High, NA (not available)

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Abrahao13TSE
data illustrate correlations between results from individual partici-
pants in a family of 2 cross-over experiments conducted by Gravino
comprehension tasks supported by UML design models: Results from
a controlled experiment and a differentiated replication," Journal of
experiments assess whether the comprehension of object oriented
source-code increases used with UML class and sequence diagrams
produced in the software design phase. If you use this data set please
comprehension tasks supported by UML design models: Results from
a controlled experiment and a differentiated replication," Journal of
[2] Barbara Kitchenham, Lech Madeyski, Giuseppe Scanniello and
Carmine Gravino, "The importance of the Correlation between
Results from Individual Participants in Crossover Experiments" (to be
submitted as of 2020).

Description

KitchenhamEtAl.CorrelationsAmongParticipants.Gravino15JVLC data illustrate correlations between
results from individual participants in a family of 2 cross-over experiments conducted by Gravino
whether the comprehension of object oriented source-code increases used with UML class and se-
quence diagrams produced in the software design phase. If you use this data set please cite: [1] C. Gravino, G. Scanniello, and G. Tortora, "Source-code comprehension tasks supported by UML
design models: Results from a controlled experiment and a differentiated replication," Journal of
Results from Individual Participants in Crossover Experiments" (to be submitted as of 2020).

Usage

KitchenhamEtAl.CorrelationsAmongParticipants.Gravino15JVLC

Format

A data frame with 64 rows and 9 variables:

**ExperimentID**  <fct>|ExperimentID: A unique identifier for each of the three experiments in the
data set.
ParticipantID  <fct>|Participant ID: An identifier for each participant, unique for a specific experiment.
SequenceGroup  <fct>|Experimental Sequence Group: A , B , C, D
System  <fct>|Software systems used in the experiment: Music shop, a system for handling the sales of a music shop. Theater ticket, a system for managing theatre reservations.
Period  <fct>|Time period of the cross-over experiment: 1 or 2
Treatment  <fct>|Experimental Treatment: Mo, design models were available, NOMo design models were not available
Comprehension  <dbl>|Dependent variable: The level of comprehension achieved by the software engineer.
Time  <dbl>|Dependent variable: The time [min] taken to complete the comprehension task.
CrossOverID  <fct>|CrossOver category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For 2 group crossover, the category is set to CO1 only

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Gravino15JVLC


Description

Usage

`KitchenhamEtAl.CorrelationsAmongParticipants.Madeyski10`

Format

`'KitchenhamEtAl.CorrelationsAmongParticipants.Madeyski10'`: a data frame with 45 rows and 10 variables:

- **ExperimentID** `<fct>` | ExperimentID: This experiment is the only cross-over experiment in the family of TDD and Pair-Programming experiments conducted by Madeyski, so all values in this column are set to "P2007".
- **ParticipantID** `<fct>` | Participant ID: An identifier for each participant, unique for a specific experiment.
- **SequenceGroup** `<fct>` | Experimental Sequence Group: A (TLSP-TFSP), B (TFSP-TLSP)
- **System** `<fct>` | Software system to develop: Smell (a tool for identifying bad code smells in Java source code through the use of a set of software metrics) or Library (a library application)
- **Period** `<fct>` | Time period of the cross-over experiment: 1 or 2
- **Treatment** `<fct>` | Experimental Treatment: Test-First Solo Programming (TFSP) vs Test-Last Solo Programming (TLSP)
- **PATP** `<dbl>` | Dependent variable: Percentage of Acceptance Tests Passed
- **NATPPH** `<dbl>` | Dependent variable: Number of Acceptance Tests Passed Per Hour
- **CBO** `<dbl>` | Dependent variable: Mean value of Coupling Between Objects (CBO), see CK set of metrics
- **WMC** `<dbl>` | Dependent variable: Mean value of Weighted Number of Methods in Class (WMC), see CK set of metrics
- **RFC** `<dbl>` | Dependent variable: Mean value of Response For a Class (RFC), see CK set of metrics
- **CrossOverID** `<fct>` | Cross-Over Code. This experiment is a simple two-group cross-over experiment with one cross-over code, so all values in this column are set to "CO1". However, four-group experiments require a code to identify the linked sequence groups (although that can be deduced from the system used in the first time period). A crossover code is also essential for non-parametric analysis.

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

`KitchenhamEtAl.CorrelationsAmongParticipants.Madeyski10`

**Description**


**Usage**

KitchenhamEtAl.CorrelationsAmongParticipants.Reggio15SSM

**Format**

A data frame with 78 rows and 9 variables:

- **ExperimentID** <fct> | ExperimentID: A unique identifier for each of the three experiments in the data set.
- **ParticipantID** <fct> | Participant ID: An identifier for each participant, unique for a specific experiment.
- **SequenceGroup** <fct> | Experimental Sequence Group: A, B, C, D
System <fct>Software systems used in the experiment: PO, a system to process orders for an online shop. DM, a system to manage an online document review process.

Period <fct>Time period of the cross-over experiment: 1 or 2

Treatment <fct>Experimental Treatment:

Comprehension <dbl>Dependent variable: The comprehension level obtained by each participant.

Time <dbl>Dependent variable: The time [min] taken by each participant to complete the comprehension task.

CrossOverID <fct>CrossOver category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For a 2 group crossover, the category is set to CO1 only

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Reggio15SSM

KitchenhamEtAl.CorrelationsAmongParticipants.Ricca10TSE

KitchenhamEtAl.CorrelationsAmongParticipants.Ricca10TSE data illustrate correlations between results from individual participants in a family of four cross-over experiments conducted by Ricca et al.: [1] F. Ricca, M. D. Penta, M. Torchiano, P. Tonella, and M. Ceccato “How developers’ experience and ability influence web application comprehension tasks supported by uml stereotypes: A series of four experiments”, IEEE Transactions on Software Engineering, vol. 36, no. 1, pp. 96-118, 2010. Although we present the full data set, only the first two experiments were used in the correlation study, because many of the observations in the final two studies were unpaired. The experiments assess whether participants performance comprehension tasks better when using source code complemented by standard UML diagrams (UML) or by diagrams stereotyped using the Conallen notation (Conallen). If you use this data set please cite: [1] F. Ricca, M. D. Penta, M. Torchiano, P. Tonella, and M. Ceccato "How developers’ experience and ability influence web application comprehension tasks supported by uml stereotypes: A series of four experiments", IEEE Transactions on Software Engineering, vol. 36, no. 1, pp. 96—118, 2010. [2] Barbara Kitchenham, Lech Madeyski, Giuseppe Scanniello and Carmine Gravino, "The Importance of the Correlation between Results from Individual Participants in Crossover Experiments" (to be submitted as of 2020).
Description

KitchenhamEtAl.CorrelationsAmongParticipants.Ricca10TSE data illustrate correlations between results from individual participants in a family of four cross-over experiments conducted by Ricca et al.: [1] F. Ricca, M. D. Penta, M. Torchiano, P. Tonella, and M. Ceccato "How developers’ experience and ability influence web application comprehension tasks supported by uml stereotypes: A series of four experiments", IEEE Transactions on Software Engineering, vol. 36, no. 1, pp. 96-118, 2010. Although we present the full data set, only the first two experiments were used in the correlation study, because many of the observations in the final two studies were unpaired. The experiments assess whether participants performance comprehension tasks better when using source code complemented by standard UML diagrams (UML) or by diagrams stereotyped using the Conallen notation (Conallen). If you use this data set please cite: [1] F. Ricca, M. D. Penta, M. Torchiano, P. Tonella, and M. Ceccato "How developers’ experience and ability influence web application comprehension tasks supported by uml stereotypes: A series of four experiments", IEEE Transactions on Software Engineering, vol. 36, no. 1, pp. 96—118, 2010. [2] Barbara Kitchenham, Lech Madeyski, Giuseppe Scanniello and Carmine Gravino, "The Importance of the Correlation between Results from Individual Participants in Crossover Experiments" (to be submitted as of 2020).

Usage

KitchenhamEtAl.CorrelationsAmongParticipants.Ricca10TSE

Format

A data frame with 176 rows and 10 variables:

- **ExperimentID**: A unique identifier for each of the four experiments in the data set.
- **ParticipantID**: An identifier for each participant, unique for a specific experiment.
- **SequenceGroup**: Experimental Sequence Group: A, B, C, D
- **System**: Software systems used in the experiment: Two Java-based Web applications, Claros and WfMS
- **Period**: Time period of the cross-over experiment: 1 or 2
- **Treatment**: Experimental Treatment: UML or Conallon
- **FMeasure**: Dependent variable: The comprehension level achieved by the participant.
- **Time**: Dependent variable: The time [min] to complete the experimental task
- **CrossOverID**: CrossOver category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For 2 group crossover, the category is set to CO1 only
- **Ability**: h: High l: Low, NA: Not available
- **Experience**: G: Master students, U: undergraduates, P: researchers

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Ricca10TSE
**Description**


**Usage**

KitchenhamEtAl.CorrelationsAmongParticipants.Ricca14TOSEM

**Format**

A data frame with 176 rows and 10 variables:

- **ExperimentID**<fct> | ExperimentID: A unique identifier for each of the three experiments in the data set.
- **ParticipantID**<fct> | Participant ID: An identifier for each participant, unique for a specific experiment.
- **SequenceGroup**<fct> | Experimental Sequence Group: A , B , C, D
- **System**<fct> | Software systems used in the experiment: AMICO, a system for management of condominiums. EasyCoin, a system for cataloguing collections of coins.
- **Period**<fct> | Time period of the cross-over experiment: 1 or 2
- **Treatment**<fct> | Experimental Treatment: Screen mockup available (S) vs Text only (T)
### Time <dbl>
Dependent variable: The time [min] taken to perform the software engineering task.

### Comprehension <dbl>
Dependent variable: The comprehension level the software engineers.

### Efficiency <dbl>
Dependent variable: The ratio of comprehension to time.

### CrossOverID <fct>
CrossOver category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For a 2 group crossover, the category is set to CO1 only.

---

#### Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Romano18ESEM

---

**Description**

Usage

KitchenhamEtAl.CorrelationsAmongParticipants.Romano18ESEM

Format

A data frame with 194 and 10 variables:

- **ExperimentID**<fct> | ExperimentID: A unique identifier for each part of the experiment. Exp.1 identifies data from the first part of the experiment, Exp.2 identifies data from the second part of the experiment.
- **ParticipantID**<fct> | Participant ID: An identifier for each participant, unique for both parts of the experiment.
- **SequenceGroup**<fct> | Experimental Sequence Group: A, B
- **System**<fct> | Software systems used in the experiment: For the first part of the experiment, M-Shop (a system for managing a music shop) and Theater (a system for managing theatre reservations). For the second part of the experiment: AveCalc (a system that manages as electronic register and LaTazza (a system for a drinks vending machine)
- **Period**<fct> | Time period of the cross-over experiment: 1 or 2
- **Treatment**<fct> | Experimental Treatment: NOISE, participants were asked to perform a comprehension task in a noisy environment. NORMAL, participants were asked to perform a comprehension task under normal working conditions.
- **Fc**<dbl> | Dependent variable: the balanced F-measure which represents the trade-off between precision and recall, measured in the first part of the experiment.
- **Avg**<dbl> | Dependent variable: The average number of fully correct answers, measured in the first part of the experiment.
- **Ff**<dbl> | Dependent variable: Effectiveness of fault correction. Measured in the second part of the experiment.
- **CrossOverID**<fct> | CrossOver category: For 4 group crossovers, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For 2 group crossover, the category is set to CO1 only.

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Romano18ESEM

### Description


### Usage

```
KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14EASE
```

### Format

A data frame with 174 rows and 9 variables:

- **ExperimentID** <fct>: A unique identifier for each experiment in the data set.
ParticipantID <fct> | Participant ID: An identifier for each participant, unique for a specific experiment.
System <fct> | Software systems used in the experiment: Automobile: A system for controlling car behavior with use cases about entering the car, anti-lock breaking or operating the climate control of a car. ESS (Enhanced Security System) a system designed to detect potential home intruders.
Period <fct> | Time period of the cross-over experiment: 1 or 2
Treatment <fct> | Experimental Treatment: RD availability of a SysML requirements diagram vs No requirements diagram (NORD)
Time <dbl> | Dependent variable: The time [min] required for the comprehension task.
Comprehension <dbl> | Dependent variable: The comprehension level the software engineer achieved.
CrossOverID <fct> | CrossOver category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For 2 group crossover, the category is set to CO1 only

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14EASE

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14JVLC

Description


Usage

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14JVLC

Format

A data frame with 36 rows and 12 variables:

- **ExperimentID** <fct> | Experiment ID: A unique identifier for each experiment in the data set.
- **ParticipantID** <fct> | Participant ID: An identifier for each team of four participants, unique for the specific experiment.
- **SequenceGroup** <fct> | Experimental Sequence Group: A, B
- **System** <fct> | Software systems used in the experiment: Library (a software system to manage books and users of a library) and FilmCollection (a software system for the selling and the rental of films in a shop) in ExperimentStage1 and Rent (a car rental software to manage cars, customers, and reservations) and ECP (an E-Commerce Platform to order CDs and books via the Internet from an online catalogue), in ExperimentStage2.
- **Treatment** <fct> | Experimental Treatment: TPS vs F2F.
- **Period** <fct> | Time period of the cross-over experiment: 1 or 2 within each stage of the experiment
- **Time** <dbl> | Dependent variable: The total time [min] to accomplish the requirement engineering task.
- **Quality** <dbl> | Dependent variable: The quality of the requirements engineering task.
- **CrossOverID** <fct> | Crossover category: For a single 2 group crossover experiment, the value is set to CO1 for each experiment stage.
- **ExperimentPeriod** <fct> | Experiment Period: The time period across both stages of the experiment.
- **ExperimentStage** <fct> | Experiment Stage: 1 first stage, 2 second stage.

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14JVLC

**Description**


**Usage**

KitchenhamEtAl.CorrRelationsAmongParticipants.Scanniello14TOSEM

**Format**

`KitchenhamEtAl.CorrRelationsAmongParticipants.Scanniello14TOSEM`: a data frame with 172 rows and 9 variables:

- **ExperimentID**  `<fct>` | ExperimentID: A unique identifier for each experiment in the data set.
ParticipantID <fct> | Participant ID: An identifier for each participant, unique for a specific experiment.

Treatment <fct> | Experimental Treatment: AM an Analysis Model with source code (AM) vs Source Code only (SC)

SequenceGroup <fct> | Experimental Sequence Group: A (AM-SC,S1-S2), B (SC-AM,S1-S2), C(AM-SC,S2-S1), D(SC-AM,S2-S1)

System <fct> | Software systems used in the experiment: S1 A system to sell and manage CDs/DVDs in a music shop, S2 A system to book and by theater tickets.

Comprehension <dbl> | Dependent variable: The comprehension level the software engineer achieved based on the F-measure

Modification <dbl> | Dependent variable: The modifiability level the software engineer achieved based on the F-measure

Period <fct> | Time period of the cross-over experiment: 1 or 2

CrossOverID <fct> | CrossOver category: For 4 group the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For 2 group crossover, the category is set to CO1 only

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello15EMSE

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello15EMSE
data illustrate correlations between results from individual participants in cross-over experiment usb2 conducted by Scanniello et al: [1] G. Scanniello, A. Marcus, and D. Pascale, "Link analysis algorithms for static concept location: an empirical assessment", Empirical Software Engineering, vol. 20, no. 6, pp. 1666–1720, 2015. The goal of the experiment is to assess whether a new technique (implemented as an Eclipse plug-in) for static concept location (proposed by the authors) supports users in identifying the places in the code where changes are to be made.

Description


Usage

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello15EMSE
Format

‘KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello15EMSE’: a data frame with 48 rows and 10 variables:

**ExperimentID** <fct> ExperimentID: A unique identifier for each experiment in the data set.

**ParticipantID** <fct> Participant ID: An identifier for each participant, unique for a specific experiment.

**SequenceGroup** <fct> Experimental Sequence Group: A (CL-NOCL,Jedit-Atunes), B (NOCL-CL,Atunes-Jedit), C(NOCL-CL,Jedit-Atunes), D(CL-NOCL,Atunes-Jedit)

**System** <fct> Software systems used in the experiment: Jedit and Atunes

**Treatment** <fct> Experimental Treatment: Use of Concept Location plug-in (CL) vs no Concept Location plug-in (NOCL)

**Period** <fct> Time period of the cross-over experiment: 1 or 2

**Correctness** <int> Dependent variable: 0, 1, 2, 3, 4. The participants are asked to indicate a single change method for each of 4 bug reports. A change method is correctly identified if that method is in the change set of the bug report.

**Time** <dbl> Dependent variable: The total time [min] to accomplish concept location tasks, i.e., to identify (four) bugs given their reports

**Efficiency** <dbl> Dependent variable: The participants’ efficiency in the execution of concept location tasks. It is computed dividing correctness by time.

**CrossOverID** <fct> Crossover category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2.

Examples

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello15EMSE
KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello17TOSEM
data illustrate correlations between results from individual partic-
ips in a family of four cross-over experiments conducted by
Romano, “Fixing faults in C and Java source code: Abbreviated vs.
full-word identifier names”, ACM Transactions on Software Engi-
neering Methodology, vol. 26, no. 2, 2017. The experiments assess
whether whether the use of abbreviated identifier names (ABBR),
impacts the effectiveness of fault fixing in C and Java source code in
comparison with full-word identifier names (FULL). If you use this
data set please cite: [1] G. Scanniello, M. Risi, P. Tramontana, and
S. Romano, “Fixing faults in C and Java source code: Abbreviated
vs. full-word identifier names”, ACM Transactions on Software Engi-
Lech Madeyski, Giuseppe Scanniello and Carmine Gravino, “On
the Importance of the Correlation between Results from Individual
Participants in Crossover Experiments” (to be submitted as of 2020).

Description

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello17TOSEM data illustrate correlations
between results from individual participants in a family of four cross-over experiments conducted
and Java source code: Abbreviated vs. full-word identifier names”, ACM Transactions on Software
Engineering Methodology, vol. 26, no. 2, 2017. The experiments assess whether whether the use of
abbreviated identifier names (ABBR), impacts the effectiveness of fault fixing in C and Java source
code in comparison with full-word identifier names (FULL). If you use this data set please cite:
code: Abbreviated vs. full-word identifier names”, ACM Transactions on Software Engineering
and Carmine Gravino, "On the Importance of the Correlation between Results from Individual
Participants in Crossover Experiments" (to be submitted as of 2020).

Usage

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello17TOSEM

Format

A data frame with 200 rows and 17 variables:

**ExperimentID** <fct> | ExperimentID: A unique identifier for each of the experiments in the data
set.

**ParticipantID** <fct> | Participant ID: An identifier for each participant, unique for a specific exper-
iment.

**SequenceGroup** <fct> | Experimental Sequence Group: A , B , C , D
Software systems used in the experiments: The Unibas experiment used Agenda (a system for tracking personal contacts) and Gas-Station (a system for managing a petrol station). The UniNa experiment used Financial (a system which is a command line option price calculator) and Hotel-Reservation. The POLINA and PROF experiments used AveCalc (a system that manages as electronic register and LaTazza (a system for a drinks vending machine).

**Period**<fct>\(\text{Time period of the cross-over experiment: 1 or 2}\)

**Treatment**<fct>\(\text{Experimental Treatment: ABBR, abbreviated names. FULL, full names}\)

**Time**<dbl>\(\text{Dependent variable: The time each participant spent performing the SE task.}\)

**FMeasure**<dbl>\(\text{Dependent variable: The effectiveness of the participants taking into account correctness and completeness of the fault fixing tasks}\)

**Efficiency**<dbl>\(\text{Dependent variable: The ratio of effectiveness to time.}\)

**CrossOverID**<fct>\(\text{CrossOver category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For 2 group crossover, the category is set to CO1 only.}\)

**Language**<fct>\(\text{Java or C. The language was the same for all participants in a specific experiment. POLINA and PROF used Java, UNIBAS and UNINA used C.}\)

**Ident**<dbl>\(\text{Dependent variable: The number of faults identified.}\)

**Fixed**<dbl>\(\text{Dependent variable: The number of faults identified.}\)

**WrongIdent**<dbl>\(\text{Dependent variable: The number of faults incorrectly identified}\)

**WronglyFixed**<dbl>\(\text{Dependent variable: The number of faults incorrectly fixed.}\)

**precision**<dbl>\(\text{Dependent variable: The ratio of number of faults correctly fixed to the number of faults correctly identified.}\)

**recall**<dbl>\(\text{Dependent variable: The ratio of number of faults correctly fixed to the total number of fault.}\)

**Examples**

KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello17TOSEM
KitchenhamEtAl.CorrelationsAmongParticipants.Torchiano17JVLC


Format

A data frame with 214 rows and 8 variables:

**ExperimentID** <fct> | ExperimentID: A unique identifier for each of the three experiments in the data set.
ParticipantID <fct> | Participant ID: An identifier for each participant, unique for a specific experiment.
SequenceGroup <fct> | Experimental Sequence Group: A, B, C, D
System <fct> | Software systems used in the experiment: File System manager (FS) for folders, files, links. Roads system (R) handles maps made up of cities connected by means of roads. Train (T) a system to manage timetables, trains, and paths. Catalogue system (C). It collects category of items (e.g., cars) and items (e.g., car models) based on a set of features (e.g., number of doors). In PoliTo2, only FS and T were administered to the participants, while in UniBas1 and UniGe1 all the four experimental objects were used.
Period <fct> | Time period of the cross-over experiment: 1 or 2
Treatment <fct> | Experimental Treatment: Object Diagram (OD) vs No Object Diagram (NoOD)
Comprehension <dbl> | Dependent variable: The comprehension level the software engineers. For PoliTo2 Comprehension was based on answering a set of 4 questions, for UniBas and UniGe comprehension was measured using the F metric.
CrossOverID <fct> | CrossOver category: For 4 group crossover designs, the crossover category specifies the matching pairs of sequence groups, CO1 and CO2. For 2 group crossover, the category is set to CO1 only

Examples
KitchenhamEtAl.CorrelationsAmongParticipants.Torchiano17JVLC

KitchenhamMadeyski.SimulatedCrossoverDataSets

Description

Usage
KitchenhamMadeyski.SimulatedCrossoverDataSets

Format
A data frame with variables:
actualSampleSize Sample size
SSFull Sample Size
CFull Correlation
ESFull Effect Size
Details

This is simulated normally distributed data from 30 subjects, with technique A being 10 units more effective than technique B, and there is a period effect equaling 5 units. Subject 1 to 15 used technique B first while subjects 16 to 30 used technique A first.

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Usage

KitchenhamMadeyskiBrereton.ABBAMetaAnalysisReportedResults

Format

A text file with variables:

- **Study** This field includes the study identifier of each of the 3 primary studies which reported results per document.
- **Type** This identifies the type of effect size used by the study authors. d or g refer to d_IG and g_IG, P is the aggregated p values, if the repeated measures (RM) estimate was obtained it is appropriately specified.
Source  Always set to Rep. This identifies that the data was as reported by the primary study authors.

mean  The overall mean effect size reported by the study authors

pvalue  The one-sided p-value associated with the overall mean reported by the study authors. NA means the authors did not report this statistic.

UB  The upper bound of the confidence interval of the overall mean as reported by the primary study authors. NA means the authors did not report this statistic.

LB  The lower bound of the confidence interval of the overall mean as reported by the primary study authors. NA means the authors did not report this statistic.

QE  The heterogeneity statistic associated with the meta-analysis as reported by the study authors. NA means the authors did not report this statistic.

Qep  The p-value of the heterogeneity statistic associated with the meta-analysis as reported by the study authors. NA means the authors did not report this statistic.

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

KitchenhamMadeyskiBrereton.ABBAReportedEffectSizes

KitchenhamMadeyskiBrereton.ABBAReportedEffectSizes data

Description

This data is used in the paper: Barbara Kitchenham, Lech Madeyski and Pearl Brereton. Meta-analysis for Families of Experiments: A Systematic Review and Reproducibility Assessment, Empirical Software Engineering (2019) doi:10.1007/s10664-019-09747-0. This file holds the individual effect sizes for the first time period (or equivalently the first document), as reported by the 3 primary studies in the systematic review that reported results for each document/time period separately.

Usage

KitchenhamMadeyskiBrereton.ABBAReportedEffectSizes
Format

A text file with variables:

Study This field includes the study identifier of each of the 3 primary studies which were included in the systematic review. The studies are S3, S7 and S11.

Type This identifies the type of effect size used by the study authors. d or g refer to dIG and gIG.

Source Always set to Rep. This identifies that the data was as reported by the primary study authors.

Design Mixed means different experiments in a particular family used different methods (only S3 used mixed methods and 4 experiments used the 4 group crossover and one used an independent groups design). ABBACO is the standard 2-group crossover design.

Exp1 This is the reported standardised effect size for the first time period and the first experiment in the family.

Exp2 This is the reported standardised effect size for the first time period and second experiment in the family.

Exp3 This is the reported standardised effect size for the first time period and the third experiment in the family.

Exp4 This is the reported standardised effect size for the first time period and the fourth experiment in the family. NA means there was no fourth experiment in the family.

Exp5 This is the reported standardised effect size for the first time period and the fifth experiment in the family. NA means there was no fifth experiment in the family.

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

KitchenhamMadeyskiBrereton.ABBAReportedEffectSizes

Description

This data is used in the paper: Barbara Kitchenham, Lech Madeyski and Pearl Brereton. Meta-analysis for Families of Experiments: A Systematic Review and Reproducibility Assessment, Empirical Software Engineering (2019) doi:10.1007/s10664-019-09747-0. This file holds the descriptive data for each document and each experiment for studies 3, 7 and 11 which include the mean, standard deviation and sample size for the control and treatment techniques. These studies performed ABBA crossover experiments and reported data for each document separately. Note Study 3 also undertook an independent groups study but data from that experiment is held in the ExpData file.
Usage

KitchenhamMadeyskiBrereton.DocData

Format

A text file with variables:

**Study**  This field includes the study identifier of each of the 3 primary studies which reported their basic statistics on a time period & document basis.

**Exp**  This identifies the experiment to which the descriptive data belongs.

**Doc**  This identifies whether the data arose from the document used in the first or second time period. The value "Doc1" identifies the data as coming from the first document or first time period. The value "Doc2" identifies the data as coming from the second time period or document. Note for Study 3 we used the analysis of a specific document that was used in all 4 ABBA experiments. For studies 7 and 11, the authors identified which we used in each time period and Doc1 refers to data from the first time period.

**Mc**  The mean value of the observations obtained using the control technique for the identified document.

**SDc**  The standard deviation of the observations obtained using the control technique for the identified document.

**Nc**  The number of participants using the control technique in the first time period for the identified document.

**Mt**  The mean value of the observations obtained using the treatment technique for the identified document.

**SDt**  The standard deviation of the observations obtained using the treatment technique for the identified document.

**Nt**  The number of participants using the treatment technique in the first time period for the identified document.

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

KitchenhamMadeyskiBrereton.DocData
**Description**

This data is used in the paper: Barbara Kitchenham, Lech Madeyski and Pearl Brereton. Meta-analysis for Families of Experiments: A Systematic Review and Reproducibility Assessment, Empirical Software Engineering (2019) doi:10.1007/s10664-019-09747-0. This file holds the descriptive data for each experiment which include the mean, standard deviation and sample size for the control and treatment techniques. Note in the case of studies 3, 7 and 11, which reported descriptive data for each time period (or equivalently each document) separately, the values for of the descriptive data were obtained by analysing the data reported in the DocData file.

**Usage**

KitchenhamMadeyskiBrereton.ExpData

**Format**

A text file with variables:

- **Study**  This field includes the study identifier of each of the 13 primary studies which were included in the systematic review.
- **Exp**  This identifies the experiment to which the descriptive data belongs.
- **Source**  Always set to Rep. This identifies that the data was as reported by the primary study authors.
- **Mc**  The mean value of the observations obtained using the control technique.
- **SDc**  The standard deviation of the observations obtained using the control technique.
- **Nc**  The number of participants using the control technique in the first time period.
- **Mt**  The mean value of the observations obtained using the treatment technique.
- **SDt**  The standard deviation of the observations obtained using the treatment technique.
- **Nt**  The number of participants using the treatment technique in the first time period.
- **r**  The correlation between repeated measures. NA if not reported. Note only study 13 reported this correlation.

**Source**

https://madeyski.e-informatyka.pl/reproducible-research/

**Examples**

KitchenhamMadeyskiBrereton.ExpData
KitchenhamMadeyskiBrereton.MetaAnalysisReportedResults data

Description

This data is used in the paper: Barbara Kitchenham, Lech Madeyski and Pearl Brereton. Meta-analysis for Families of Experiments: A Systematic Review and Reproducibility Assessment (to be submitted). This data set reports the meta-analysis results reported by the authors of the 13 primary studies included in the systematic review.

Usage

KitchenhamMadeyskiBrereton.MetaAnalysisReportedResults

Format

A text file file with variables:

**Study** This field includes the study identifier of each of the 13 primary studies which were included in the systematic review.

**Type** This identifies the type of effect size used by the study authors. d or g refer to d_IG and g_IG, P is the aggregated p values, if the repeated measures estimate was obtained it is appropriately specified, r refers to the point bi-serial correlation.

**Source** Always set to Rep. This identifies that the data was as reported by the primary study authors.

**mean** The overall mean effect size reported by the study authors

**pvalue** The one-sided p-value associated with the overall mean reported by the study authors. NA means the authors did not report this statistic.

**UB** The upper bound of the confidence interval of the overall mean as reported by the primary study authors. NA means the authors did not report this statistic.

**LB** The lower bound of the confidence interval of the overall mean as reported by the primary study authors. NA means the authors did not report this statistic.

**QE** The heterogeneity statistic associated with the meta-analysis as reported by the study authors. NA means the authors did not report this statistic.

**Qep** The p-value of the heterogeneity statistic associated with the meta-analysis as reported by the study authors. NA means the authors did not report this statistic.

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

KitchenhamMadeyskiBrereton.MetaAnalysisReportedResults
Description

This data is used in the paper: Barbara Kitchenham, Lech Madeyski and Pearl Brereton. Meta-analysis for Families of Experiments: A Systematic Review and Reproducibility Assessment, Empirical Software Engineering (2019) doi:10.1007/s10664-019-09747-0. This file holds the individual effect sizes for each experiment, as reported by 13 primary studies in the systematic review.

Usage

KitchenhamMadeyskiBrereton.ReportedEffectSizes

Format

A text file with variables:

**Study**  This field includes the study identifier of each of the 13 primary studies which were included in the systematic review.

**Type**  This identifies the type of effect size used by the study authors. d or g refer to dIG and gIG, p is the p-value used for aggregation, if the repeated measures estimate was obtained it is appropriately specified as gRM, r refers to the point bi-serial correlation.

**Source**  Always set to Rep. This identifies that the data was as reported by the primary study authors.

**Design**  The refers to the design method used by the study author. 4GroupCO is a 4-group crossover design. Mixed means different experiments in a particular family used different methods (only S3 used mixed methods and 4 experiments used the 4 group crossover and one used an independent groups design). ABBACO is the standard 2-group crossover design. IndGroups is the independent groups design also called between groups design or a randomised design. PrePost is pretest and posttest design with a post test control.

**Exp1**  This is the reported standardised effect size for the first experiment in the family.

**Exp2**  This is the reported standardised effect size for the second experiment in the family.

**Exp3**  This is the reported standardised effect size for the third experiment in the family.

**Exp4**  This is the reported standardised effect size for the fourth experiment in the family. NA means there was no fourth experiment in the family.

**Exp5**  This is the reported standardised effect size for the fifth experiment in the family. NA means there was no fifth experiment in the family.

Source

https://madeyski.e-informatyka.pl/reproducible-research/
Examples

KitchenhamMadeyskiBrereton.ReportedEffectSizes

KitchenhamMadeyskiBudgen16.COCOMO

KitchenhamMadeyskiBudgen16.COCOMO data

Description


Usage

KitchenhamMadeyskiBudgen16.COCOMO

Format

A data frame with variables:

- **Project**: Project ID
- **Type**: A categorical variable describing the type of the project
- **Year**: The year the project was completed
- **Lang**: A categorical variable describing the development language used
- **Rely**: Ordinal value defining the required software reliability
- **Data**: Ordinal value defining the data complexity / Data base size
- **Cplx**: Ordinal value defining the complexity of the software / Process complexity
- **Aaf**: ??
- **Time**: Ordinal value defining the stringency of timing constraints / Time constraint for cpu
- **Stor**: Ordinal value defining the stringency of the data storage requirements / Main memory constraint
- **Virt**: Virtual Machine volatility
- **Turn**: Turnaround time
- **Type2**: A categorical variable defining the hardware type: mini, max=mainframe, midi
- **Acap**: Ordinal value defining the analyst capability
- **Aexp**: Ordinal value defining the analyst experience / application experience
- **Pcap**: Ordinal value defining the programming capability of the team / Programmers capability
- **Vexp**: Ordinal value defining the virtual machine experience of the team
Lexp Ordinal value defining the programming language experience of the team

Cont ??

Modp / Modern programming practices

Tool Ordinal value defining the extent of tool use / Use of software tools

ToolCat Recoding of Tool to labelled ordinal scale

Sced Ordinal value defining the stringency of the schedule requirements / Schedule constraint

Rvol Ordinal value defining the requirements volatility of the project

Select Categorical value calculated by BAK for an analysis example

Rvolcat Recoding of Rvol to a labelled ordinal scale

Modecat Mode of the projects: O=Organic, E=Embedded, SD-Semi-Detached

Mode1 Dummy variable calculated by BAK: 1 if the project is Organic, 0 otherwise

Mode2 Dummy variable calculated by BAK: 1 if the project is Semi-detached, 0 otherwise

Mode3 Dummy variable calculated by BAK: 1 if the project is Embedded, 0 otherwise

KDSI Product Size Thousand of Source Instructions

AKDSI Adjusted Product Size for Project in Thousand Source Instructions - differs from KDSI for enhancement projects

Effort Project Effort in Man months

Duration Duration in months

Productivity Productivity of project calculated by BAK as AKDSI/Effort, so the the larger the value the better the productivity

Details


Explanations by Barbara Kitchenham / https://terapromise.csc.ncsu.edu:8443/#!/repo/view/head/effort/cocomo/cocomo1/nasa93/nasa93.arff

COCOMO.txt: pro type year Lang Rely Data CPLX aaf time store virt turn type2 acap aexp pcap vexp lexp cont modp TOOL TOOLcat SCED RVOL Select rvolcat Modecat Mode1 Mode2 Mode3 KDSI AKDSI Effort Dur Productivity

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

KitchenhamMadeyskiBudgen16.COCOMO
Description


Usage

KitchenhamMadeyskiBudgen16.DiffInDiffData

Format

A data frame with variables:

- **Abstract**  The abstract identifier
- **Site**  A numeric identifier of the site
- **Treatment**  A three character alphanumeric identifying the journal and time period of the abstract
- **Journal**  The journal in which the abstract was published: IST or JSS
- **Timeperiod**  The time period in which the abstract: 1 or 2
- **J1**  The identifier for the judge who made the next 2 assessments
- **J1Completeness**  The average completeness made by judge J1 based on the 8 completeness questions
- **J1Clarity**  The clarity assessment made by judge J1
- **J2**  The identifier for the judge who made the next 2 assessments
- **J2Completeness**  The average completeness made by judge J2 based on the 8 completeness questions
- **J2Clarity**  The clarity assessment made by judge J2
- **J3**  The identifier for the judge who made the next 2 assessments
- **J3Completeness**  The average completeness made by judge J3 based on the 8 completeness questions
- **J3Clarity**  The clarity assessment made by judge J3
- **J4**  The identifier for the judge who made the next 2 assessments
- **J4Completeness**  The average completeness made by judge J4 based on the 8 completeness questions
- **J4Clarity**  The clarity assessment made by judge J4
- **MeanCompleteness**  The mean of J1Completeness, J2Completeness, J3Completeness, J4Completeness
**MedianCompleteness**  The median of J1Completeness, J2Completeness, J3Completeness, J4Completeness

**MedianClarity**  The median clarity of J1Clarity, J2Clarity, J3Clarity, J4Clarity

**MeanClarity**  The mean clarity of J1Clarity, J2Clarity, J3Clarity, J4Clarity

**VarCompleteness**  The variance of J1Completeness, J2Completeness, J3Completeness, J4Completeness

**VarClarity**  The variance clarity of J1Clarity, J2Clarity, J3Clarity, J4Clarity

**Details**

Data set was derived from the data reported in the SubjectData data set (subjectdata.txt). It contains the summary completeness and clarity data from 4 judges who assessed the same abstract. Only the initial 5 sites are included.

dinddata.txt

**Source**

[https://madeyski.e-informatyka.pl/reproducible-research/](https://madeyski.e-informatyka.pl/reproducible-research/)

**Examples**

KitchenhamMadeyskiBudgen16.DiffInDiffData

---

**Description**


**Usage**

KitchenhamMadeyskiBudgen16.FINNISH

**Format**

A data frame with variables:

- **Project**  Project ID
- **DevEffort**  Development Effort measured in hours
- **UserEffort**  Effort provided by the customer/user organisation measured in hours
- **Duration**  Project duration measured in months
**HWType** A categorical variable defining the hardware type

**AppType** A categorical variable defining the application type

**FP** Function Points measured using the TIEKE organisation method

**Co** A categorical variable defining the company

**Details**

Data set collected from 9 Finish companies by Mr Hanna Mäki from the TIEKE organisation see Barbara Kitchenham and Kari Kansala, Inter-item correlations among function points, Proceedings ICSE 15, 1983, pp 477-480

**Source**

https://madeyski.e-informatyka.pl/reproducible-research/

**Examples**

KitchenhamMadeyskiBudgen16.FINNISH

---

**Description**


**Usage**

KitchenhamMadeyskiBudgen16.PolishData

**Format**

A data frame with variables:

- **Abstract** The abstract identifier
- **Site** Numeric identifier for the site
- **Treatment** The first three characters of the Abstract field which identifies the journal and time period of the abstract
- **Journal** An acronym for the journal from which the abstract was obtained: IST or JSS
- **Timeperiod** The Time period in which the abstract was found: 1 or 2
J1 The identifier for the judge who made the next 2 assessments
J1Completeness The average completeness made by judge J1 based on the 8 completeness questions
J1Clarity The clarity assessment made by judge J1
J2 The identifier for the judge who made the next 2 assessments
J2Completeness The average completeness made by judge J2 based on the 8 completeness questions
J2Clarity The clarity assessment made by judge J2
J3 The identifier for the judge who made the next 2 assessments
J3Completeness The average completeness made by judge J3 based on the 8 completeness questions
J3Clarity The clarity assessment made by judge J3
J4 The identifier for the judge who made the next 2 assessments
J4Completeness The average completeness made by judge J4 based on the 8 completeness questions
J4Clarity The clarity assessment made by judge J4
MedianCompleteness The median of J1Completeness, J2Completeness, J3Completeness, J4Completeness
MedianClarity The median of J1Clarity, J2Clarity, J3Clarity, J4Clarity

Details
Data set derived from PolishSubjects data set collected at Wroclaw University. It summarizes the completeness and clarity data collected from 4 judges about the same abstract.

PolishData.txt

Source
https://madeyski.e-informatyka.pl/reproducible-research/

Examples
KitchenhamMadeyskiBudgen16.PolishData

Description
Usage

KitchenhamMadeyskiBudgen16.PolishSubjects

Format

A data frame with variables:

Judge The identifier for each subject

Abstract The identifier for each abstract - the code starts with a three alphanumeric string that defines the source of the abstract

OrderViewed Each judge assessed 4 abstracts in sequence, this data item identifies the order in which the subject viewed the specified abstract

Completness1 Assessment by judge of question 1: Is the reason for the project clear? Can take values: Yes/No/Partly

Completness2 Assessment by judge of question 2: Is the specific aim/purpose of the study clear? Can take values: Yes/No/Partly

Completness3 Assessment by judge of question 3: If the aim is to describe a new or enhanced software technology (e.g. method, tool, procedure or process) is the method used to develop this technology defined? Can take values: Yes/No/Partly/NA

Completness4 Assessment by judge of question 4: Is the form (e.g. experiment, general empirical study, data mining, case study, survey, simulation etc.) that was used to evaluate the technology made clear? Can take values: Yes/No/Partly

Completness5 Assessment by judge of question 5: Is there a description of how the evaluation process was organised? Can take values: Yes/No/Partly

Completness6 Assessment by judge of question 6: Are the results of the evaluation clearly described? Can take values: Yes/No/Partly

Completness7 Assessment by judge of question 7: Are any limitations of the study reported?: Yes/No/Partly

Completness8 Assessment by judge of question 8: Are any ideas for future research presented?: Yes/No/Partly

Clarity Assessment by judge of question regarding the overall understandability of the abstract: Please give an assessment of the clarity of this abstract by circling a number on the scale of 1-10 below, where a value of 1 represents Very Obscure and 10 represents Extremely Clearly Written.

Completness1NumValue A numerical value for completeness question 1 where 0=No, Partly=0.5, yes =1

Completness2NumValue A numerical value for completeness question 2 where 0=No, Partly=0.5, yes =1, NA means not applicable

Completness3NumValue A numerical value for completeness question 3 where 0=No, Partly=0.5, yes =1, NA means not applicable or not answered

Completness4NumValue A numerical value for completeness question 4 where 0=No, Partly=0.5, yes =1, NA means not applicable

Completness5NumValue A numerical value for completeness question 5 where 0=No, Partly=0.5, yes =1, NA means not applicable
**Completeness6NumValue**  A numerical value for completeness question 6 where 0=No, Partly=0.5, yes =1, NA means not applicable

**Completeness7NumValue**  A numerical value for completeness question 7 where 0=No, Partly=0.5, yes =1, NA means not applicable

**Completeness8NumValue**  A numerical value for completeness question 8 where 0=No, Partly=0.5, yes =1, NA means not applicable

**Sum**  The sum of the numerical completeness questions excluding those labelled NA

**TotalQuestions**  The count of the number of question related to completeness excluding questions considered not applicable

**Completeness**  Sum/TotalQuestions

### Details

Data set collected at Wroclaw University of Technology (POLAND) by Lech Madeyski includes separate entries for each abstract assessed by a judge, that is 4 entries for each judge. Data collected from 16 subjects recruited from Wroclaw University of Technology who were each asked to assess 4 abstracts.

Note Only completeness question 2 was expected to be context dependent and have a NA (not applicable) answer, if other completeness answers were left blank, BAK coded the answer as NA polishsubjects.txt

### Source

https://madeyski.e-informatyka.pl/reproducible-research/

## Examples

KitchenhamMadeyskiBudgen16.PolishSubjects

---

**Description**


### Usage

KitchenhamMadeyskiBudgen16.SubjectData
### Format

A data frame with variables:

- **Judge**: Alphanumeric identifier for each judge
- **Institution**: Numerical value identifying each site from which data was collected
- **JudgeID**: Numerical value identifying each judge
- **Age**: Age of the judge in years
- **Eng1st**: Whether the judge’s first language was English: Yes/No
- **YearsStudy**: The number of years have student been studying computing at University: 1, 2, 3, 4
- **AbstractsRead**: Number of abstracts the judge had read prior to the study: 0, 1 to 10, 10+
- **AbstractsWritten**: Whether the judge had ever written an abstract for a scientific report/article
- **AbstractID**: Alphanumeric identifier for an abstract. The first character identifies the journal, I=IST, J=JSS, the third digit identifies the time period as 1 or 2, the remaining digits identify the abstract number within the set of abstracts found for the specified journal and time period
- **Treat**: The initial 3 characters of AbstractID
- **TreatID**: A numeric identifier for the journal and time period, 1=IB1, 2=IB2, 3=JB1, 4=JB2
- **Order**: The order in which the judge should have viewed the specified abstract
- **Completness1NumValue**: The numeric answer to completeness question 1
- **Completness2NumValue**: The numeric answer to completeness question 2
- **Completness3NumValue**: The numeric answer to completeness question 3
- **Completness4NumValue**: The numeric answer to completeness question 4
- **Completness5NumValue**: The numeric answer to completeness question 5
- **Completness6NumValue**: The numeric answer to completeness question 6
- **Completness7NumValue**: The numeric answer to completeness question 7
- **Completness8NumValue**: The numeric answer to completeness question 8
- **Clarity**: The response to the clarity question or NA if not answered
- **NumberOfAnsweredCompletnessQuestions**: The number of completeness questions excluding those with NA
- **TotalScore**: Sum of the numeric values of the 8 completeness questions
- **MeanScore**: Sum of the completeness questions 1 to 8 divided by TotalScore
- **Site**: The name of the site which provided the data. Hong Kong refers to the Polytechnic University, Hong Kong 2 refers to the City University

### Details

Data set collected from 16 judges assessing 4 abstracts at 6 sites: Lincoln University NZ=1, Hong Kong Polytechnic University=2, PSu Thailand=3, Durham=4, Keele=5, Hong Kong City University=6

subjectdata.txt: Judge Institution JudgeID age eng1st years.study abs.read Absid Treat TreatID Order Com.1 Com.2 Com.3 Com.4 Com.5 Com.6 Com.7 Com.8 Clarity num.questions total.score av.score Site
Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

KitchenhamMadeyskiBudgen16.SubjectData

Description

Returns a sample of N observations from a Laplace distribution with specified mean and spread.

Usage

LaplaceDist(N, mean, spread, max = 0.5, min = -0.5)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>is the required sample size</td>
</tr>
<tr>
<td>mean</td>
<td>is the required mean</td>
</tr>
<tr>
<td>spread</td>
<td>is the spread of the function</td>
</tr>
<tr>
<td>max</td>
<td>upper limits of the distribution, must be finite</td>
</tr>
<tr>
<td>min</td>
<td>lower limit of the distribution, must be finite</td>
</tr>
</tbody>
</table>

Value

N values from a Laplace distribution

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

```r
set.seed(123)
LaplaceDist(10,0,1)
# [1] -0.55311564 0.85946218 -0.20094937 1.45258293 2.12808209 -2.39565480 0.05785263...
```
Madeyski15EISEJ.OpenProjects

Madeyski15EISEJ.OpenProjects data

Description


Usage

Madeyski15EISEJ.OpenProjects

Format

A data frame with variables:

- **PROP** The percentage of classes of proprietary (i.e., industrial) projects that must be tested in order to find 80% of defects in case of software defect prediction models built on open source projects.
- **NOTOPEN** The percentage of classes of projects which are not open source projects that must be tested in order to find 80% of defects in case of software defect prediction models built on open source projects.
- **STUD** The percentage of classes of student (i.e., academic) projects that must be tested in order to find 80% of defects in case of software defect prediction models built on open source projects.
- **OPEN** The percentage of classes of open source projects that must be tested in order to find 80% of defects in case of software defect prediction models built on open source projects.

Details

This paper presents an analysis of 84 versions of industrial, open-source and academic projects. We have empirically evaluated whether those project types constitute separate classes of projects with regard to defect prediction. The predictions obtained from the models trained on the data from the open source projects were compared with the predictions from the other models (built on proprietary, i.e. industrial, student, open source, and not open source projects).

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

Madeyski15EISEJ.OpenProjects
Madeyski15EISEJ.PropProjects

Madeyski15EISEJ.PropProjects data

Description


Usage

Madeyski15EISEJ.PropProjects

Format

A data frame with variables:

NOTPROP The percentage of classes of non-proprietary (i.e., non-industrial) projects that must be tested in order to find 80% of defects in case of software defect prediction models built on proprietary (i.e., industrial) projects.

OPEN The percentage of classes of open source projects that must be tested in order to find 80% of defects in case of software defect prediction models built on proprietary (i.e., industrial) projects.

STUD The percentage of classes of student (i.e., academic) projects that must be tested in order to find 80% of defects in case of software defect prediction models built on proprietary (i.e., industrial) projects.

PROP The percentage of classes of proprietary (i.e., industrial) projects that must be tested in order to find 80% of defects in case of software defect prediction models built on proprietary (i.e., industrial) projects.

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

Madeyski15EISEJ.PropProjects

A data frame with variables:

**PROP** The percentage of classes of proprietary (i.e., industrial) projects that must be tested in order to find 80% of defects in case of software defect prediction models built on student (i.e., academic) projects.

**NOTSTUD** The percentage of classes of projects which are not student projects that must be tested in order to find 80% of defects in case of software defect prediction models built on student (i.e., academic) projects.

**STUD** The percentage of classes of student (i.e., academic) projects that must be tested in order to find 80% of defects in case of software defect prediction models built on student (i.e., academic) projects.

**OPEN** The percentage of classes of open source projects that must be tested in order to find 80% of defects in case of software defect prediction models built on student (i.e., academic) projects.

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

Madeyski15EISEJ.StudProjects
Description


Usage

Madeyski15SQJ.NDC

Format

A data frame with variables:

**Project** In case of open source projects this field includes the name of the project as well as its version. In case of industrial projects this field includes the string "proprietary" (we were not allowed to disclose the names of the analyzed industrial software projects developed by Capgemini Polska).

**simple** The percentage of classes that must be tested in order to find 80% of defects in case of simple defect prediction models, i.e., using only software product metrics as predictors.

**advanced** The percentage of classes that must be tested in order to find 80% of defects in case of advanced defect prediction models, using not only software product metrics but also the NDC (Number of distinct committers) process metric.

Details

"This paper presents an empirical evaluation in which several process metrics were investigated in order to identify the ones which significantly improve the defect prediction models based on product metrics. Data from a wide range of software projects (both, industrial and open source) were collected. The predictions of the models that use only product metrics (simple models) were compared with the predictions of the models which used product metrics, as well as one of the process metrics under scrutiny (advanced models). To decide whether the improvements were significant or not, statistical tests were performed and effect sizes were calculated. The advanced defect prediction models trained on a data set containing product metrics and additionally Number of Distinct Committers (NDC) were significantly better than the simple models without NDC, while the effect size was medium and the probability of superiority (PS) of the advanced models over simple ones was high (p=.016, r=-.29, PS=.76), which is a substantial finding useful in defect prediction. A similar result with slightly smaller PS was achieved by the advanced models trained on a data set containing product metrics and additionally all of the investigated process metrics (p=.038, r=-.29, PS=.68). The advanced models trained on a data set containing product metrics and additionally Number of Modified Lines (NML) were significantly better than the simple models without NML, but the effect size was small (p=.038, r=.06). Hence, it is reasonable to recommend the NDC process metric in building the defect prediction models." [https://dx.doi.org/10.1007/s11219-014-9241-7]
Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

Madeyski15SQJ.NDC

Description

If you use this data set please cite this R package and the paper where we analyze the data set: Lech Madeyski and Barbara Kitchenham, "Effect Sizes and their Variance for AB/BA Crossover Design Studies". Empirical Software Engineering, vol. 24, no.4, p. 1982-2017, 2018. DOI: 10.1007/s10664-017-9574-5

Usage

MadeyskiKitchenham.EUBASdata

Format

A data frame with variables:

- **ID**  Project ID
- **TimePeriod**  Period of time (run): R1, R2
- **SequenceGroup**  Sequence group: G1, G2, G3, G4
- **System**  Software system identifier indicates the system (i.e., S1 or S2) used as the experimental object: S1. A software system to sell and manage CDs/DVDs in a music shop, S2. A software system to book and buy theater tickets
- **Technique**  The independent variable. It is a nominal variable that can assume the following two values: AM (analysis models plus source code) and SC (source code alone)
- **Comp_Level**  This denotes the comprehension level of the source code achieved by a software engineer
- **Modi_Level**  This denotes the capability of a maintainer to modify source code
Details
Data set comes from an experiment conducted in Italy at the University of Basilicata (with 24 first-year students from the Master’s Program in Computer Science) to answer the question "Do the software models produced in the requirements analysis process aid in the comprehensibility and modifiability of source code?", see G. Scanniello, C. Gravino, M. Genero, J. A. Cruz-Lemus, and G. Tortora, “On the Impact of UML Analysis Models on Source-code Comprehensibility and Modifiability,” ACM Transactions on Software Engineering and Methodology, vol. 23, pp. 13:1-13:26, Apr. 2014. However, the inconsistent subject data for subject 2 was removed, see the aforementioned paper by Madeyski and Kitchenham.

Source
https://madeyski.e-informatyka.pl/reproducible-research/

Examples
MadeyskiKitchenham.EUBASdata

Description
If you use this data set please cite: Lech Madeyski and Barbara Kitchenham, "How variations in experimental designs impact the construction of comparable effect sizes for meta-analysis", 2015.

Usage
MadeyskiKitchenham.MetaAnalysis.PBRvsCBRorAR

Format
A data frame with 17 rows and 26 variables:

- **Study**  Name of empirical study
- **Ref.** Reference to the paper reporting primary study or experimental run where data were originally reported
- **Teams**  The number of teams including both, PBR and Control teams
- **DesignDesc**  Experimental design description: Before-after, Between-groups, Cross-over
- **ExpDesign**  Experimental design: between-groups (BG), within-subjects cross-over (WSCO), within-subjects before-after (WSBA)
M_PBR The average proportion of defects found by teams using PBR

M_C The average proportion of defects found by teams using Control treatment: Check-Based Reading (CBR) or Ad-Hoc Reading (AR)

Diff The difference between M_PBR and M_C, i.e. Diff = M_PBR - M_C

Inc The percentage increase in defect rate detection, i.e. Inc=100*[(M_PBR-M_C)/M_C]

SD_C_ByAuthors The standard deviation of the control group values reported by the original Authors, i.e., obtained from the papers/raw data

SD_C The standard deviation of the control group values equals SD_C_ByAuthors for studies for which the data was available OR the weighted average of SD_C_ByAuthors (i.e., 0.169) for studies where SD_C_ByAuthors is missing.

V_C The variance of the Control group observations, i.e., the variance obtained from the teams using the Control method V_C=SD_C^2

V_D The variance of the unstandardized mean difference D (between the mean value for the treatment group and the mean value for the Control group)

SD_C_Alt This is the equivalent of SD_C (the standard deviation of the control group) based on a different variance for the student studies or the practitioner studies depending on the subject type of the study with the missing value.

V_Alt The variance of the mean difference in the meta-analysis based on SD_C_Alt

SS_C The sum of squares of the Control group values. For within subjects studies SS=V_C*(n-1).
For between subjects studies SS=V_C*(n_C-1)

n_PBR The number of PBR teams

n_C The number of Control (CBR or AR) teams

ControlType Type of Control treatment: CRB or AR

ParticipantsType Type of participants: Engineers or Students

TeamType Type of team: Nominal or Real

TwoPersonTeamVsLargerTeam Reflects size of the teams: 2-PersonTeam or LargerTeam

ArtefactType The type of artefact: Requirements or Other

AssociatedWithBasili Whether study is associated with Basili (the forerunner): Yes or No

ControlType_Basili Combined ControlType and AssociatedWithBasili: AH_AssociatedWithBasili, CBR_AssociatedWithBasili, CBR_NotAssociatedWithBasili

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

MadeyskiKitchenham.MetaAnalysis.PBRvsCBRorAR
Description

This data is used in the paper: Tomasz Lewowski and Lech Madeyski, "Creating Evolving Project Data Sets in Software Engineering", vol. 851 of Studies in Computational Intelligence, pp. 1–14. Cham: Springer, 2020. DOI: 10.1007/978-3-030-26574-8_1

Usage

Format

A text file with variables:

- **rowID**: unique id assigned to projects before filtering (source: API)
- **id**: GitHub repository ID (source: API)
- **repository owner**: the organization or user owning the repository (source: API)
- **project name**: name of the project (source: API)
- **manual**: link to best found project documentation - wiki, webpage, documentation directory or readme. Projects with limited documentation were marked with (limited) and ones that had documentation in Chinese - (Chinese) (source: manual)
- **installation**: the recommended installation medium(s) for the project. Some mediums may be missing for projects with multiple recommendations. (source: manual)
- **support**: channel(s) that can be used to get support and/or report bugs. Some channels may be missing for projects with multiple ones. Abbreviations used (source: manual): GH GitHub Issues SO Stack Overflow GG Google Groups ML Mailing list FB Facebook MM Mattermost LI LinkedIn ? not found
- **is not sample/playground/docs/...**: 1 if the project is an actual application or library, 0 if it is a set of samples, only documentation or some experimental area (source: manual)
- **is industrial**: whether the project can be treated as industrial quality one. Values and their meanings: 1 the repository can be classified as industrial grade; 0,5 the repository can sometimes be classified as industrial grade, but it is either a minor project or its documentation or support may be lacking the depth; 0 the repository cannot be classified as industrial-grade; -1 the repository is no longer actively maintained as of the date of data acquisition; -2 the repository is no longer in Java as of the date of data acquisition. (source: manual)
- **createdAt**: the date at which the repository was created (source: API)
- **updatedAt**: the date of last repository update - including changes in projects, watchers, issues etc. (source: API)
pushedAt  the date of last push to the repository - NOT the date of last pushed commit (source: API)  
diskUsage  total number of bytes on disk that are needed to store the repository (source: API)  
dforkCount  number of existing repository forks (independent copies managed by other entities) (source: API)  
isArchived  true if the repository is archived (no longer maintained), false otherwise (source: API)  
isFork  true if the repository is a fork (not the main repository), false otherwise (source: API)  
isMirror  true if the repository is a mirror, false otherwise (source: API)  
sshUrlOfRepository  URL that can be used to immediately clone the repository (source: API)  
licenseInfo.name  name of license under which the project is distributed. Names are the same as in https://choosealicense.com/appendix/ (source: API)  
commitSHA  unique Git identifier of commit that was top of the main branch at the time of data acquisition (source: API)  
defaultBranchRef.target.history.totalCount  number of commits on the default branch in the repository (usually master) at the time of data acquisition (source: API)  
stargazers.totalCount  number of stargazers for the repository at the time of data acquisition (source: API)  
watchers.totalCount  number of watchers for the repository at the time of data acquisition (source: API)  
languages.totalSize  total size of all source code files (source: API)  
Java.byte.count  total size of Java files (source: API)  
Language  main programming language used in the repository, i.e. one that the most code is written in (source: API)  
searchQuery  query used during search that obtained this project (source: API)  

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20190324

Description

This data is used in the paper: Tomasz Lewowski and Lech Madeyski, "How do software engineering data sets evolve? A reproduction study", 2020 (submitted). Generated by: token <- "...
MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20190324<-searchForIndustryRelevantGitHubProjects(token, "2019-03-01", "2018-08-01") usethis::use_data(MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20190324)
Usage

MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20191022

Format

A text file with variables:

- **rowID** unique id assigned to projects before filtering (source: API)
- **id** GitHub repository ID (source: API)
- **repository owner** the organization or user owning the repository (source: API)
- **project name** name of the project (source: API)
- **manual** link to best found project documentation - wiki, webpage, documentation directory or readme. Projects with limited documentation were marked with (limited) and ones that had documentation in Chinese - (Chinese) (source: manual)
- **installation** the recommended installation medium(s) for the project. Some mediums may be missing for projects with multiple recommendations. (source: manual)
- **support** channel(s) that can be used to get support and/or report bugs. Some channels may be missing for projects with multiple ones. Abbreviations used (source: manual): GH GitHub Issues SO Stack Overflow GG Google Groups ML Mailing list FB Facebook MM Mattermost LI LinkedIn ? not found
- **is not sample/playground/docs/...** 1 if the project is an actual application or library, 0 if it is a set of samples, only documentation or some experimental area (source: manual)
- **is industrial** whether the project can be treated as industrial quality one. Values and their meanings: 1 the repository can be classified as industrial grade; 0,5 the repository can sometimes be classified as industrial grade, but it is either a minor project or its documentation or support may be lacking the depth; 0 the repository cannot be classified as industrial-grade; -1 the repository is no longer actively maintained as of the date of data acquisition; -2 the repository is no longer in Java as of the date of data acquisition. (source: manual)
- **createdAt** the date at which the repository was created (source: API)
- **updatedAt** the date of last repository update - including changes in projects, watchers, issues etc. (source: API)
- **pushedAt** the date of last push to the repository - NOT the date of last pushed commit (source: API)
- **diskUsage** total number of bytes on disk that are needed to store the repository (source: API)
- **forkCount** number of existing repository forks (independent copies managed by other entities) (source: API)
- **isArchived** true if the repository is archived (no longer maintained), false otherwise (source: API)
- **isFork** true if the repository is a fork (not the main repository), false otherwise (source: API)
- **isMirror** true if the repository is a mirror, false otherwise (source: API)
- **sshUrlOfRepository** URL that can be used to immediately clone the repository (source: API)
- **licenseInfo.name** name of license under which the project is distributed. Names are the same as in https://choosealicense.com/appendix/ (source: API)
commitSHA  unique Git identifier of commit that was top of the main branch at the time of data acquisition (source: API)

defaultBranchRef.target.history.totalCount  number of commits on the default branch in the repository (usually master) at the time of data acquisition (source: API)

stargazers.totalCount  number of stargazers for the repository at the time of data acquisition (source: API)

watchers.totalCount  number of watchers for the repository at the time of data acquisition (source: API)

languages.totalSize  total size of all source code files (source: API)

Java.byte.count  total size of Java files (source: API)

Language  main programming language used in the repository, i.e. one that the most code is written in (source: API)

searchQuery  query used during search that obtained this project (source: API)

Source

https://madeyski.e-informatyka.pl/reproducible-research/

Examples

MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20191022

---

**MetaAnalysisSimulations**

This function simulates data from many families of experiments. The number of families simulated is defined by the Replications parameter. The parameter Exp determines the number of experiments in each family. The function simulates data from one of four distributions and uses the data to construct two of groups of equal size (GroupSize). The experimental design of individual experiments in each family is determined by the FourGroup parameter. If FourGroup=FALSE, the basic experimental design is a balanced two group randomized experiment, otherwise the experimental design is a balanced four group experiment corresponding to a randomized blocks experiment. The function calls either NP2GroupMetaAnalysisSimulation or NP4GroupMetaAnalysisSimulation to generate and analyse data for each individual family. The function either returns the meta-analysed data from each experiment or provides summary statistics.
Usage

```r
MetaAnalysisSimulations(
    mean = 0,
    sd = 1,
    diff = 0.5,
    GroupSize = 10,
    type = "n",
    Replications = 50,
    Exp = 5,
    seed = 456,
    alpha = 0.05,
    FourGroup = FALSE,
    StdAdj = 0,
    BlockEffect = 0,
    BlockStdAdj = 0,
    StdExp = 0,
    MAMethod = "PM",
    returnES = F
)
```

Arguments

- **mean**: the value used for the mean of control group in the simulated data. It can be any real number including zero.
- **sd**: the value used for the spread of the control group and the spread of the treatment group in the simulated data. The value must be a real value greater than 0.
- **diff**: mean+diff is the value used for the mean of the treatment group. It can be zero.
- **GroupSize**: is the size of each of the groups comprising one experiment. GroupSize should be an integer of 4 or more.
- **type**: specifies the distribution being simulated. The permitted values are "n" for the normal distribution, "l" for the lognormal distribution, "g" for the gamma distribution and "lap" for the Laplace distribution. The parameter defaults to "n".
- **Replications**: The number of times the set of experiments is simulated. It defaults to 500.
- **Exp**: the number of experiments being simulated. Exp should be an integer of 2 or more. It defaults to 5.
- **seed**: specifies the seed to be used to initiate the simulation, so the simulation is repeatable. It defaults to 456.
- **alpha**: This is the Type 1 probability level used in significance tests. It defaults to 0.05.
- **FourGroup**: is a Boolean variable that determines whether the experiment is a two group experiments or a 4-Group randomised block experiment. It defaults to FALSE which means a two group experiment is the default condition.
- **StdAdj**: specifies the adjustment made to the treatment variance if it is required to simulate variance heterogeneity due to the treatment. It defaults to 0.
- **BlockEffect**: specified a fixed block effect that can be applied to randomized blocks designs.
BlockStdAdj specifies the adjustment made to the variance of one block if it is required to simulate heterogeneity due to blocks in a randomized blocks experiment. The variance adjustment is made to the same block that any Block effect is applied to. It defaults to 0.

StdExp specifies the adjustment made to the variance of each experiment if it is required to simulate variance heterogeneity due to individual experiments in a family. It defaults to 0.

MAMethod specifies the model to be used when experimental effect sizes are aggregated using the R metafor package.

returnES if TRUE the function outputs the summary statistics otherwise it outputs the meta-analysis results for each family. The parameter defaults to FALSE

Value

The parameter either returns the meta-analysis values obtained from each family or the average values of the meta-analysis over all replications.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

```r
MetaAnalysisSimulations(mean=0,sd=1,diff=0.5,GroupSize=10,type="n",Replications=25,Exp=5, seed=456,alpha=0.05,FourGroup=FALSE,StdAdj=0,BlockEffect=0,BlockStdAdj=0,StdExp=0,MAMethod="PM")
# A tibble: 1 x 26
# Averagektau Averagektauctvar AveragetauSigCVt AverageCliffd AverageCliffdvar AverageCliffdsig
# <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
# 1 0.164 0.00375 0.72 0.311 0.0134 0.72
#. ... with 20 more variables: Averagephat <dbl>, Averagephatvar <dbl>, Averagephatsig <dbl>,
# MAaverage <dbl>, MAVariance <dbl>, MASignificant <dbl>, HetSignificant <dbl>, PMAverage <dbl>,
# PMASignificant <dbl>, PMASignificant <dbl>, Mean.phat <dbl>, phat.sig <dbl>, Mean.d <dbl>,
# d.sigmoid <dbl>, Mean.g.exact <dbl>, g.exact.sig <dbl>, Mean.g.approx <dbl>, g.approx.sig <dbl>,
# Mean.Cohend <dbl>, Cohend.sig <dbl>

MetaAnalysisSimulations(mean=0,sd=1,diff=0.734,GroupSize=100,type="l",Replications=25,Exp=5, seed=456,alpha=0.05,FourGroup=FALSE,StdAdj=0,BlockEffect=0,BlockStdAdj=0,StdExp=0,MAMethod="PM")
# A tibble: 1 x 26
# Averagektau Averagektauctvar AveragetauSigCVt AverageCliffd AverageCliffdvar AverageCliffdsig
# <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
# 1 0.200 0.000276 1 0.399 0.00109 1
#. ... with 20 more variables: Averagephat <dbl>, Averagephatvar <dbl>, Averagephatsig <dbl>,
# MAaverage <dbl>, MAVariance <dbl>, MASignificant <dbl>, HetSignificant <dbl>, PMAverage <dbl>,
# PMASignificant <dbl>, PMASignificant <dbl>, Mean.phat <dbl>, phat.sig <dbl>, Mean.d <dbl>,
# d.sigmoid <dbl>, Mean.g.exact <dbl>, g.exact.sig <dbl>, Mean.g.approx <dbl>, g.approx.sig <dbl>,
# Mean.Cohend <dbl>, Cohend.sig <dbl>

MetaAnalysisSimulations(mean=0,sd=1,diff=0.5,GroupSize=10,type="n",Replications=10,Exp=5, seed=456,alpha=0.05,FourGroup=TRUE,StdAdj=0,BlockEffect=0.5,BlockStdAdj=0,StdExp=0,
# MAMethod="PM",returnES=TRUE)
# A tibble: 10 x 31
# Family NumExp GroupSize AveKtau AveKtauctvar tauSigCVt AveCliffd AveCliffdvar AveCliffdsig
# <int> <dbl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl>
# 1 0.200 0.000276 1 0.399 0.00109 1
#. ... with 20 more variables: Averagephat <dbl>, Averagephatvar <dbl>, Averagephatsig <dbl>,
# MAaverage <dbl>, MAVariance <dbl>, MASignificant <dbl>, HetSignificant <dbl>, PMAverage <dbl>,
# PMASignificant <dbl>, PMASignificant <dbl>, Mean.phat <dbl>, phat.sig <dbl>, Mean.d <dbl>,
# d.sigmoid <dbl>, Mean.g.exact <dbl>, g.exact.sig <dbl>, Mean.g.approx <dbl>, g.approx.sig <dbl>,
# Mean.Cohend <dbl>, Cohend.sig <dbl>
```
### Description

This function simulates data from a family of experiments. The parameter `Exp` determines the number of experiments in the family. The function simulates data from one of four distributions and uses the data to construct two groups of equal size (`GroupSize`). The distribution for one of the groups corresponds to the control and is based on the given mean and spread, the distribution for the other group corresponds to the treatment group and is based on the mean+diff and the spread plus any variance adjustment (determined by the parameter `StdAdj`). The data from each experiment is analysed separately to estimate three non-parametric effect sizes: the point bi-serial version of Kendall’s tau, Cliff’s d and the probability of superiority referred to as phat and their variances. Parametric effect sizes Cohen’s d (also known as the standardized means difference, SMD) and the small sample size adjusted standardized mean difference g are also calculated together with their variances. The effect sizes are then meta-analysed using two main methods: the simple average of the effect size and the variance weighted average. The function uses the metafor package for formal meta-analysis, and the specific method of formal meta-analysis used is determined by the `MAAMethod`. All tests of significance are done at the 0.05 level. If the parameter `returnES` is TRUE, the function returns the effect sizes for each experiment in the family, otherwise it returns the meta-analysis results.

### Usage

```r
NP2GroupMetaAnalysisSimulation(
  mean,
  sd,
  diff,
  GroupSize,
  Exp = 5,
  type = "n",
)```

---

NP2GroupMetaAnalysisSimulation

---

NP2GroupMetaAnalysisSimulation

---
\[ \text{StdAdj} = 0, \]
\[ \text{alpha} = 0.05, \]
\[ \text{seed} = 123, \]
\[ \text{StdExp} = 0, \]
\[ \text{MAMethod}, \]
\[ \text{returnES} = \text{FALSE} \]

Arguments

- **mean**: the value used for the mean of control group in the simulated data. It can be any real number including zero.
- **sd**: the value used for the spread of the control group and the spread of the treatment group in the simulated data. The value must be a real value greater than 0.
- **diff**: mean + diff is the value used for the mean of the treatment group. It can be zero.
- **GroupSize**: is the size of each of the 2 groups comprising one experiment. GroupSize should be an integer of 4 or more.
- **Exp**: is the number of experiments being simulated. Exp should be an integer of 2 or more. It defaults to 5.
- **type**: specifies the distribution being simulated. The permitted values are "n" for the normal distribution, "l" for the lognormal distribution, "g" for the gamma distribution and "lap" for the Laplace distribution. The parameter defaults to "n".
- **StdAdj**: specifies a level used to adjust the treatment variance. It allows heterogeneity to be modelled. It defaults to zero meaning no variance heterogeneity is introduced.
- **alpha**: the Type 1 error rate level use for statistical tests itb defaults to 0.05.
- **seed**: specifies the seed to be used to initiate the simulation, so the simulation is repeatable. It defaults to 123.
- **StdExp**: specifies the adjustment made to the variance of each experiment, if it is required to simulate variance heterogeneity due to individual experiments in a family. It defaults to 0.
- **MAMethod**: the meta-analysis method needed for the call to the metafor package rma algorithm
- **returnES**: Determines the format of the output. It defaults to FALSE which causes the function to output the meta-analysis results for the family of experiments. If set to TRUE it returns the effect sizes for each experiment.

Value

Depending on the value of the returnES parameter, the function either returns the effect sizes for each experiment or the aggregated results for the family

Author(s)

Barbara Kitchenham and Lech Madeyski
Examples

NP2GroupMetaAnalysisSimulation(mean=0, sd=1, diff=0.5, GroupSize=100, Exp=5, type="n", StdAdj=0, alpha=0.05, seed=123, StdExp=1, MAMethod="PM", returnES=FALSE)
# A tibble: 1 x 30
# NumExp GroupSize AveKtau AveKtauvar tauSigCVt AveCliffd AveCliffdvar AveCliffdsig AvePhat AvePhatvar AvePhatsig
# <dbl> <dbl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl> <dbl>
# <dbl> <lgl>
# 1 5 100 0.126 0.000315 TRUE 0.250 0.00125 TRUE 0.625
# 0.000310 TRUE
# . . . with 19 more variables: MAMean <dbl>, MAvar <dbl>, MASig <lgl>, QE <dbl>, QEp <dbl>,
# HetSig <lgl>, P.mean <dbl>,
NP2GroupMetaAnalysisSimulation(mean=0, sd=1, diff=0.5, GroupSize=10, Exp=5, type="n", StdAdj=0, alpha=0.05, seed=123, StdExp=1, MAMethod="PM", returnES=TRUE)
# A tibble: 5 x 17
# MeanExp VarExp StdESExp df tval tpval tciL tciU Cliffd Cliffdvar PHat PHatvar PHatdf g gvar.exact
# <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
# <dbl> <dbl> <lgl>
#1 0.226 1.03 0.223 17.9 -0.498 0.624 -1.18 0.728 0.1 0.0808 0.55 0.0197
# 17.8 0.213 NA
#2 1.00 0.622 1.27 15.4 -2.84 0.0122 -1.75 -0.251 0.64 0.0400 0.82 0.00947
# 15.0 1.21 NA
#3 0.432 0.848 0.469 18.0 -1.05 0.308 -1.30 0.434 0.36 0.0725 0.68 0.0177
# 15.5 0.449 NA
#4 0.434 0.967 0.441 13.7 -0.986 0.341 -1.38 0.512 0.22 0.0759 0.61 0.0184
# 14.6 0.416 NA
#5 -0.0342 0.782 -0.0387 14.9 0.0865 0.932 -0.809 0.878 -0.2 0.0817 0.4 0.02
# 17.3 -0.0367 NA
# . . . with 2 more variables: gvar.approx <dbl>, . . .17 <dbl>
NP2GroupMetaAnalysisSimulation(mean=0, sd=1, diff=0.724, GroupSize=10, Exp=5, type="l", StdAdj=0, alpha=0.05, seed=123, StdExp=1, MAMethod="PM", returnES=FALSE)
# A tibble: 1 x 30
# NumExp GroupSize AveKtau AveKtauvar tauSigCVt AveCliffd AveCliffdvar AveCliffdsig AvePhat AvePhatvar AvePhatsig
# <dbl> <dbl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl> <dbl>
# <dbl> <lgl>
# 1 5 10 0.181 0.00360 TRUE 0.344 0.0129 TRUE 0.672
# 0.00312 TRUE
# . . . with 19 more variables: MAMean <dbl>, MAvar <dbl>, MASig <lgl>, QE <dbl>, QEp <dbl>,
# HetSig <lgl>, P.mean <dbl>,
# P.rsig <lgl>, P.hetsig <lgl>, Mean.phat <dbl>, phat.sig <lgl>, Mean.d <dbl>, d.sig <lgl>,
# Mean.g.exact <dbl>,
# g.exact.sig <lgl>, Mean.g.approx <dbl>, g.approx.sig <lgl>, Cohend.mean <dbl>,
# Cohend.sig <lgl>
Description
This function simulates data from a family of experiments, where the number of experiments in a family is defined by the parameter \( \text{Exp} \). It simulates data from one of four distributions and uses the data to construct four of groups of equal size (\( \text{GroupSize} \)). Two groups are assigned as control groups and their distribution is based on the parameter mean and the parameter spread, however, the mean and spread for the control group in Block 2 can be adjusted using the parameters \( \text{BlockEffect} \) and \( \text{BlockStdAdj} \) respectively. The other two groups are treatment groups and their distribution is based on the mean+\( \text{diff} \) and the spread parameter, but the distributions can be adjusted using the \( \text{StdAdj} \), \( \text{BlockEffect} \) and \( \text{BlockStdAdj} \) parameters. The data from each experiment is analysed separately to estimate the non-parametric statistics P-hat, Cliff’s d and Kendall’s tau and their variances. The statistics are then meta-analysed using the method specified by the \( \text{MAMethod} \) parameter. We output the average non-parametric effect statistics across the \( \text{Exp} \) experiment as if from a single large experiment and also the results of meta-analysing each non-parametric effect size. We use the standard parametric effect sizes and their meta-analysis as baselines. All tests of significance are done at the 0.05 level.

Usage
```r
NP4GroupMetaAnalysisSimulation(
  mean,
  sd,
  \text{diff},
  \text{GroupSize},
  \text{Exp} = 5,
  \text{type} = \text{"n"},
  \text{alpha} = 0.05,
  \text{seed} = 123,
  \text{StdAdj} = 0,
  \text{BlockEffect} = 0,
  \text{BlockStdAdj} = 0,
  \text{StdExp} = 0,
  \text{MAMethod},
  \text{returnES} = \text{FALSE}
)
```

Arguments

- **mean**: The default value used for the group means in the simulated data. It can be any real number including zero.
- **sd**: The default value used for the spread of the control group and the spread of the treatment group in the simulated data. The value must be a real value greater than 0.
- **\text{diff}**: \( \text{mean} + \text{diff} \) is the value used for the mean of the treatment group. It can be zero.
- **\text{GroupSize}**: is the size of each of the 4 groups comprising one experiment. \( \text{GroupSize} \) should be an integer of 4 or more
- **\text{Exp}**: is the number of experiments being simulated. \( \text{Exp} \) should be an integer of 2 or more. It defaults to 5.
NP4GroupMetaAnalysisSimulation

**type** specifies the distribution being simulated. The permitted values are "n" for the normal distribution, "l" for the lognormal distribution, "g" for the gamma distribution and "lap" for the Laplace distribution. The parameter defaults to "n".

**alpha** The Type 1 value used in all significance tests. It defaults to 0.05.

**seed** specifies the seed to be used to initiate the simulation, so the simulation is repeatable. It defaults to 123.

**StdAdj** The value used to introduce heterogeneity into the treatment groups variance if required.

**BlockEffect** is the effect of having two different blocks

**BlockStdAdj** is the variance associated with the Block. If BlockStdAdj is zero it means we are treat the block effect as a fixed effect. If BlockStdAdj>0, we treat the block effect as a random effect and increase the variance of Block 2 data.

**StdExp** The value used to introduce heterogeneity into experiments in a family required.

**MAMethod** defines the method used for meta-analysis

**returnES** default to FALSE

**Value**

If returnES is true, the function returns the summary meta-analysis summary statistics otherwise the function returns the effect sizes for each experiment

**Author(s)**

Barbara Kitchenham and Lech Madeyski

**Examples**

NP4GroupMetaAnalysisSimulation(mean=0,sd=1,diff=0.5,GroupSize=10,Exp=5,type="n",alpha=0.05, seed=123,StdAdj=0,BlockEffect=0.5,BlockStdAdj=0,StdExp=0,MAMethod="PM")

# A tibble: 1 x 30
# NumExp GroupSize AveKtau AveKtauctvar tauSigCVt AveCliffd AveCliffdvar AveCliffdsig Avephat Avephatvar Avephatsig 
# Avephatvar Avephatsig 
# <dbl> <dbl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl> <dbl> 
# <dbl> <lgl> 
# 1 5 10 0.182 0.00188 TRUE 0.346 0.00673 TRUE 0.673 
# 0.00163 TRUE 
# . . . with 19 more variables: MAMean <dbl>, MAvar <dbl>, MASig <lgl>, QE <dbl>, QEp <dbl>, 
# HetSig <lgl>, P.mean <dbl>,

NP4GroupMetaAnalysisSimulation(mean=0,sd=1,diff=0.724,GroupSize=10,Exp=5,type="l",alpha=0.05, seed=123,StdAdj=0,BlockEffect=0.5,BlockStdAdj=0,StdExp=0,MAMethod="PM")

# A tibble: 1 x 30
# NumExp GroupSize AveKtau AveKtauctvar tauSigCVt AveCliffd AveCliffdvar AveCliffdsig Avephat Avephatvar Avephatsig 
# Avephatvar Avephatsig 
# <dbl> <dbl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl> <dbl> 
# <dbl> <lgl> 
# 1 5 10 0.244 0.00167 TRUE 0.464 0.00593 TRUE 0.732 
# 0.00144 TRUE 
# . . . with 19 more variables: MAMean <dbl>, MAvar <dbl>, MASig <lgl>, QE <dbl>, QEp <dbl>, 
# HetSig <lgl>, P.mean <dbl>,
NP4GroupMetaAnalysisSimulation(mean=0, sd=1, diff=0.5, GroupSize=10, Exp=5, type="n", alpha=0.05, seed=123, StdAdj=0, BlockEffect=0.5, BlockStdAdj=0, StdExp=0, MAMethod="PM", returnES=TRUE)
# A tibble: 5 x 16
#  MeanExp VarExp StdESExp df tval tpval tciL tciU Cliffd Cliffdvar PHat PHatvar
#  g gvar.approx g gvar.approx g gvar.approx g gvar.approx g gvar.approx g gvar.approx
#1 0.940 0.783 1.06 31.3 3.36 0.00206 0.370 1.51 0.58 0.0243 0.29 0.00587
#2 0.372 0.943 0.383 35.0 1.21 0.234 -0.251 0.996 0.21 0.0380 0.105 0.00927
#3 0.598 0.619 0.761 28.6 2.40 0.0229 0.0892 1.11 0.37 0.0336 0.185 0.00813
#4 0.873 1.13 0.821 28.1 2.60 0.0148 0.184 1.56 0.440 0.0333 0.220 0.00813
#5 0.243 1.03 0.240 31.5 0.758 0.454 -0.410 0.896 0.13 0.0390 0.065 0.00946
# . . . with 1 more variable: Cohendvar <dbl>

percentageInaccuracyOfLargeSampleVarianceApproximation
percentageInaccuracyOfLargeSampleVarianceApproximation

Description
Plot the extent of inaccuracy using the large sample approximate effect size variance on 4 related graphs corresponding to the four different correlation values. Plot visualizes the relationship between sample size and effect size and the percentage inaccuracy of the large sample variance approximation. Function is used in a paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham.

Usage
percentageInaccuracyOfLargeSampleVarianceApproximation(data)

Arguments
data - data behind the plot returned by getSimulatedCrossoverDataSets() or stored in reproducer::KitchenhamMadeyski.SimulatedCrossoverDataSets

Value
plot described in description

Author(s)
Barbara Kitchenham and Lech Madeyski
Examples

```r
data <- KitchenhamMadeyski.SimulatedCrossoverDataSets
myPlot <- percentageInaccuracyOfLargeSampleVarianceApproximation(data)
```

Description

Function to plot a figure on the outcomes for individuals in each sequence group used in a paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham

Usage

```r
plotOutcomesForIndividualsInEachSequenceGroup(
  var,
  covar,
  meanA1,
  treatmentDiff,
  periodEffect,
  numOfSamples
)
```

Arguments

- `var`: Variance among subjects is a sum of the between subjects variance and the within subjects variance
- `covar`: Covariance equal to the between subjects variance
- `meanA1`: Mean for treatment sequence A1
- `treatmentDiff`: Technique effect which is the difference between the effect of technique A and technique B
- `periodEffect`: Period effect which is the difference between period 1 and period 2
- `numOfSamples`: Number of samples ("rows" of data) required for each technique and period

Value

- `plot`

Author(s)

Lech Madeyski and Barbara Kitchenham

Examples

```r
myPlot <- plotOutcomesForIndividualsInEachSequenceGroup(25, 18.75, 50, 10, 5, 15)
```
PrepareForMetaAnalysisGtoR

Description

This function calculates the standardized effect sizes and their confidence intervals, the equivalence point biserial effect size and the \( Z_r \) and \( \text{var}(Z_r) \) needed for input into the metafor rma function (meta analysis). In this function the point bi-serial effect size is based on the adjusted Hedges g value. The function uses the Hedges g to r transformation to prepare for meta-analysing the data where the mean values, the standard deviations, and the number of observations are available.

Usage

\[
\text{PrepareForMetaAnalysisGtoR}(Mc, Mt, SDc, SDt, Nc, Nt)
\]

Arguments

- \( Mc \): is a vector containing the mean value of the control group for each experiment.
- \( Mt \): is a vector containing the mean value of the treatment group for each experiment.
- \( SDc \): is a vector of the standard deviations of the control group for each experiment.
- \( SDt \): is a vector of the standard deviations of the treatment group for each experiment.
- \( Nc \): is a vector containing the number of observations (participants) in the control group for each experiment.
- \( Nt \): is a vector of the number of observations (participants) in the treatment group for each experiment.

Value

data frame incl. calculated effect sizes (Hedges’ g, Hedges’ g adjusted), upper and lower confidence bounds on Hedges’ g, \( Z_r \), \( \text{var}(Z_r) \), \( r \) and \( p \)-value

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

\[
\begin{align*}
\text{PrepareForMetaAnalysisGtoR}(c(10, 10), c(12, 14), c(4, 4), c(4, 4), c(20, 20), c(40, 40)) \\
\text{HGvalues.Hg} & \quad \text{HGvalues.HgAdjusted} & \quad \text{Hgupper} & \quad \text{Hglower} & \quad \text{Zr} & \quad \text{vi} & \quad r & \quad pvalue \\
\# & 0.5 & 1.082017 & 0.4935018 & -0.06156572 & 0.2305901 & 0.01754386 & 0.2265882 & 0.0006813222 \\
\# & 1.0 & 1.634701 & 0.9870036 & 0.40620071 & 0.4499419 & 0.01754386 & 0.4218513 & 0.0816981743
\end{align*}
\]
Description

print data table using xtable R package

Usage

printXTable(
  data,
  selectedColumns,
  tableType = "latex",
  alignCells,
  digits,
  caption,
  label,
  fontSize,
  captionPlacement = "bottom",
  alignHeader
)

Arguments

data	Data structure including columns to be printed.
selectedColumns	Columns selected to be printed.
tableType	Type of table to produce. Possible values are "latex" or "html". Default value is "latex".
alignCells	Defines how to align data cells.
digits	Defines the number of decimal points in each column.
caption	Caption of the table.
label	Label of the table.
fontSize	Size of the font used to produce a table.
captionPlacement	The caption will be have placed at the bottom of the table if captionPlacement is "bottom" and at the top of the table if it equals "top". Default value is "bottom".
alignHeader	Defines how to align column headers of a table.

Value

A table generated on the fly on a basis of passed data (data, selectedColumns etc.).

Author(s)

Lech Madeyski
Examples

```r
d <- reproducer::MadeyskiKitchenham.MetaAnalysis.PBRvsCBRorAR
reproducer::printXTable(d, "Study", "latex", "cc", 0, "C", "L", "tiny", "top", "l")
```

Description

Plots visualize the relationship between sample size, effect size and the proportion of significant t-values using the correct analysis. Function is used in a paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham.

Usage

```r
proportionOfSignificantTValuesUsingCorrectAnalysis(data)
```

Arguments

- `data` - data behind the plot returned by getSimulatedCrossoverDataSets() or stored in reproducer::KitchenhamMadeyski.SimulatedCrossoverDataSets

Value

plot described in description

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

```r
data <- KitchenhamMadeyski.SimulatedCrossoverDataSets
myPlot <- proportionOfSignificantTValuesUsingCorrectAnalysis(data)
```
Description

Plots visualize the relationship between sample size, effect size and the proportion of significant t-values using the incorrect analysis. Function is used in a paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham.

Usage

proportionOfSignificantTValuesUsingIncorrectAnalysis(data)

Arguments

data - data behind the plot returned by getSimulatedCrossoverDataSets() or stored in reproducer::KitchenhamMadeyski.SimulatedCrossoverDataSets

Value

plot described in description

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

data <- KitchenhamMadeyski.SimulatedCrossoverDataSets
myPlot <- proportionOfSignificantTValuesUsingIncorrectAnalysis(data)

Description

This function performs multiple simulations of two-group balanced experiments for one of four distributions and a specific group size. It identifies the average value of phat, Cliff’s d and Kendall’s point biserial tau and their variances. It either returns the effect sizes for each non-parametric effect size or it reports the number of times the each non-parametric effect size is assessed to be significantly different from zero. We also present the values for the t-test as a comparison. For log-normal data the results of analysing the transformed data are also reported.
**RandomExperimentSimulations**

**Usage**

```r
RandomExperimentSimulations(
  mean,
  sd,
  diff,
  N,
  reps,
  type = "n",
  seed = 123,
  StdAdj = 0,
  returnData = FALSE
)
```

**Arguments**

- **mean**
  - The default mean used for both groups (one treatment group and one control group). It can be changed for the treatment group using the parameter **diff**

- **sd**
  - This is the default spread for both groups. It must be a real value greater than 0. It can be adjusted for the treatment group using the parameter **StdAdj**

- **diff**
  - This is added to the treatment group mean. It can be a real value and can take the value zero.

- **N**
  - This is the number of observations in each group. It must be an integer greater than 3.

- **reps**
  - This identifies the number of times each experiment simulation is replicated.

- **type**
  - This specifies the underlying distribution used to generate the data. It takes the values "n" for a normal distribution, "l" for lognormal distribution, "g" for a gamma distribution, "lap" for a Laplace distribution.

- **seed**
  - This specifies the initial seed for the set of replications (default 123).

- **StdAdj**
  - This specifies the extent of variance instability introduced by the treatment and it must be non-negative but can be 0.

- **returnData**
  - If TRUE, the function returns the individual effect sizes and their variances, otherwise it returns summary statistics (default FALSE).

**Author(s)**

Barbara Kitchenham and Lech Madeyski

**Examples**

```r
#RandomExperimentSimulations(mean=0,sd=1,diff=0.5,N=20, reps=500,type="n",seed=123,StdAdj=0)
#reps=500 may take a bit too for CRAN so let us use reps=50
RandomExperimentSimulations(mean=0,sd=1,diff=0.5,N=20, reps=50,type="n",seed=123,StdAdj=0)
# A tibble: 1 x 17
# phat phatvar sigphat emp.phat.var d dvar sigd emp.d.var ktau ktauvar emp.tau.var
# kpowerCVt tpower ES
# <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
# <dbl> <dbl> <dbl>
RandomizedBlockDesignEffectSizes

Description

This function finds the theoretical effect sizes for a four-group randomized block experiments assuming one of four different underlying distributions specified by the type parameter. The design assumes two blocks each comprising a control and treatment group. If required a fixed Blocking effect is added to the mean for Block 2.

Usage

RandomizedBlockDesignEffectSizes(m1, std1, m2, std2,
RandomizedBlockDesignEffectSizes

m3,
std3,
m4,
std4,
BE = 0,
type = "n"
)

Arguments

m1 The theoretical mean for the control group in Block 1
std1 The theoretical variance for the control group in Block 1
m2 The theoretical mean for the treatment group in Block 1
std2 The theoretical variance for the treatment group in Block 1
m3 The theoretical mean for the control group in Block 2
std3 The theoretical variance for the control group in Block 2
m4 The theoretical mean for the treatment group in Block 2
std4 The theoretical variance for the treatment group in Block 2
BE A fixed block effect to be added to the Block 2 mean values.
type String identifying the distribution, "n" for normal, "ln" for lognormal, "lap" for Laplace, "g" for Gamma

Value

dataframe holding the expected unstandardized mean difference effect size, the pooled within group variance, the standardized effect size and the point bi-serial correlation.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

RandomizedBlockDesignEffectSizes(m1=0,std1=1,m2=1,std2=1,m3=0,std3=1,m4=1,std4=1,BE=1,type="n")
# A tibble: 1 x 4
#  ES  Var StdES rPBS
#1  1   1   1  0.447
RandomizedBlockDesignEffectSizes(m1=0,std1=1,m2=1,std2=1,m3=0,std3=1,m4=1,std4=1,BE=1,type="l")
# A tibble: 1 x 4
#  ES  Var StdES rPBS
#1  5.27 82.2 0.581 0.279
RandomizedBlockDesignEffectSizes(m1=0,std1=1,m2=0.266,std2=1,m3=0,std3=1,m4=0.266,std4=1,BE=0,
type = "l")
# A tibble: 1 x 4
#  ES  Var StdES rPBS
#1  0.502 6.31 0.200 0.0995
RandomizedBlocksAnalysis

Description

The function performs a heteroscedastic test of a two treatment by J blocks randomized blocks
effect size. The data are assumed to be stored in $x$ in list mode, a matrix or a data frame. If in list
mode, length(x) is assumed to correspond to the total number of groups. All groups are assumed to
be independent. Missing values are automatically removed.

Usage

RandomizedBlocksAnalysis(x, con = c(-0.5, 0.5, -0.5, 0.5), alpha = 0.05)

Arguments

x
the structure holding the data. In list format, for a 2 treatment by J block ran-
domized blocks experiments, there are 2J list elements each one specifying the
outcome for a specific block and a specific treatment.

con
is a 2J list containing the contrast coefficients that are used to calculate the mean
effect size.

alpha
default to 0.05 is the Type 1 error level used for the test of significance

Value

The t-test and its associated metrics (i.e., critical value standard error and degrees of freedom) and
the estimate of the contrast with its upper and lower confidence interval bounds and p-value.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

set.seed(123)
x=list()
x[[1]]=stats::rnorm(10,0,1)
x[[2]]=stats::rnorm(10,0.8,1)
x[[3]]=stats::rnorm(10,0.5,1)
x[[4]]=stats::rnorm(10,1.3,1)
vec=c(-1,1,-1,1)/2
RandomizedBlocksAnalysis(x,con=vec,alpha=0.05)
# $n
# [1] 10 10 10 10
# $test
# test  crit  se   df
# [1,] 4.432644 2.038622 0.2798104 31.33793
# $\psi_{\text{ihat}}$
# \psi_{\text{ihat}} \ ci.lower \ ci.upper \ \text{p.value}
# [1,] 1.2403 0.6698721 1.810728 0.0001062952
# dat=c(x[[1]],x[[2]],x[[3]],x[[4]])
# matx=matrix(dat,nrow=10,ncol=4)
# RandomizedBlocksAnalysis(matx,con=c(-1,1,-1,1)/2,alpha=0.05)
#$n$
# [1] 10 10 10 10
#$\text{test}$
# test \ crit \ se \ df
# [1,] 4.432644 2.038622 0.2798104 31.33793
#$\psi_{\text{ihat}}$
# \psi_{\text{ihat}} \ ci.lower \ ci.upper \ \text{p.value}
# [1,] 1.2403 0.6698721 1.810728 0.0001062952

---

Description

This function performs multiple simulations of 4 group balanced randomised Block experiments with two control groups and two treatment groups where one control group and one treatment group are assigned to block 1 and the other control group and treatment group are assigned to block 2. The simulations are based on one of four distributions and a specific group size. The function identifies the average value of the non-parametric effect sizes $P$-hat, Cliff’s d and Kendall’s point biserial tau and their variances and whether or not the statistics were significant at the 0.05 level. We also present the values of the t-test as a comparison.

Usage

RandomizedBlocksExperimentSimulations(
  mean,
  sd,
  diff,
  N,
  reps,
  type = "n",
  alpha = 0.05,
  Blockmean = 0,
  BlockStdAdj = 0,
  StdAdj = 0,
  seed = 123,
  returnData = FALSE
)
Arguments

**mean**
The default mean for all 4 groups. The default for the two treatment groups can be altered using the parameter `diff` and the block mean for block 2 can be altered using the parameter `Blockmean`.

**sd**
The default spread for all 4 groups. It must be a real value greater than 0. If can be altered for treatment groups using the parameter `StdAdj` and for Block 2 groups using `BlockStdAdj`.

**diff**
The is is added to the parameter mean, to define the mean of the other treatment group. It can be a real value and can take the value zero.

**N**
this is the number of observations in each group. It must be an integer greater than 3.

**reps**
this identifies the number of time the simulation is replicated.

**type**
this specifies the underlying distribution used to generate the data. it takes the values "n" for a normal distribution, "l" for lognormal distribution,"g" for a gamma distribution, "lap" for a Laplace distribution.

**alpha**
is the Type 1 error level used for constructing confidence intervals

**Blockmean**
is the effect of having two different blocks

**BlockStdAdj**
is the variance associated with the Block mean. If Blockvar is zero it means we are treat the block effect as a fixed effect. If BlockStdAdj>0, we treat the block effect as a random effect.

**StdAdj**
The value used to introduce heterogeneity into the treatment groups variance if required.

**seed**
this specifies the seed value for the simulations and allows the experiment to be repeated.

**returnData**
= FALSE if TRUE the function returns the generated data otherwise it returns summary statistics.

Value

depending on the parameter `returnData` it returns the generated nonparametric and parametric values or the summary statistics examples RandomizedBlocksExperimentSimulations(mean=0,sd=1,diff=0.5,N=10, reps=500,type="n",alpha=0.05,Blockmean=0.5, BlockStdAdj=0,StdAdj=0,seed=123) # A tibble: 1 x 18 # phat varphat sigphat emp.phat.var d vard sigd emp.d.var ktau kconsistentvar emp.tau.var kpowerCVt StdES # <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> # 1 0.640 0.00832 0.326 0.00773 0.279 0.0343 0.282 0.0309 0.147 0.00956 0.00856 0.168 0.513 # . . . with 5 more variables: ES <dbl>, Var <dbl>, emp.StdESvar <dbl>, MedDiff <dbl>, tpower <dbl> RandomizedBlocksExperimentSimulations(mean=0,sd=1,diff=0.5,N=10, reps=10,type="n",alpha=0.05,Blockmean=0.5, BlockStdAdj=0,StdAdj=0,seed=123,returnData=T) # A tibble: 10 x 3 # Cliffd PHat StdES # <dbl> <dbl> <dbl> # 1 0.58 0.79 1.06 # 3 0.37 0.685 0.761 # 4 0.440 0.72 0.821 # 5 0.13 0.565 0.240 # 6 0.16 0.58 0.222 # 7 0.38 0.69 0.580 # 8 0.48 0.74 0.882 # 9 0.11 0.555 0.181 #10 -0.03 0.485 0.124

Author(s)

Barbara Kitchenham and Lech Madeyski
RandomizedDesignEffectSizes

Description

This function creates the theoretical effect sizes for random samples from one of four different distributions for specified parameter values for the distribution specified by the type parameter. It assumes there are two samples, one corresponding to a control group and the other to the treatment group. It returns the theoretical effect sizes for a fully randomized experiment.

Usage

RandomizedDesignEffectSizes(m1, std1, m2, std2, type = "n")

Arguments

- m1: The theoretical mean for the control group
- std1: The theoretical variance for the control group
- m2: The theoretical mean for the treatment group
- std2: The theoretical variance for the treatment group
- type: String identifying the distribution, "n" for normal, "ln" for lognormal, "lap" for Laplace, "g" for Gamma

Value

dataframe containing the expected values of the unstandardized mean difference effect size, the pooled within group variance, the standardized mean difference effect size and the point bi-serial correlation.

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

```r
RandomizedDesignEffectSizes(m1=0, std1=1, m2=1, std2=3, type = "n")
# A tibble: 1 x 4
#  ES Var StdES rPBS
#1 1  5 0.447 0.218
RandomizedDesignEffectSizes(m1=0, std1=1, m2=1, std2=3, type = "ln")
# A tibble: 1 x 4
#  ES Var StdES rPBS
#1 243. 242552663. 0.0156 0.00780
RandomizedDesignEffectSizes(m1=0, std1=1, m2=0.266, std2=1, type = "lap")
```
readExcelSheet

# A tibble: 1 x 4
#  ES  Var StdES  rPBS
#1  0.502 6.31 0.200 0.0995

Description
Function reads data from an Excel file from a specified sheet

Usage
readExcelSheet(path, sheet, colNames)

Arguments
path Path to an Excel file, e.g. /User/lma/datasets/MyDataSet.xls
sheet Name of a sheet within an Excel file we want to read
colNames If TRUE, first row of data will be used as column names.

Author(s)
Lech Madeyski

Examples
myPath=system.file("extdata", "DataSet.xlsx", package = "reproducer")
Madeyski15SQJ.NDC<-readExcelSheet(path=myPath, sheet="Madeyski15SQJ.NDC", colNames=TRUE)

reproduceForestPlotRandomEffects
reproduceForestPlotRandomEffects()

Description
Function reproduces Forest Plot of a Random-Effects Meta-analysis of Mean Differences.

Usage
reproduceForestPlotRandomEffects()

Author(s)
Lech Madeyski
Description
Function reproduces Mixed-Effects Analysis using Subject Specific Estimated Variance with Experimental Design as a Moderator.

Usage
reproduceMixedEffectsAnalysisWithEstimatedVarianceAndExperimentalDesignModerator()

Author(s)
Lech Madeyski

Examples
reproduceMixedEffectsAnalysisWithEstimatedVarianceAndExperimentalDesignModerator()
reproduceMixedEffectsForestPlotWithExperimentalDesignModerator

Description

Function reproduces Forest Plot of a Mixed Effects Meta-analysis of Mean Differences with Experimental Design as a Moderator Variable.

Usage

reproduceMixedEffectsForestPlotWithExperimentalDesignModerator()

Author(s)

Lech Madeyski

Examples

reproduceMixedEffectsForestPlotWithExperimentalDesignModerator()

reproduceSimulationResultsBasedOn500Reps1000Obs

Description

Function to calculate simulation results based on 500 repetitions of 1000 observation samples. Function is used in a paper "Effect Sizes and their Variance for AB/BA Crossover Design Studies" by Lech Madeyski and Barbara Kitchenham.

Usage

reproduceSimulationResultsBasedOn500Reps1000Obs()

Value

data frame including the following simulation results: # treatmentEffect.Ave - Average Technique Effect # dRM.Ave - Average dRM # dRM.Var - Variance of dRM # dRM.Var.Ave - Average of var(dRM) # dRM.Var.ModerateSampleSizeApprox - # dIG.Ave - Average dIG # dIG.Var - Variance of dIG # dIG.Var.Ave - Average of var(dIG) # dIG.Var.ModerateSampleSizeApprox -

Author(s)

Lech Madeyski and Barbara Kitchenham
Examples

```r
# return simulation results based on 500 repetitions of 1000 observation samples
simulationResultsTable500x1000<-reproduceSimulationResultsBasedOn500Reps1000Obs()
```

**Description**

This function reproduces five of the output tables used in the systematic review paper "Meta-analysis for Families of Experiments: A Systematic Review and Reproducibility Assessment". It extracts the reported values for effect sizes, meta-analysis and descriptive statistics in the primary studies. It uses the descriptive statistics to re-calculate effect sizes and then performs a meta-analyses using the constructed effect sizes and compares the calculated values with the reported values.

**Usage**

```r
reproduceTablesOfPaperMetaAnalysisForFamiliesOfExperiments()
```

**Value**

list incl. the data presented in five of the tables presented in the paper.

**Author(s)**

Barbara Kitchenham and Lech Madeyski

**Examples**

```r
rrData = reproduceTablesOfPaperMetaAnalysisForFamiliesOfExperiments()
# Reproduce Table "Overall Mean Values of Effect Sizes Reported and Calculated":
xtable::xtable(rrData$MAStats)
# Reproduce Table "Calculated and Reported Effect Sizes":
xtable::xtable(rrData$ESdata)
# Report values for 3 papers that reported per document
rrData$MAStatsTP1=data.frame(rrData$MAStatsTP1,row.names=NULL)
rrData$ESTP1res=data.frame(rrData$ESTP1res,row.names=NULL)
xtable::xtable(rrData$MAStatsTP1)
xtable::xtable(rrData$ESTP1res)
# Report extra results for Study 8
# Reproduce Table "Calculating r_PB Effect Size from Probabilities"
xtable::xtable(rrData$GH2015extra)
```
Description
Function reproduces Table, which shows the effect sizes based on mean differences.

Usage
reproduceTableWithEffectSizesBasedOnMeanDifferences()

Author(s)
Lech Madeyski

Examples
reproduceTableWithEffectSizesBasedOnMeanDifferences()

Description
Function reproduces Table with possible moderating factors.

Usage
reproduceTableWithPossibleModeratingFactors()

Author(s)
Lech Madeyski

Examples
reproduceTableWithPossibleModeratingFactors()
Description

Function reproduces Table, which shows the effect sizes reported by Ciolkowski identifying the type of design used in each study.

Usage

reproduceTableWithSourceDataByCiolkowski()

Author(s)

Lech Madeyski

Examples

reproduceTableWithSourceDataByCiolkowski()

rSimulations

Description

This function simulates many datasets from the same bivariate distribution to investigate the distribution of correlations for specific sample sizes.

Usage

rSimulations(
  mean,
  var,
  diff,
  r,
  N,
  reps,
  VarAdj = 0,
  seed = 123,
  returntSignificant = F,
  returndata = F,
  plothist = F
)
rSimulations

Arguments

mean  The mean used for one of bivariate distributions - assumed to be the control condition in an experiment.

var  The variance used for both treatment groups. It must be a real value greater than 0.

diff  This value is added to the parameter mean to specify the mean for the other bivariate distribution - assumed to be the treatment condition in an experiment.

r  This specifies the correlation coefficient to be used for the bivariate normal distribution it must be a value in the range [-1, 1].

N  The number of observations in each simulated bivariate normal data set.

reps  The number of bivariate data sets that will be simulated.

VarAdj  This value will be added to the variance of the treatment condition.

seed  This specifies the seed value for the simulations and allows the experiment to be repeated.

returntSignificant  If set to true the percentage of times the t-test delivered a value significant at the 0.05 level is reported (default returntSignificant=F).

returndata  If set to FALSE, the function returns the summary information across all the replications (default returndata=F). If set to TRUE the function outputs the r and variance ratio, and variance accuracy values generated in each replication.

plothist  If set to T, the function outputs a histogram of the r-values, the varprop values and the accuracy values (default plothist=F).

Value

output If returndata=F, the output returns summary information about the average of r and the variance properties across the replicated data sets. If returndata=T, the function returns the r-values obtained for each of the simulated data sets to gather with the variance ratio, the variance accuracy measure and a dummy variable indicating whether a test of significance between the mean values was significant (which is indicated by the dummy variable being set to 1) or not (which is indicated by the dummy variable being set to 0).

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

# output=rSimulations(mean=0,var=1,diff=0,r=0.25,N=4,reps=10000)
# reduced reps to pass CRAN time limits
output=rSimulations(mean=0,var=1,diff=0,r=0.25,N=4,reps=1000)
output=signif(output,4)
output
# r.Mean r.Median Var.r PercentNegative Mean.VarProp Variance.VarProp ...
# 1 0.2132 0.3128 0.3126 34.21 0.5036 0.06046 ...

#output=rSimulations(mean=0,var=1,diff=0.8,r=0.25,N=60,reps=10000,returntSignificant=TRUE)
```r
# reduced reps to pass CRAN time limits
output=rSimulations(mean=0,var=1,diff=0.8,r=0.25,N=60,reps=1000,returntSignificant=TRUE)
output=signif(output,4)
output

output=rSimulations(mean=0,var=1,diff=0,r=0.25,N=30,reps=10,returndata=TRUE)
output
```

### Description

Function searches for industry relevant software projects available from GitHub. The function was used to deliver data set of software projects in an NCBiR project. More details are described in a report: Lech Madeyski, “Training data preparation method,” tech. rep., code quest (research project NCBiR POIR.01.01.01-00-0792/16), 2019, as well as a paper: Tomasz Lewowski and Lech Madeyski, "Creating evolving project data sets in software engineering", 2019. If you use this function or the returned data set than please cite: Tomasz Lewowski and Lech Madeyski, "Creating evolving project data sets in software engineering", 2019

### Usage

```r
searchForIndustryRelevantGitHubProjects(
  myToken,
  earliestPushDate,
  latestCreationDate
)
```

### Arguments

- **myToken**: A private token used to access GitHub
- **earliestPushDate**: Only repositories which were pushed after this date will be included in the results (i.e., repositories for which the latest push was before this date will not be included in the results)
latestCreationDate
Only repositories which were created before this date will be included in the results (i.e., repositories created after this date will not be included in the results)

Value
selected GitHub projects

Author(s)
Lech Madeyski and Tomasz Lewowski

Examples
#to run this function you need to use your own token as a parameter of the function
#use your own token as the first parameter of the function
#searchForIndustryRelevantGitHubProjects("...", "2019-03-01", "2018-08-01")

Description
This simulates one of four distributions, and finds the values of ktau and Cliffs d and their variances. It simulates a randomised blocks experiment with two treatment groups and two control groups each of which being divided into two blocks. By default it assumes equal group sizes but group spread (standard deviation can be changed, see Stadj). It returns values of both parametric and non-parametric effect sizes and their variance for simulated experiments. It returns the number of times each effect size was significant. For the logarithmic distribution it calculates effect sizes based on the log transformed data as well as the raw data.

Usage
simulateRandomizedBlockDesignEffectSizes(
    mean,
    sd,
    diff,
    N,
    type = "n",
    alpha = 0.05,
    Blockmean = 0,
    BlockStdAdj = 0,
    StdAdj = 0
)
Arguments

mean  The default value for all groups which can be changed for the two treatment groups using the parameter diff and for the two block 2 groups using the parameter Blockmean

sd    The default spread used for all four groups unless adjusted by the StdAdj. It must be a real value greater than 0.

diff  This is added to the parameter mean to obtain the required mean for treatment groups. It can be a real value and can take the value zero.

N     this is the number of observations in each group. It must be an integer greater than 3.

type  this specifies the underlying distribution used to generate the data. it takes the values "n" for a normal distribution, "l" for lognormal distribution,"g" for a gamma distribution, "lap" for a Laplace distribution.

alpha is the significance level for statistical tests

Blockmean if >0 an adjustment made to both group means in Block 2

BlockStdAdj if >0, an adjustment that can be made to the sd of each group in block 2

StdAdj this specifies the extent of variance instability introduced by the treatment and if >0 will be used to amend the sd parameter for both treatment groups.

Value
data frame incl. the non-parametric and parametric effect sizes and whether the effect sizes are significant at the 0.05 level.

Author(s)
Barbara Kitchenham and Lech Madeyski

Examples

```
set.seed(123)
simulateRandomizedBlockDesignEffectSizes(mean=0,sd=1,diff=.5,N=10,type="n",alpha=0.05,
Blockmean=0.5,BlockStdAdj=0,StdAdj=0)
# A tibble: 1 x 23
#   phat phat.var phat.df phat.test phat.pvalue phat.sig d vard d.sig cor sqse
#    <dbl>  <dbl>  <dbl>    <dbl>       <dbl> <lgl> <dbl> <dbl> <lgl> <dbl> <dbl>
#1     0.79 0.00587 30.2    3.79      0.000681 TRUE 0.58 0.0243 TRUE 0.305 0.0132

set.seed(123)
simulateRandomizedBlockDesignEffectSizes(mean=0,sd=1,diff=.5,N=10,type="l",alpha=0.05,
Blockmean=0.5,BlockStdAdj=0,StdAdj=0)
# A tibble: 1 x 27
#   phat phat.var phat.df phat.test phat.pvalue phat.sig d vard d.sig cor sqse
```

**simulateRandomizedDesignEffectSizes**

**Description**

This simulates one of four distributions, and finds the values of ktau, phat and Cliffs d and their variances. It assumes equal group sizes. It returns values of the effect sizes and their variance for a simulated randomized experiment with two treatments. It returns whether to not each non-parametric effect size was significant. It also returns the parametric (unstandardized and unstandardized) Effect Size and the whether the t-test was significant.

**Usage**

```r
simulateRandomizedDesignEffectSizes(mean, sd, diff, N, type = "n", StdAdj = 0)
```

**Arguments**

- **mean**
  The mean used for one of the treatment groups
- **sd**
  The spread used for both treatment groups. It must be a real value greater than 0.
- **diff**
  This is added to the parameter mean, to define the mean of the other treatment group. It can be a real value and can take the value zero.
- **N**
  This is the number of observations in each group. It must be an integer greater than 3.
- **type**
  This specifies the underlying distribution used to generate the data. It takes the values "n" for a normal distribution, "l" for lognormal distribution, "g" for a gamma distribution, "lap" for a Laplace distribution.
- **StdAdj**
  This specifies the extent of variance instability introduced by the treatment.

**Value**

data frame incl. the non-parametric and parametric effect sizes and whether the effect sizes are significant at the 0.05 level.

**Author(s)**

Barbara Kitchenham and Lech Madeyski
Examples

```
set.seed(123)
simulateRandomizedDesignEffectSizes(mean=0, sd=1, diff=0.8, N=10, type="n", StdAdj=0)
# A tibble: 1 x 15
# phat varphat dfphat sigphat d vard sigd cor varcor sigCVt ttestp ES Variance StdES
# MedDiff
# <dbl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl> <dbl> <dbl> <dbl> <dbl>
# 1 0.75 0.0152 17.5 FALSE 0.5 0.0624 FALSE 0.263 0.0175 FALSE 0.0507 0.934 0.994 0.937
# 1.26
```

```
set.seed(123)
simulateRandomizedDesignEffectSizes(mean=0, sd=1, diff=0.8, N=10, type="l", StdAdj=0)
# A tibble: 1 x 19
# phat varphat dfphat sigphat d vard sigd cor varcor sigCVt ttestp ES Variance StdES
# MedDiff transt
# <dbl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl> <dbl> <dbl> <lgl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
# 1 0.75 0.0152 17.5 FALSE 0.5 0.0624 FALSE 0.263 0.0175 FALSE 0.0958 2.41 9.01 0.802 2.32 0.0507
# ... with 3 more variables: EStrans <dbl>, StdEStrans <dbl>, VarTrans <dbl>
```

transformHgtoR

transformHgtoR

Description

The functions transforms a vector of Hedges g values to their equivalent point bi-serial values.

Usage

```
transformHgtoR(g, Nc, Nt)
```

Arguments

- `g`: A vector of Hegdes g values.
- `Nc`: A vector of numbers identifying the number of control condition participants in each group.
- `Nt`: A vector of numbers identifying the number of treatment condition participants in each group.

Value

value of point biserial r

Author(s)

Barbara Kitchenham and Lech Madeyski
Examples

```r
transformHgtoZr(0.4, 20, 20)
# [1] 0.1961161
```

Description

The functions transforms a vector of Hedges g values to their normal approximation of point biserial values.

Usage

```r
transformHgtoZr(g, Nc, Nt)
```

Arguments

- **g**: value of Hedges’ g
- **Nc**: the number of observations (participants) in the first (control) group
- **Nt**: the number of observations (participants) in the second (treatment) group

Value

value of normal approximation of point biserial r

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

```r
transformHgtoZr(0.5, 20, 20)
# [1] 0.2474665
```
Description

This function converts a vector of point bi-serial r values with associated sample size information back to the mean difference effect size Hedges g.

Usage

transformRtoHg(r, Nc, Nt)

Arguments

r
A vector of point bi-serial correlation values.

Nc
A vector of the number of observations in the control condition for the related experiments.

Nt
A vector of the number of observations in the treatment condition for the related experiments.

Value

value of Hedges' g

Author(s)

Barbara Kitchenham and Lech Madeyski

Examples

transformRtoHg(c(0.4, 0.2), c(20, 20), c(20, 20))
# [1] 0.8728716 0.4082483

transformRtoZr

Description

The function transforms a vector of point biserial r values to their normal approximation. It also works for the correlation r.

Usage

transformRtoZr(r)
**transformZrtoHg**

**Arguments**
- \( r \)  
  A vector of \( r \)-values

**Value**

value of normal approximation of point biserial \( r \)

**Author(s)**
Barbara Kitchenham and Lech Madeyski

**Examples**

```r
toiler::transformRtoZr(0.4)
# [1] 0.4236489
Zr=toiler::transformRtoZr(c(0.4,0.2))
Zr
# [1] 0.4236489 0.2027326
```

---

**transformZrtoHg**

**Description**

Transforms \( Z_r \) to Hedge’s \( g \).

**Usage**

```r
transformZrtoHg(Zr, Nc, Nt)
```

**Arguments**
- \( Z_r \)  
  the normal variate
- \( N_c \)  
  the number of observations (participants) in the first (control) group
- \( N_t \)  
  the number of observations (participants) in the second (treatment) group

**Value**

value of Hedges’ \( g \)

**Author(s)**
Barbara Kitchenham and Lech Madeyski

**Examples**

```r
transformZrtoHg(0.5, 20, 20)
# [1] 1.042191
```
transformZrtoHgapprox

Description
This function provides an approximate transformation from $Z_r$ to Hedges' $g$ when the number of observations in the treatment and control group are unknown. It is also used to allow the forest plots to display Hedge’s $g$ when they are based on $r$. It is necessary because the transformation function in the forest plot function does not allow any parameters other than effect size used. The function assumes that $N_c=N_t$ and gives the same results as transformZrtoHg when $N_c=N_t$.

Usage
transformZrtoHgapprox(Zr)

Arguments
Zr
A vector of normalised point bi-serial values

Value
approx. value of Hedges’ $g$

Author(s)
Barbara Kitchenham and Lech Madeyski

Examples
transformZrtoHgapprox(c(0.4,0.2))
# [1] 0.8215047 0.4026720

transformZrtoR

Description
The function transforms a vector of standardized normal variates to their equivalent $r$-values.

Usage
transformZrtoR(zr)

Arguments
zr
A vector of standard normal variates.
**Value**

value of point biserial r

**Author(s)**

Barbara Kitchenham and Lech Madeyski

**Examples**

```r
transformZrtoR(0.4236489)
# [1] 0.4
transformZrtoR(c(0.4236489, 0.2027326))
# [1] 0.4 0.2
```

**Description**

Function calculates the exact variance of a standardized effect size based on the relationship between \( t \) and the standardized effect size, see Morris and DeShon, Combining Effect Size Estimates in Meta-Analysis With Repeated Measures and Independent-Groups Designs, Psychological Methods, 7 (1), pp 105-125.

**Usage**

```r
varStandardizedEffectSize(d, A, f, returnVarg = TRUE)
```

**Arguments**

- **d**: An unadjusted standardized effect size
- **A**: The squared constant linking \( t \) and \( d \) i.e. \( t\sqrt{A} = d \)
- **f**: The degrees of freedom of the \( t \) value
- **returnVarg**: if set to TRUE return the variance of the small sample size adjusted standardized effect size \( (g) \), otherwise returns \( \text{var}(d) \) where \( d \) is the input parameter

**Value**

if `returnVarg` if set to TRUE, return `var(g)` otherwise `var(d)`

**Author(s)**

Barbara Kitchenham and Lech Madeyski
Examples

d=0.5
varStandardizedEffectSize(d,sqrt(1/30),29,returnVarg=FALSE)
# [1] 0.2007699
varStandardizedEffectSize(d,sqrt(1/30),29,returnVarg=TRUE)
# [1] 0.1904167
Index

* datasets
Ciolkowski09ESEM.MetaAnalysis.PBRvsCBRorAR, 17
KitchenhamEtAl.CorrelationsAmongParticipants.Abrahao13TSE, 36
KitchenhamEtAl.CorrelationsAmongParticipants.Gravino15JVLC, 38
KitchenhamEtAl.CorrelationsAmongParticipants.Madeyski10, 39
KitchenhamEtAl.CorrelationsAmongParticipants.Reggio15SSM, 41
KitchenhamEtAl.CorrelationsAmongParticipants.Ricca10TSE, 42
KitchenhamEtAl.CorrelationsAmongParticipants.Ricca14TOSEM, 44
KitchenhamEtAl.CorrelationsAmongParticipants.Romano18ESEM, 45
KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14EASE, 47
KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14JVLC, 48
KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello14TOSEM, 50
KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello15EMSE, 51
KitchenhamEtAl.CorrelationsAmongParticipants.Scanniello17TOSEM, 53
KitchenhamEtAl.CorrelationsAmongParticipants.Torchiano17JVLC, 55
KitchenhamMadeyskiSimulatedCrossoverDataSets, 62
KitchenhamMadeyskiBrereton.MetaAnalysisReportedResults, 64
KitchenhamMadeyskiBrereton.ReportedEffectSizes, 66
KitchenhamMadeyskiBudgen16.COCOMO, 68
KitchenhamMadeyskiBudgen16.DiffInDiffData, 69
KitchenhamMadeyskiBudgen16.PolishData, 70
KitchenhamMadeyskiBudgen16.PolishSubjects, 71
KitchenhamMadeyskiBudgen16.SubjectData, 72
Madeyski15EISEJ.OpenProjects, 74
Madeyski15EISEJ.PropProjects, 75
Madeyski15EISEJ.StudProjects, 76
Madeyski15SQJ.NDC, 77
MadeyskiKitchenham.EUBASdata, 78
MadeyskiKitchenham.MetaAnalysis.PBRvsCBRorAR, 79
MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20190324, 81
MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20191022, 82
aggregateIndividualDocumentStatistics, 4
boxplotAndDensityCurveOnHistogram, 5
boxplotHV, 6
calc4GroupNPStats, 6
calculateBasicStatistics, 7
calculateGroupSummaryStatistics, 8
calculateHg, 9
calculatePhat, 12
CalculateLevel2ExperimentRData, 10
123
INDEX

CalculateRLevel1, 13
CalculateSmallSampleSizeAdjustment, 15
CalculateTheoreticalEffectSizes, 16
Ciolkowski09ESEM.MetaAnalysis.PBRvsCBRorAR, 17
Cliffd, 18
constructEffectSizes, 19
ConstructLevel1ExperimentRData, 20
densityCurveOnHistogram, 22
effectSizeCI, 23
ExtractExperimentData, 24
ExtractGroupSizeData, 25
ExtractMAStatistics, 27
ExtractSummaryStatisticsRandomizedExp,
fmt, 30
getEffectSizesABBA, 31
getEffectSizesABBAIgnoringPeriodEffect, 32
getSimulationData, 33
getTheoreticalEffectSizeVariancesABBA,
Kendalltaupb, 35
KitchenhamEtAl.CalculationsCorrelationsAmongParticipants.Scanniello17TOSEM, 52
KitchenhamEtAl.CalculationsCorrelationsAmongParticipants.Torchiano17JVLC, 54
KitchenhamMadeyski.SimulatedCrossoverDataSets, 56
KitchenhamMadeyskiBrereton.ABBAMetaAnalysisReportedResults, 57
KitchenhamMadeyskiBrereton.ABBAReportedEffectSizes, 58
KitchenhamMadeyskiBrereton.DocData, 59
KitchenhamMadeyskiBrereton.ExpData, 61
KitchenhamMadeyskiBrereton.MetaAnalysisReportedResults, 62
KitchenhamMadeyskiBrereton.ReportedEffectSizes, 63
KitchenhamMadeyskiBudgen16.COCOMO, 64
KitchenhamMadeyskiBudgen16.DiffInDiffData, 66
KitchenhamMadeyskiBudgen16.FINNISH, 67
KitchenhamMadeyskiBudgen16.PolishData, 68
KitchenhamMadeyskiBudgen16.PolishSubjects, 69
KitchenhamMadeyskiBudgen16.SubjectData, 71
Klapa10TSMP, 73
Madeyski15EISEJ.OpenProjects, 74
Madeyski15EISEJ.PropProjects, 75
Madeyski15EISEJ.StudProjects, 76
Madeyski15SQJ.NDC, 77
Madeyski16EISEJ.EUBASdata, 78
MadeyskiKitchenham.MetaAnalysis.PBRvsCBRorAR, 79
MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20190324, 81
MadeyskiLewowski.IndustryRelevantGitHubJavaProjects20191022, 82
MetaAnalysisSimulations, 84
NP2GroupMetaAnalysisSimulation, 87
NP4GroupMetaAnalysisSimulation, 89
percentageInaccuracyOfLargeSampleVarianceApproximation, 92
plotOutcomesForIndividualsInEachSequenceGroup, 93
INDEX

PrepareForMetaAnalysisGtoR, 94
printXTable, 95
proportionOfSignificantTValuesUsingCorrectAnalysis, 96
proportionOfSignificantTValuesUsingIncorrectAnalysis, 97
RandomExperimentSimulations, 97
RandomizedBlockDesignEffectSizes, 99
RandomizedBlocksAnalysis, 101
RandomizedBlocksExperimentSimulations, 102
RandomizedDesignEffectSizes, 104
readExcelSheet, 105
reproduceForestPlotRandomEffects, 105
reproduceMixedEffectsAnalysisWithEstimatedVarianceAndExperimentalDesignModerator, 106
reproduceMixedEffectsAnalysisWithExperimentalDesignModerator, 106
reproduceMixedEffectsForestPlotWithExperimentalDesignModerator, 107
reproduceSimulationResultsBasedOn500Reps1000Obs, 107
reproduceTablesOfPaperMetaAnalysisForFamiliesOfExperiments, 108
reproduceTableWithEffectSizesBasedOnMeanDifferences, 109
reproduceTableWithPossibleModeratingFactors, 109
reproduceTableWithSourceDataByCiolkowski, 110
rSimulations, 110
searchForIndustryRelevantGitHubProjects, 112
simulateRandomizedBlockDesignEffectSizes, 113
simulateRandomizedDesignEffectSizes, 115
transformHgtoR, 116
transformHgtoZr, 117
transformRtoHg, 118
transformRtoZr, 118
transformZrtoHg, 119
transformZrtoHgapprox, 120
transformZrtoR, 120
varStandardizedEffectSize, 121