Package 'CRTSize'

October 12, 2022

Version 1.1

Date 2022-10-05

Title Sample Size Estimation Functions for Cluster Randomized Trials

Author Michael A Rotondi <mrotondi@yorku.ca>

Maintainer Michael A Rotondi <mrotondi@yorku.ca>

Depends R (>= 2.01)

Description Sample size estimation in cluster (group) randomized trials. Contains traditional power-based methods, empirical smoothing (Rotondi and Donner, 2009), and updated meta-analysis techniques (Rotondi and Donner, 2012).

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2022-10-12 14:02:32 UTC

R topics documented:

`fixedMetaAnalMD` *Fixed Effects Meta-Analysis for Mean Differences*

Description

This function performs a fixed effects meta-analysis of mean differences (a continuous effect measure) from a matrix of mean differences, lower and upper confidence limits.

Usage

```
fixedMetaAnalMD(data, alpha=0.05)
```

Arguments

`data`	A matrix with individual studies in each row. The first column contains the estimate of the Mean Difference (Treatment - Control), the second column contains the lower 95 % confidence limit and the third contains the 95 % upper limit. For simplicity, this function assumes that each of the inputed lower and upper limits that are 95 % confidence limits.
`alpha`	The desired type I error rate for calculation of confidence limits for the fixed effects mean difference.

Details

This function performs an elementary fixed effects meta-analysis of continous outcome measures from an input matrix of mean differences and their respective confidence limits for any number of studies. This function is designed for use with `n4meansMeta` to provide sample size estimation based on an updated meta-analysis. Note that it is assumed that the 95 % confidence limits are correctly calculated, and that this function, being elementary in nature only accepts the mean difference as the effect measure. Additional utilities for a meta-analysis or meta-regression can be found in Viechbauer (2006).

Value

`data`	The data matrix is returned.
`thetaF`	The fixed effects outcome measure, the mean difference.
`lF`	The 100(1 - α) % lower limit of the pooled mean difference.
`uF`	The 100(1 - α) % upper limit of the pooled mean difference.
`Var`	The variance of the fixed effects treatment mean difference.
`Sig`	Does this show a statistically significant benefit or harm? (Binary: zero (non-significant) or one (significant)).
`alpha`	The desired type I error rate.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

White IR and Thomas J. Standardized mean differences in individually-randomized and cluster-randomized trials, with applications to meta-analysis. Clinical Trials 2005. 2:141-151.

Viechtbauer, W. Metafor: A meta-analysis package for R. 2010. Available at: http://www.metafor-project.org/

See Also

`fixedMetaAnalRROR`, `n4meansMeta`

Examples

```
fixedMetaAnalMD(data=rbind(c(100, 50, 150), c(25, -100, 150), c(-90, -190, 10),
c(-125, -200, -50)), alpha=0.05);
```

`fixedMetaAnalRROR` *Fixed Effects Meta-Analysis for Relative Risks/Odds Ratios*

Description

This function provides a detailed fixed effects meta-analysis of relative risks/odds ratios from a matrix of effect measures, lower and upper confidence limits.

Usage

```
fixedMetaAnalRROR(data, alpha=0.05)
```

Arguments

`data`	A matrix with individual studies in each row. The first column contains the estimate of the relative risk/odds ratio, second column contains the 95 % lower limit and the third contains the 95 % upper limit.
`alpha`	The desired type I error rate for calculation of confidence limits for the pooled fixed effects measure. For simplicity, this function assumes that each of the inputed lower and upper limits are 95 % confidence limits.

Details

This function performs an elementary fixed effects meta-analysis of relative risks/odds ratios from an input matrix of relative risks and their respective confidence limits for any number of studies. This function is designed to be used in conjunction with `n4propsMeta` in order to provide a sample size calculation based on an updated meta-analysis. Note that this function works on the log scale to calculate variances and the pooled effect measure. Additional utilities for a meta-analysis or meta-regression can be found in Viechbauer (2006).

Value

data	The data matrix is returned.
thetaF	The fixed effects Log Relative Risk (RR)/Odds Ratio (OR).
lF	The 100(1 - α) % lower limit of the pooled Log RR/OR.
uF	The 100(1 - α) % upper limit of the pooled Log RR/OR.
Var	The variance of the Log RR/OR.
Sig	Is the result statistically significant (Binary zero or one)?
alpha	The desired type I error rate.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

Viechtbauer, W. Metafor: A meta-analysis package for R. 2010. Available at: http://www.metafor-project.org/

See Also

fixedMetaAnalMD, n4propsMeta

Examples

```
fixedMetaAnalRROR(data=rbind(c(0.672, 0.342, 1.321), c(0.942, 0.761, 1.165),
c(1.228, 0.384, 3.934)), alpha=0.05);
```

---

n4incidence *Number of Subjects Required for a Cluster Randomized Trial Comparing Incidence Rates*

Description

This function provides detailed sample size estimation information to determine the number of subjects that must be enrolled in a cluster randomized trial to test for a significant difference in incidence rates.

Usage

```
n4incidence(le, lc, m, t, CV, alpha=0.05, power = 0.80, AR=1, two.tailed=TRUE, digits=3)
```

Arguments

le	The anticipated incidence rate, λ_E, in the experimental group with the outcome.
lc	The anticipated incidence rate, λ_C, in the control group with the outcome.
m	The anticipated average (or actual) cluster size.
t	The planned follow-up time for the study (in weeks, months, etc.)
CV	The coefficient of variation, assumed constant over both the treatment and control groups. Note that CV = $\sigma_1/\lambda_E = \sigma_2/\lambda_C$, where σ_E and σ_C represent the between-cluster variation in incidence rates for each group.
AR	The Allocation Ratio: AR=1 implies an equal number of subjects per treatment and control group (maximum efficiency), > 1, implies more subjects will be enrolled in the control group (e.g. in the case of costly intervention), < 1 implies more subjects in the tretment group (rarely used).
alpha	The desired type I error rate.
power	The desired level of power, recall power = 1 - type II error.
two.tailed	Logical, If TRUE calculations are based on a two-tailed type I error, if FALSE, a one-sided calculation is performed.
digits	Number of digits to round calculations.

Details

This function provides detailed information, similar to PROC POWER in SAS, but with less functionality and more concise output. It is used for sample size estimation in a cluster randomized trial where the outcome of interest is an incidence rate. A simple example may include whether a new treatment can successfully reduce the incidence of heart attacks over a six month period. In epidemiological terms, λ_E and λ_C are the expected incidence rate of the outcome in the experimental and control group. Note that if the results suggest a small number of clusters is required, an iterative procedure will include the T distribution instead of the normal critical value for alpha, iterating until convergence. In some cases, such as small ICC values, the algorithm may fail to converge and may need to be stopped.

Value

nE	The minimum number of clusters required in the experimental group.
nC	The minimum number of clusters required in the control group.
le	The anticipated incidence rate, λ_E, in the experimental group with the outcome.
lc	The anticipated incidence rate, λ_C, in the control group with the outcome.
m	The anticipated average (or actual) cluster size.
t	The planned follow-up time for the study.
CV	The coefficient of variation.
AR	The Allocation Ratio: One implies an equal number of subjects per treatment and control groups.
alpha	The desired type I error rate.
power	The desired level of power.
AR	The Allocation Ratio.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

See Also

n4means, n4props

Examples

```
## Not run:
Suppose a new drug is thought to reduce the incidence of HIV from 0.01 per person-year to 0.005
per person-year. Assume the coefficient of variation is 0.25 and that 1000 subjects will be
followed for a two year period.  Calculate the required number of subjects that must be enrolled
in a study to detect this difference with alpha = 0.05 and power = 0.80.

## End(Not run)
n4incidence(le=0.01, lc=0.005, m=1000, t=2, CV=0.25);
```

---

n4means *Number of Subjects Required for a Cluster Randomized Trial with a Continuous Outcome*

Description

This function provides detailed sample size estimation information to determine the number of subjects that must be enrolled in a cluster randomized trial to compare two means.

Usage

```
n4means(delta, sigma, m, ICC, alpha=0.05, power=0.8, AR=1, two.tailed=TRUE, digits=3)
```

Arguments

delta	The minimum detectable difference between population means.
sigma	The standard deviation of the outcome.
m	The anticipated average (or actual) cluster size.
ICC	The anticipated value of the intraclass correlation coefficient, ρ.

AR	The Allocation Ratio: AR=1 implies an equal number of subjects per treatment and control group (maximum efficiency), AR > 1, implies more subjects will be enrolled in the control group (e.g. in the case of costly intervention), AR < 1 implies more subjects in the treatment group (rarely used).
alpha	The desired type I error rate.
power	The desired level of power, recall power = 1 - type II error.
two.tailed	Logical, If TRUE calculations are based on a two-tailed type I error, if FALSE, a one-sided calculation is performed.
digits	Number of digits to round calculations.

Details

This function provides detailed sample size information, similar to PROC POWER in SAS, but with less functionality and more concise output, and adapted for the design of cluster randomized trial. It is used for sample size estimation in a cluster randomized trial where the outcome is continuous, e.g. blood pressure, or weight. Note that if the results suggest a small number of clusters is required, an iterative procedure will include the T distribution instead of the normal critical value for alpha, iterating until convergence. In some cases, such as small ICC values, the algorithm may fail to converge and may need to be stopped.

Value

nE	The minimum number of clusters required in the experimental group.
nC	The minimum number of clusters required in the control group.
delta	The minimum detectable difference between population means.
sigma	The standard deviation of the outcome.
alpha	The desired type I error rate.
power	The desired level of power, recall power = 1 - type II error.
AR	The Allocation Ratio.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

See Also

n4props, n4incidence

Examples

```
## Not run: Suppose we wish to test whether a blood pressure medication reduces diastolic blood
pressure by 10 mm Hg, at standard significance and power, assume the standard deviation is 10 mm Hg.
## End(Not run)
n4means(delta=10, sigma=1, m=25, ICC=0.05, alpha=0.05, power=0.80);
```

`n4meansEB` *Number of Subjects Required for a Cluster Randomized Trial with a Continuous Outcome Using Empirical Smoothing*

Description

This function provides detailed sample size estimation information to determine the required number of clusters that must be enrolled in a cluster randomized trial using the empirical smoothing density for the ICC. The method applies a smoothed density function (including optional weighting) to obtain an empirical distribution for the ICC. Output includes quantiles of values of the required number of clusters to obtain a prespecified power level. This version assumes the outcome of interest is continuous.

Usage

```
n4meansEB(ICC, varICC=0, delta, from, to, sigma, m, iter=1000, alpha=0.05, power=0.8,
two.tailed=TRUE, digits=3, plot=TRUE)
```

Arguments

`ICC`	A vector of possible ICC values, obtained from a reasonable number of independent studies. These values form the basis of the empirical density function for the ICC.
`varICC`	A vector of variances of the ICC estimates. In some cases, it may be desirable to give greater weight (smaller variances) to estimates of the ICC that are obtained from larger samples. The default value is zero, which implies that all estimates are weighted equally.
`delta`	The anticipated mean difference in the planned study.
`from`	A lower limit representing the lowest plausible value for the ICC. This is used in the estimation of the ICC's empirical density function. The default value is zero as the ICC is assumed to be non-negative.
`to`	An upper bound for the plausible range of the ICC. A default value is not specified as this may range depending on the circumstances.
`sigma`	The anticipated variance in each group.
`m`	The anticipated average (or actual) cluster size.
`iter`	The total number of iterations.
`alpha`	The desired type I error rate.
`power`	The desired level of power, recall power = 1 - type II error.

two.tailed	Logical, If TRUE calculations are based on a two-tailed type I error, if FALSE, a one-sided calculation is performed.
digits	Number of digits to round calculations.
plot	Logical: Would you like a plot of the estimated density of the ICC and Histogram?

Details

This function estimates an empirical density for the ICC using the Gaussian kernel. Weights can be incorporated through the use of the varICC parameter. Values are sampled from this empirical density a large (iter) number of times and the resulting number of clusters that must be randomized to achieve a pre-specified power level is then calculated. The resulting output is the quantiles of the required number of clusters, illustrating the most likely values of the ICC and number of clusters required. Additional details are in Rotondi and Donner (2009).

Value

ResRho	A vector of values of sampled values of the ICC. This is of length iter.
ResK	A vector of values of the required number of clusters k, using the ICC values in ResRho. This is also of length iter.
pe	The anticipated proportion of individuals in the experimental group with the outcome.
pc	The anticipated proportion of individuals in the control group with the outcome.
ICC	The specified vector of values for the ICC.
varICC	A vector of variances of the ICC (study weights).
from	Lower bound in the ICC density estimation. Default of zero.
to	Upper bound in ICC Density Estimation.
m	The size of each cluster.
alpha	The desired type I error rate.
power	The desired level of power, recall power = 1 - type II error.
two.tailed	TRUE or FALSE; Depending on whether the alpha level is one or two sided.
digits	Number of digits to round results.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

Rotondi M and Donner A. (2009) Sample Size Estimation in Cluster Randomized Trials: An Empirical Bayes Approach, Journal of Educational and Behavioral Statistics, 34:229-237.

See Also

n4propsEB

Examples

```
## Not run: ICC values are from Rotondi and Donner (2009). Suppose classrooms of size 25 are randomized
with hypothetical experimental rates of 0.05 and control rates of 0.18. Plots are suppressed,
and iter = 50 for testing purposes.
## End(Not run)
n4meansEB(delta=0.5, sigma=1, m=25, ICC=c(0.162, 0.205, 0.234, 0.253),
varICC= c(0.030, 0.032, 0.010, 0.026)^2, from=0.15, to=0.28, iter=50, plot=FALSE);
```

---

n4meansMeta *Empirical Power and Variance Reduction for an Updated Fixed Effects Meta-Analysis in Cluster Randomized Trials*

Description

This function provides the empirical power/reduction in variance in an updated meta-analysis for a vector of number of clusters to randomize per group and a vector of estimates of the ICC.

Usage

```
n4meansMeta(data, k, ICC, ICCDistn="unif", lower=0, upper=0.25, varRed=FALSE, m, sdm,
meanC, sdC, sdT=sdC, iter=1000, alpha=0.05)
```

Arguments

data	A matrix of individual studies (each row). The first column contains the estimate of the relative risks, second column contains the lower 95 % confidence limit and the third contains the upper 95 % confidence limit.
k	A vector of potential number of clusters to randomize to each of the treatment and control groups. Note that this function assumes an equal allocation to treatment and control group status.
ICC	A vector of potential values of the ICC, these can be obtained from the cluster trials themselves, or from the literature.
ICCDistn	The hypothetical distribution of the ICC values. This can be set to "fixed" (note that only one ICC value is accepted for this option), "unif" on the range [lower, upper], "normal", corresponding to the truncated normal distribution (Turner et al, 2004), and "smooth" corresponding to the empirical smoothing option. (Rotondi and Donner, 2009)
lower	The lower bound for the smoothing or unif options. Default value is zero.
upper	The upper bound for the smoothing or unif options. Default value is 0.25.
varRed	Logical; If varRed is set to TRUE, the proportionate reduction of variance is displayed for the fixed effects meta-analysis.

m	The mean cluster size.
sdm	The standard deviation of the mean cluster size. This adds additional real-world variation in the simulated study.
meanC	The anticipated mean response level in the control group. The anticipated treatment mean is calculated from the simulated effect size of the preliminary meta-analysis.
sdC	The standard deviation of the control rate. This adds real-world variation in the simulated study and can be precise or imprecise depending on the investigators preference.
sdT	The standard deviation of the treatment rate. By default, this is set to the same sdC.
iter	The number of iterations for each value of k and the ICC. This has a large impact on computational time. Default is 1000.
alpha	The desired type I error rate for calculation of confidence limits for the fixed effects level. For simplicity, this function assumes that each of the inputed lower and upper limits that are 95 % confidence limits.

Details

This function calculates the empirical power of an updated meta-analysis by a generalization of Sutton et al. (2007). The procedure is summarized in the accompanying manuscript. In short, a hypothetical new study of a given size is simulated, then added to the meta-analysis. The results are re-meta-analyzed and it is verified whether the result is statistically significant. Note that the proportion of variance reduction and power may not always strictly decrease with k, as the simulation exhibits individual-level variation.

Value

power	The power of the updated meta-analysis. Presented as a matrix of number of clusters by ICC values.
data	The data matrix is returned.
newMean	The preliminary fixed effects mean difference.
newVar	The variance of the preliminary fixed effects mean difference (MD).
lF	The 100(1 - α) % lower limit of the MD in the Original meta-analysis.
uF	The 100(1 - α) % upper limit of the MD in the Original meta-analysis.
Var	The variance of the effect measure.
Sig	Is the result statistically significant (Binary zero or one).
k	The number of clusters randomized per group (vector).
ICC	A vector of ICC values.
ICCDistn	The distributional assumption about the ICC.
varRed	Variance Reduction: Logical.
varianceReduction	The proportionate reduction in variance for the number of clusters in the fixed effects meta-analysis.

m	The mean cluster size.
sdm	The standard deviation of the mean cluster size.
meanC	The control mean.
sdC	The standard deviation of the control mean.
alpha	The desired type I error rate.
iter	The total number of iterations.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

Sutton AJ et al. (2007) Evidence-based sample size calculations based upon updated meta-analysis. Statistics in Medicine, 26(12):2479-2500.

Turner R et al. (2004) Allowing for imprecision in the intracluster correlation coefficient in the design of cluster randomized trials. Statistics in Medicine, 23(8):1195-1214.

Rotondi M and Donner A. (2009) Sample Size Estimation in Cluster Randomized Trials: An Empirical Bayes Approach. Journal of Educational and Behavioral Statistics. DOI: 10.3102/1076998609332756.

Rotondi M and Donner A. (2012) Sample Size Estimation in Cluster Randomized Trials: An Evidence-Based Perspective. Computational Statistics and Data Analysis 56:1174-1187.

See Also

fixedMetaAnalMD, n4propsMeta

Examples

```
## Not run: A brief example with 5 iterations.
n4meansMeta(data=rbind(c(100, 50, 150), c(25, -100, 150), c(-90, -190, 10), c(-125, -200, -50)),
k=c(10, 20), ICC=c(0.1, 0.15, 0.18), m=100, sdm=0, meanC=100, sdC=10, iter=5, alpha=0.05,
varRed=TRUE, ICCDistn="smooth");
```

n4props *Number of Subjects Required for a Cluster Randomized Trial with a Binary Outcome*

Description

This function provides detailed sample size estimation information to determine the number of subjects that must be enrolled in a cluster randomized trial with a binary outcome.

Usage

```
n4props(pe, pc, m, ICC, alpha=0.05, power = 0.80, AR=1, two.tailed=TRUE, digits=3)
```

Arguments

pe	The anticipated proportion of individuals with the outcome of interest in the experimental group.
pc	The anticipated proportion of individuals with the outcome of interest in the control group.
m	The anticipated average (or actual) cluster size.
ICC	The anticipated value of the intraclass correlation coefficient, ICC.
AR	The Allocation Ratio: AR$=$1 implies an equal number of subjects per treatment and control group (maximum efficiency), AR $>$ 1, implies more subjects will be enrolled in the control group (e.g. in the case of costly intervention), AR $<$ 1 implies more subjects in the treatment group (rarely used).
alpha	The desired type I error rate.
power	The desired level of power, recall power = 1 - type II error.
two.tailed	Logical, If TRUE calculations are based on a two-tailed type I error, if FALSE, a one-sided calculation is performed.
digits	Number of digits to round calculations

Details

This function provides detailed information, similar to PROC POWER in SAS, but with less functionality and more concise output. It is used for sample size estimation in a cluster randomized trial where the outcome of interest is binary. A simple example may include whether an individual dies from a heart attack. In epidemiological terms, pe and pc can be thought of as the expected prevalence of the outcome in the experimental and control group. Note that if the results suggest a small number of clusters is required, an iterative procedure will include the T distribution instead of the normal critical value for alpha, iterating the procedure until convergence. Thus on some occasions, the algorithm may not converge. In some cases, such as small ICC values or proportions, this fails to converge and may need to be stopped.

Value

nE	The minimum number of clusters required in the experimental group.
nC	The minimum number of clusters required in the control group.
pe	The anticipated proportion of individuals in the experimental group with the outcome.
pc	The anticipated proportion of individuals in the control group with the outcome.
alpha	The desired type I error rate.
power	The desired level of power, recall power = 1 - type II error.
AR	The Allocation Ratio.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

See Also

n4means, n4incidence

Examples

```
## Not run: Suppose a new drug is thought to reduce heart attack mortality from 0.10 to 0.03.
Calculate the required number of subjects that must be enrolled in a study to detect this
difference with alpha = 0.05 and power = 0.80.
## End(Not run)
n4props(pe=0.03, pc=0.10, m=25, ICC=0.20, AR=1, alpha=0.05, power=0.80);
```

---

n4propsEB *Number of Subjects Required for a Cluster Randomized Trial with Binary Outcomes Using Empirical Smoothing*

Description

This function provides detailed sample size estimation information to determine the required number of clusters that must be enrolled in a cluster randomized trial using the empirical smoothing model. The method applies a smoothed density function (including optional weighting) to obtain an empirical distribution for the ICC. Output includes quantiles of values of the required number of clusters to obtain a prespecified power level.

Usage

```
n4propsEB(ICC, varICC=0, from=0, to, pe, pc, m, iter=1000, alpha=0.05,
power=0.8, two.tailed=TRUE, digits=3, plot=TRUE)
```

Arguments

ICC	A vector of possible ICC values, obtained from a reasonable number of independent studies. These values form the basis of the empirical density function for the ICC.
varICC	A vector of variances of the estimates of the ICC. In some cases, it may be desirable to give greater weight (smaller variances) to estimates of the ICC that are obtained from larger samples. The default value is zero, which implies that all estimates are weighted equally.
from	A lower limit representing the lowest plausible value for the ICC. This is used in the estimation of the ICC's empirical density function. The default value is zero as the ICC is assumed to be non-negative.
to	An upper bound for the plausible range of the ICC. A default value is not specified as this may range depending on the circumstances.
pe	The anticipated proportion of individuals in the experimental group with the outcome.
pc	The anticipated proportion of individuals in the control group with the outcome.
m	The anticipated average (or actual) cluster size.
iter	The total number of iterations.
alpha	The desired type I error rate.
power	The desired level of power, recall power = 1 - type II error.
two.tailed	Logical, If TRUE calculations are based on a two-tailed type I error, if FALSE, a one-sided calculation is performed.
digits	Number of digits to round calculations.
plot	Logical: Would you like a plot of the estimated density of the ICC and Histogram?

Details

This function estimates an empirical density for the ICC using a Gaussian kernel. Weights can be incorporated through specification of the varICC parameter. ICC values are sampled from this empirical density a large (iter) number of times and the resulting number of clusters that must be randomized to achieve a pre-specified power level is then calculated. The resulting output is the quantiles of the required number of clusters, illustrating the most likely values of the ICC and number of clusters required. Additional details are in Rotondi and Donner (2009).

Value

ResRho	A vector of values of sampled values of the ICC. This is of length iter.
ResK	A vector of values of the calculated required number of clusters k, using the ICC values in ResRho. This is also of length iter.

pe	The anticipated proportion of individuals in the experimental group with the outcome.
pc	The anticipated proportion of individuals in the control group with the outcome.
ICC	The specified vector of values for the ICC.
varICC	A vector of variances of the ICC, these can be used as study weights.
from	Lower bound in ICC Density Estimation. Default of zero.
to	Upper bound in ICC Density Estimation. Default of zero.
m	The size of each cluster
alpha	The desired type I error rate.
power	The desired level of power, recall power = 1 - type II error.
two.tailed	TRUE or FALSE; Depending on whether the alpha level is one or two sided.
digits	Number of digits to round results.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

Rotondi M and Donner A. (2009) Sample Size Estimation in Cluster Randomized Trials: An Empirical Bayes Approach. Journal of Educational and Behavioral Statistics, 34:229-237.

See Also

n4meansEB

Examples

```
## Not run: ICC values are from Rotondi and Donner (2009).  Suppose classrooms of size 25
are randomized with hypothetical experimental rates of 0.05 and control rates of 0.18.
Plots are suppressed, and iter = 50 for testing purposes.
## End(Not run)
n4propsEB(pe=0.10, pc=0.18, m=25, ICC=c(0.162, 0.205, 0.234, 0.253),
varICC= c(0.030, 0.032, 0.010, 0.026)^2, from=0.15, to=0.28, iter=50, plot=FALSE);
```

n4propsMeta *Empirical Power and Variance Reduction of an Updated Fixed Effects Meta-Analysis with Binary Outcomes*

Description

This function provides the empirical power/variance reduction of an updated meta-analysis for a vector of the number of clusters to randomize per group and a vector of estimates of the ICC with a binary outcome measured using the (log) relative risk or odds ratio.

Usage

```
n4propsMeta(data, k, ICC, ICCDistn="unif", lower=0, upper=0.25, varRed=FALSE,
m, sdm, pC, sdpC, iter=1000, alpha=0.05, RR=TRUE)
```

Arguments

data	A matrix with completed studies in each row. The first column contains the estimate of the relative risk or odds ratio, the second column contains the 95 % lower limit and the third contains the 95 % upper limit.
k	A vector of the potential number of clusters to randomize to each of the treatment and control groups. Note that this function assumes an equal allocation to treatment and control group status.
ICC	A vector of potential values of the ICC, these can be obtained from the cluster trials themselves, or from the literature.
ICCDistn	The hypothetical distribution of the ICC values. This can be set to "fixed" (note that only one ICC value is accepted for this option), "unif" on the range [lower, upper], "normal", corresponding to the truncated normal distribution (Turner et al, 2004), and "smooth" corresponding to the empirical smoothing option (Rotondi and Donner, 2009).
lower	The lower bound for the smoothing or unif options. Default value is zero.
upper	The upper bound for the smoothing or unif options. Default value is 0.25.
varRed	Logical; If varRed is set to TRUE, the proportionate reduction of variance is displayed for the fixed effects meta-analysis.
m	The mean cluster size.
sdm	The standard deviation of the mean cluster size. This adds additional real-world variation in the simulated study, using a normal model for large cluster sizes.
pC	The anticipated event in the control group. The anticipated treatment event is calculated from the simulated effect size of the preliminary meta-analysis.
sdpC	The standard deviation of the control rate. This is to generate real-world variation in the simulated study and can be precise or imprecise depending on the investigators preference.
iter	The number of iterations for each value of k and the ICC. This has a large impact on computational time. Default is 1000.

alpha	The desired type I error rate for calculation of confidence limits for the fixed effects level. For simplicity, this function assumes that each of the inputed lower and upper limits that are 95 % confidence limits.
RR	Logical; Are the effect measures Relative Risks (TRUE) (default) or Odds Ratios (FALSE)? This is necessary only for the calculation of variances.

Details

This function calculates the empirical power of an updated meta-analysis by a generalization of Sutton et al. (2007) to the context of cluster randomized trials with a binary outcome. The procedure is summarized in the accompanying manuscript. In short, a hypothetical new study of a given size is simulated, then added to the meta-analysis. The results are re-meta-analyzed and it is verified whether the pooled result is statistically significant, or the appropriate reduction in variance of the pooled effect measure is recorded. Note that the proportion of variance reduction and power may not always (strictly) decrease with k, as the simulation exhibits individual-level variation.

Value

power	The power of the updated meta-analysis. Presented as a vector corresponding to the number of clusters.
varianceReduction	The proportionate reduction in variance for the number of clusters in the fixed effects meta-analysis.
m	The mean cluster size.
data	The data matrix is returned.
newMean	The preliminary fixed Effects log relative risk (RR) or odds ratio (OR).
newVar	The variance of the preliminary fixed effects log RR or log OR.
lF	The 100(1 - α) % lower limit of the log RR/log OR in the original meta-analysis.
uF	The 100(1 - α) % upper limit of the log RR/log OR in the original meta-analysis.
Var	The variance of the overall log RR/log OR.
k	The number of clusters randomized per group (vector).
ICC	A vector of ICC values.
ICCDistn	The distributional assumption about the ICC.
varRed	Variance Reduction: Logical.
sdm	The standard deviation of the mean cluster size.
pC	The mean control rate.
sdpC	The standard deviation of the control rate.
alpha	The desired type I error rate.
iter	The total number of iterations.

Author(s)

Michael Rotondi, <mrotondi@yorku.ca>

References

Matthews JNS. Introduction to Randomized Controlled Clinical Trials (2nd Ed.) Chapman & Hall: New York, 2006.

Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold: London, 2000.

Sutton AJ et al. (2007) Evidence-based sample size calculations based upon updated meta-analysis. Statistics in Medicine, 26(12):2479-2500.

Turner R et al. (2004) Allowing for imprecision in the intracluster correlation coefficient in the design of cluster randomized trials. Statistics in Medicine, 23(8):1195-1214.

Rotondi M and Donner A. (2009) Sample Size Estimation in Cluster Randomized Trials: An Empirical Bayes Approach. Journal of Educational and Behavioral Statistics, 34:229-237.

Rotondi M and Donner A. (2011) Sample Size Estimation in Cluster Randomized Trials: An Evidence-Based Perspective. Computational Statistics and Data Analysis 56:1174-1187.

See Also

`fixedMetaAnalRROR`, `n4meansMeta`

Examples

```
## Not run: A brief example with 10 iterations.
n4propsMeta(data=rbind(c(0.672, 0.342, 1.321), c(0.942, 0.761, 1.165), c(1.228, 0.384, 3.934)),
k=c(10, 20, 30), ICC=c(0.1, 0.15, 0.14), m=10, sdm=0, pC=0.1, sdpC=0, iter=10, alpha=0.05,
varRed=TRUE, ICCDistn="unif");
```

Index