Package

July 17, 2019

Title Calculate Power and Sample Size with Beta Regression.

Version 1.0-1

Date 2019-07-10

Description Power calculations are a critical component of any research study to determine the minimum sample size necessary to detect differences between multiple groups. Researchers often work with data taking the form of proportions that can be modeled with a beta distribution. Here we present an R package, BetaPASS, and analogous SAS macro, that perform power and sample size calculations for data following a beta distribution with comparative nonparametric output. This package allows flexibility with multiple options for link functions to fit the data and graphing functionality for visual comparisons.

Depends R (>= 3.5.1)

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.1.0

Imports betareg, lmtest, reshape, ggplot2, Rcpp, stats

Suggests knitr,
rmarkdown

VignetteBuilder knitr

R topics documented:

betapower *Find Power with Beta distribution*

Description

Find the power for a given sample size when testing the null hypothesis that the means for the control and treatment groups are equal against a two-sided alternative.

Usage

```
betapower(mu0, sd0, mu1.start, mu1.end = NULL, mu1.by = NULL,
ss.start, ss.end = NULL, ss.by = NULL, sig.level = 0.05,
trials = 100, seed = 1, link.type="logit",
equal.precision=TRUE, sd1 = NULL)
```

Arguments

`mu0`	the mean for the control group
`sd0`	the standard deviation for the control group
`mu1.start`	the starting value of mean for the treatment group under the alternative mu1
`mu1.end`	the ending value of mean for the treatment group under the alternative mu1
`mu1.by`	the step length of mean for the treatment group under the alternative mu1
`ss.start`	the starting value of sample size
`ss.end`	the ending value of sample size
`ss.by`	the step length of sample size
`sig.level`	significant level; default value is 0.05
`trials`	the number of trials
`seed`	the seed used in the simulation
`link.type`	the type of link used in the beta regression. Default value is "logit", or you can use "all" or choose one or more of the following: "logit", "probit", "cloglog", "cauchit", "log", "loglog"
`equal.precision`	equal dispersion parameter assumption in simulation
`sd1`	the standard deviation for the treatment group. Only applicable when equal.precision = FALSE

Details

betapower function allows you to control the number of trials in the simulation, the sample sizes used, and the alternative means. You can fix the alternative and vary sample size to match a desired power; You can fix the sample size and vary the alternative to see which will match a desired power; You can vary both; Start with a small number of trials (say 100) to determine the rough range of sample sizes or alternatives; Use a larger number of trials (say 1000) to get better estimates.

Value

Return a matrix with 7 to 12 columns:

`power.of.GLM: link name`
: the power using regression method; it will return the power with every links if you use link.type = "all" statement.

`power.of.Wilcoxon.test`
: the power from Wilcoxon Rank sum test.

`sample size` sample size.

`mu1` the mean for the treatment group under the alternative.

`mu0` the mean for the control group.

`sd0` the standard deviation for the control group.

`trials` the number of trials.

Examples

```
betapower(mu0 = 0.56, sd0 = 0.255, mu1.start = .70, mu1.end = .75, mu1.by = .05,
ss.start = 30, ss.end = 50, ss.by = 20, trials = 40)
betapower(0.56, 0.255, .70, ss.start = 30, ss.end = 50, ss.by = 20,trials = 40,
link.type = c("logit","loglog","log"))
```

plot_betapower *Plots of Beta power*

Description

Generate several comparison plots of power.

Usage

```
plot_betapower(betapower.matrix,link.type,by)
```

Arguments

`betapower.matrix`
: a matrix obtained by the function betapower.(the formula was described as the output formula in the function betapower)

`link.type` the type of link used in the beta regression. You can choose one or more of the following: "logit", "probit", "cloglog", "cauchit", "log", "loglog", "all"

`by` the type of plot. see details.

Details

plot_betapower() returns different plots depends on by
by = "linktype": plot_betapower() returns graphs that plot power against mu1, where mu1 is the mean for the treatment group under the alternative. The number of plots will vary depending on the number of link types selected with the last plot showing power based on Wilcoxon Rank Sum Test. The first one or several plots show comparisons of power with different sample size, using GLM method with one or several link types. The last plot shows a comparison of the power with different sample size using Wilcoxon Rank Sum Test. Y-axis denotes power and X-axis denotes mu1, the

mean for the treatment group under the alternative.
by = "samplesize": plot_betapower() returns a number of plots equal to the number of sample sizes tested. Each plot compares power calculated with different link types and the Wilcoxon Rank Sum Test. Y-axis denotes power and X-axis denotes mu1, the mean for the treatment group under the alternative.
by = "mu1": plot_betapower() returns a number of plots equal to the number of mu1 used in the procedure. Each plot compares power calculated with different link types and the Wilcoxon Rank Sum Test. Y-axis denotes power and X-axis denotes sample size.

Examples

```
BPmat <- betapower(mu0 = 0.56, sd0 = 0.255, mu1.start = .70, mu1.end = .75, mu1.by = .05,
ss.start = 30, ss.end = 50, ss.by = 20, trials = 40, link.type = "all")
plot_betapower(BPmat,link.type = "all",by="linktype")
plot_betapower(BPmat,link.type = "all",by="samplesize")
plot_betapower(BPmat,link.type = "all",by="mu1")
BPmat2 <- betapower(mu0 = 0.56, sd0 = 0.255, mu1.start =.61, mu1.end = .76, mu1.by =.05,
ss.start = 30, ss.end = 45, ss.by = 5,trials = 200,link.type = c("logit","loglog","log"))
plot_betapower(BPmat2,link.type = c("logit","loglog","log"),by="linktype")
plot_betapower(BPmat2,link.type = c("logit","loglog","log"),by="samplesize")
plot_betapower(BPmat2,link.type = c("logit","loglog","log"),by="mu1")
```

plot_samplesize *Plots by mu1*

Description

Generate the comparison plots using GLM method and Wilcoxon Rank Sum Test with different mu1.

Usage

```
plot_samplesize(SS.matrix,link.type)
```

Arguments

SS.matrix	the matrix obtained by the function samplesize.(the formula was described as the output formula in the function samplesize)
link.type	the type of link used in the beta regression(or Wilcoxon Rank Sum Test). You can choose one or more of the following: "logit", "probit", "cloglog", "cauchit", "log", "loglog", "wilcoxon", "all"

Details

plot_samplesize() returns a series of plots equal to the number of mu1 used in the procedure. Y-axis denotes minimum sample size and X-axis denotes minimum power.

Examples

```
SSmat <- samplesize(mu0=0.56, sd0=0.255, mu1.start = 0.60, mu1.end = 0.70, mu1.by = 0.05,
power.start = 0.7, power.end = 0.9, power.by = 0.1, link.type = "all")
plot_samplesize(SSmat, "all")
SSmat2 <- samplesize(mu0=0.56, sd0=0.255, mu1.start = 0.60, mu1.end = 0.70, mu1.by = 0.05,
power.start = 0.7, power.end = 0.9, power.by = 0.1, link.type = c("logit","loglog","log"))
plot_samplesize(SSmat2,link.type = c("logit","loglog","log"))
```

`samplesize` *Find minimum sample size with Beta distribution*

Description

Find minimum sample sizes with Beta distribution and given mu0,sd0,mu1 and target powers.

Usage

```
samplesize(mu0, sd0, mu1.start, mu1.end = NULL, mu1.by = NULL,
power.start, power.end = NULL, power.by = NULL, sig.level = 0.05,
trials = 100, delta = 1, seed = 1, link.type = "logit",
equal.precision = TRUE, sd1 = NULL)
```

Arguments

`mu0`	the mean for the control group
`sd0`	the standard deviation for the control group
`mu1.start`	the starting value of mean for the treatment group under the alternative mu1
`mu1.end`	the ending value of mean for the treatment group under the alternative mu1
`mu1.by`	the step length of mean for the treatment group under the alternative mu1
`power.start`	the starting value of target power
`power.end`	the ending value of target power
`power.by`	the step length of target power
`sig.level`	significant level; default value is 0.05
`trials`	the number of trials; default value is 100
`delta`	the accuracy of the result; must be integer
`seed`	the seed used in the simulation
`link.type`	default link is "logit". Other link options include: "logit", "probit", "cloglog", "log", "loglog", "wilcoxon", or you can use "all" for all types of link
`equal.precision`	equal dispersion parameter assumption in simulation
`sd1`	the standard deviation for the treatment group. Only applicable when equal.precision = FALSE

Details

The samplesize function allows you to control the number of trials in the simulation, the target power, delta, and the alternative means. You can fix the alternative and vary power to match a desired sample size; Use default values for the number of trials for a quick view; Use a larger number of trials (say 1000) and a smaller delta (say 1) to get better estimates.

Value

Return a table including minimum sample size and power, as well as the target power and mu1:

`minimum sample size:  link type:` minimum sample size for given given mu0, sd0, mu1, target power and type of link.

`minimum power:  link type:` the minimum power greater than or equal to target power.

`target power:` the target power.

`mu1:` mean for the treatment group under the alternative.

`mu0:` the mean for the control group.

`sd0:` the standard deviation for the control group.

Examples

```
samplesize(mu0=0.56, sd0=0.255, mu1.start = 0.65, mu1.end = 0.75, mu1.by = 0.05,
power.start = 0.7, power.end = 0.9, power.by = 0.1)
samplesize(mu0=0.56, sd0=0.255, mu1.start = 0.65, mu1.end = 0.75, mu1.by = 0.05,
power.start = 0.7, power.end = 0.9, power.by = 0.1, link.type = c("logit","loglog","log"))
samplesize(mu0=0.56, sd0=0.255, mu1.start = 0.65, mu1.end = 0.75, mu1.by = 0.05,
power.start = 0.7, power.end = 0.9, power.by = 0.1, link.type = "all")
```

Index