Package 'DesignCTPB'

October 12, 2022

Type Package

Title Design Clinical Trials with Potential Biomarker Effect

Version 1.1.3

Description Applying 'CUDA' 'GPUs' via 'Numba' for optimal clinical design. It allows the user to utilize a 'reticulate' 'Python' environment and run intensive Monte Carlo simulation to get the optimal cutoff for the clinical design with potential biomarker effect, which can guide the realistic clinical trials.

License GPL (>= 2)

URL <https://github.com/ubcxzhang/DesignCTPB>, Y Lu (2020) <doi:10.1002/sim.8868>

BugReports <https://github.com/ubcxzhang/DesignCTPB/issues>

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

NeedsCompilation no

Config/reticulate list(packages = list(list(package = ``scipy'', pip = TRUE), list(package = ``numba'', pip = TRUE), list(package = ``pandas'', pip = TRUE)))

Imports reticulate, mnormt, fields, plotly, dplyr

Suggests knitr, rmarkdown

VignetteBuilder knitr

Depends R (>= 3.5.0)

SystemRequirements OpenSSL(>= 1.0.1), NVIDIA CUDA GPU with compute capability 3.0 or above and NVIDIA CUDA Toolkit 9.0 or above

Author Yitao Lu [aut, cre] (<https://orcid.org/0000-0002-0523-7416>),
Belaid Moa [aut],
Julie Zhou [aut],
Li Xing [aut] (<https://orcid.org/0000-0002-4186-7909>),
Xuekui Zhang [aut] (<https://orcid.org/0000-0003-4728-2343>)

Maintainer Yitao Lu <yitaolu@uvic.ca>

Repository CRAN

Date/Publication 2021-09-21 05:10:02 UTC

R topics documented:

---

alpha.split   *The optimal design given one set of proportion for each sub-population*

Description

First, the function fits a smooth surface given grid values of alpha(that's sig.lv for each sub-population) and the corresponding power values, and we suggest thin plate splines here. Second, we apply a L-BFGS-B optimization method to estimate the optimal power values and the corresponding alpha value on the estimated thin plate spline surface.

Usage

```
alpha.split(
  r = c(1, 0.5, 0.3),
  N1 = 20480,
  N2 = 10240,
  N3 = 2000,
  E = NULL,
  sig = NULL,
  sd_full = 1/base::sqrt(20),
  delta = NULL,
  delta_linear_bd = c(0.2, 0.8),
  seed = NULL
)
```

Arguments

r	vector for the proportion for each sub-population, r_1 is 1, r_i>r_i+1
N1	integer, which is fixed as 10240 in our package
N2	integer, which is fixed as 20480 in our package

N3	integer, the number of grid point for the sig.lv, which should be the multiples of 5, because we apply 5 stream parallel
E	integer, the total number of events for the Phase 3 clinical trail, if not specified by user, then an estimation will apply
sig	the vector of standard deviation of each sub-population
sd_full	a numeric number, which denotes the prior information of standard deviation for the harzard reduction if sig is not specified, then sd_full must has an input value to define the standard deviation of the full population
delta	vector,the point estimation of harzard reduction in prior information, if not specified we apply a linear scheme by giving bound to the linear harzard reduction
delta_linear_bd	
	vector of length 2, specifying the upper bound and lower bound for the harzard reduction; if the delta is not specified for each sub-population, then the linear scheme will apply and the input is a must.
seed	integer, seed for random number generation

Value

list of the optimal results given specific r: optimal alpha split and the corresponding optimal power value

Examples

```
## Not run:
#In the example, we apply a linear scheme for the harzard reduction
alpha.split(r=c(1,0.4,0.1), N3=2000, sd_full=1/sqrt(20),delta_linear_bd = c(0.2,0.8))

## End(Not run)
```

---

ctpbs   *The clinical trial design for strong biomarker effect*

Description

The dataset contains the 3-D rotable figure for the simulated nested two-subset design(strong biomarker effect), from which the optimal choice of each population's proportion and the optimal alpha split can be abtianed.

Format

The format is: List of 5 $ plot_power : plotly object $ plot_alpha : plotly object $ opt_r_split : numeric vector $ opt_power : numeric $ opt_alpha_split: numeric vector

plot_power the 3-d rotatable plot of optimal power versus r2 and r3.

plot_alpha the 3-d rotatable plot of optimal alpha versus r2 and r3

opt_r_split The optimal cutoffs for the 3 dimensional clinical design

opt_power The optimal power values corresponding to the optimal r split

opt_alpha_split The optimal alpha split corresponding to the optimal r split

---

ctpbw   *The clinical trial design for weak biomarker effect*

Description

The dataset contains the 3-D rotable figure for the simulated nested two-subset design(weak biomarker effect), from which the optimal choice of each population's proportion and the optimal alpha split can be abtianed.

Format

The format is: List of 5 $ plot_power :plotly object $ plot_alpha :plotly object $ opt_r_split : numeric vector $ opt_power : numeric $ opt_alpha_split: numeric vector

plot_power the 3-d rotatable plot of optimal power versus r2 and r3.

plot_alpha the 3-d rotatable plot of optimal alpha versus r2 and r3

opt_r_split The optimal cutoffs for the 3 dimensional clinical design

opt_power The optimal power values corresponding to the optimal r split

opt_alpha_split The optimal alpha split corresponding to the optimal r split

---

designCTPB   *Optimal design for 3-dimensional with visulization*

Description

This function uses GPU parallel computing to calculate the high dimensional integral and apply the smoothing method(thin plate splines) to get the optimum of power values given the prior information: the harzard reduction distribution. This function guides to choose the size of nested populations, i.e. find optimal r-values. The function visualizes and optimizes r-values, but only supports 3-dimension. The optimization of r-values in more than 3-dimension is trivial, but visualization can be too hard.

Usage

```
designCTPB(
  m = 24,
  r_set = NULL,
  n_dim = 3,
  N1 = 20480,
  N2 = 10240,
  N3 = 2000,
  E = NULL,
  SIGMA = NULL,
  sd_full = 1/base::sqrt(20),
```

```
    DELTA = NULL,
    delta_linear_bd = c(0.2, 0.8),
    seed = NULL
)
```

Arguments

m	integer, the number of grid points in each dimension for r, and we suggest m around 20 is enough for 3 dimension
r_set	the matrix of proportion for each sub-population, r_1 is 1, r_i>r_i+1
n_dim	integer, the number of dimension
N1	integer, which is fixed as 10240 in our package
N2	integer, which is fixed as 20480 in our package
N3	integer, the number of grid point for the sig.lv, which should be the multiples of 5, because we apply 5 stream parallel
E	integer, the total number of events for the Phase 3 clinical trail, if not specified by user, then an estimation will apply
SIGMA	the matrix of standard deviation of each sub-population, which should coincide with r_set or the default setting of each sub-population(i.e each entry of each row coincides to the corresponding entry in r_set)
sd_full	a numeric number, which denotes the prior information of standard deviation for the harzard reduction if sig is not specified by user, then sd_full must has an input value to define the standard deviation of the full population
DELTA	matrix, each row is an vector stands for the point estimation of harzard reduction in prior information corresponds to the r setting, if not specified we apply a linear scheme by giving bound to the linear harzard reduction
delta_linear_bd	
	vector of length 2, specifying the upper bound and lower bound for the harzard reduction; if user don't specify the delta for each sub-population, then the linear scheme will apply and the input is a must.
seed	integer, seed for random number generation

Details

the standard deviation of each population can be specified by giving SIGMA as input, and specify the harzard reduction rate DELTA for each population. Just enter values to SIGMA and DELTA, but note that the entered matrix should coincides with the matrix of r-split setting.

Value

list of 5 parts: plot_power: 3-d plot of the optimal power values versus r2 and r3; plot_alpha: 3-d plot of the optimal alpha-split values versus r2 and r3; opt_r_split: the optimal choice of proportion for each sub-population; opt_power: the optimal power values with the optimal r choice; opt_alpha_split: the optimal alpha split with the optimal r choice

See Also

Grid setting of proportions for each sub-population proportion() and alpha.split()

Examples

```
## Not run:
# the default setting of our paper's strong biomarker effect
res <- designCTPB()
res$plot_power # to see 3-d plot for the optimal power versus r2 and r3
res$plot_alpha # to see 3-d plot for the optimal alpha versus r2 and r3
res$opt_r_split #  to see the optimal cutoff of the sub-population,
#and here suggesting not cutoff at the 2-nd sub-population
res$opt_power
res$opt_alpha_split

## End(Not run)
```

---

phat   *Point estimator for the power value*

Description

This function is to estimate the power values given fixed proportion r for each sub-population, which we utilize Monte Carlo method and GPU accelerator to estimate the power value. The user can specify the standard deviation and harzard reduction for each sub-population as the prior information of harzard reduction distribution, when not specified, we apply a default setting of linear harzard reduction scheme and the sd for each sub-population is inversely proportional to sqrt(r_i)

Usage

```
phat(
  r,
  N1,
  N2,
  N3,
  E = NULL,
  sig = NULL,
  sd_full,
  delta = NULL,
  delta_linear_bd,
  seed = NULL
)
```

Arguments

r	vector for the proportion for each sub-population, r_1is 1, r_i>r_i+1
N1	integer, which is fixed as 10240 in our package

N2	integer, which is fixed as 20480 in our package
N3	integer, the number of grid point for the sig.lv, which should be the multiples of 5, because we apply 5 stream parallel
E	integer, the total number of events for the Phase 3 clinical trail, if not specified, then an estimation will be applied
sig	the vector of standard deviation of each sub-population
sd_full	a numeric number, which denotes the prior information of standard deviation for the harzard reduction. If sig is not specified, then sd_full must has an input value to define the standard deviation of the full population
delta	vector, the point estimation of harzard reduction in prior information, if not specified we apply a linear scheme by giving bound to the linear harzard reduction
delta_linear_bd	
	vector of length 2, specifying the upper bound and lower bound for the harzard reduction; if user don't specify the delta for each sub-population, then the linear scheme will apply and the input is a must.
seed	integer, seed for random number generation

Details

We interface python by reticulate package to utilize numba(cuda version) module to accelerate calculation.

Value

list of 2 parts of the sampling points given specific r; alpha is the matrix as each row is the given sig.lv for each population; power is the corresponding power values given each row of the alpha

---

proportion	*Grid setting of proportions for each sub-population*

Description

This function is to decide the r setting given specific density in each dimension

Usage

```
proportion(m, n_dim)
```

Arguments

m	integer, the number of grid points in each dimension, and we suggest m around 20 for 3 dimension
n_dim	integer for the dimension, which is equal to the number of sub-population plus 1

Value

matrix of setting the proportion of the population by given specific dimension and density in each dimension

Index