

Package ‘debest’

May 8, 2026

Type Package

Title Duration Estimation for Biomarker Enrichment Studies and Trials

Version 0.1.0

Depends R (>= 3.5.0)

Maintainer Hong Zhang <hzhang@wpi.edu>

Description A general framework using mixture Weibull distributions to accurately predict biomarker-guided trial duration accounting for heterogeneous population. Extensive simulations are performed to evaluate the impact of heterogeneous population and the dynamics of biomarker characteristics and disease on the study duration. Several influential parameters including median survival time, enrollment rate, biomarker prevalence and effect size are identified. Efficiency gains of biomarker-guided trials can be quantitatively compared to the traditional all-comers design. For reference, see Zhang et al. (2024) <[doi:10.48550/arXiv.2401.00540](https://doi.org/10.48550/arXiv.2401.00540)>.

License GPL-2

Imports stats, survival, flexsurv

Encoding UTF-8

LazyData true

RoxygenNote 7.2.0

NeedsCompilation no

Author Hong Zhang [aut, cre] (ORCID: <<https://orcid.org/0000-0002-8869-8671>>)

Repository CRAN

Date/Publication 2024-01-10 14:23:06 UTC

Contents

calcDuration	2
dat_udca	3
getBeta	4
getWeibull	5
Index	6

calcDuration *Calculate the study duration based on Weibull distributions.*

Description

Calculate the study duration based on Weibull distributions.

Usage

```
calcDuration(  
  d,  
  n,  
  proportion,  
  SHAPEV,  
  SCALEV,  
  LAMW,  
  A,  
  BETA,  
  conf.level = 0.9,  
  nsim = 10000  
)
```

Arguments

d	- scalar, target number of events.
n	- scalar, sample size.
proportion	- vector of percentages of the subgroup.
SHAPEV	- vector of shape parameters of Weibull survival distributions of the subgroups.
SCALEV	- vector of scale parameters of Weibull survival distributions of the subgroups.
LAMW	- vector of exponential drop-out distribution parameters of the subgroups.
A	- vector of enrollment durations of the subgroups.
BETA	- vector of beta distribution parameters of the subgroups.
conf.level	- scalar, confidence level, default 0.9.
nsim	- scalar, number of repetitions, default 1e4.

Value

study duration estimate, d_med, and the confidence interval (d_lower, d_upper), as well as all the realizations, Z_d, of the study duration from the simulation.

References

Hong Zhang, Jie Pu, Shibing Deng, Satrajit Roychoudhury, Haitao Chu and Douglas Robinson. "Study Duration Prediction for Clinical Trials with Time-to-Event Endpoints Using Mixture Distributions Accounting for Heterogeneous Population", arXiv:2401.00540.

Examples

```

res_weibull = getWeibull(dat_udca)
res_beta = getBeta(as.numeric(dat_udca$entry.dt))
prop = c(table(dat_udca$group)/length(dat_udca$group))
SHAPEV = res_weibull$shape
SCALEV = res_weibull$scale
LAMW = rep(-log(1 - 0.1)/6, 4)
A = rep(res_beta$a/30.416, 4) # convert days to months
BETA = rep(res_beta$b_Mean, 4)
myres1 = calcDuration(d=50, n=169, proportion=prop, SHAPEV, SCALEV, LAMW=LAMW, A, BETA)
c(myres1$d_lower, myres1$d_med, myres1$d_upper)
# drop-out will make the target number of events not achievable
myres2 = calcDuration(d=80, n=169, proportion=prop, SHAPEV, SCALEV, LAMW=LAMW, A, BETA)
c(myres2$d_lower, myres2$d_med, myres2$d_upper)
# If there is no drop-out
myres3 = calcDuration(d=80, n=169, proportion=prop, SHAPEV, SCALEV, LAMW=rep(0, 4), A, BETA)
c(myres3$d_lower, myres3$d_med, myres3$d_upper)

```

dat_udca

A derived biomarker dataset from a trial of usrodeoxycholic acid

Description

dat_udca is derived from the udca data from the R survival package with biomarker variables defined as the baseline bilirubin level > 1.

Usage

```
dat_udca
```

Format

A data frame with 169 rows and 12 variables:

id subject id

trt treatment assignment

stage disease stage

bili bilirubin level

riskscore disease risk score

futime follow up time in days

status event indicator

time follow up time in months

entry.dt enrollment date

last.dt last visit date

BMpos biomarker positive indicator (bili>1)

group subgroup indicator. 1 : trt=0&BMpos=FALSE; 2 : trt=1&BMpos=FALSE; 3 : trt=0&BMpos=TRUE;
4 : trt=1&BMpos=TRUE.

Source

udca data from the R survival package. <https://CRAN.R-project.org/package=survival>

Examples

```
data("dat_udca")
```

getBeta

Estimate parameters of the distribution of $a \cdot \text{Beta}(1, b)$.

Description

Estimate parameters of the distribution of $a \cdot \text{Beta}(1, b)$.

Usage

```
getBeta(x)
```

Arguments

x - vector of real positive numbers.

Value

estimates of a and b. For a, it is the range of x. For b, two estimates are provided: the MLE, and the method of moment (matching mean).

References

Hong Zhang, Jie Pu, Shibing Deng, Satrajit Roychoudhury, Haitao Chu and Douglas Robinson. "Study Duration Prediction for Clinical Trials with Time-to-Event Endpoints Using Mixture Distributions Accounting for Heterogeneous Population", arXiv:2401.00540.

Examples

```
getBeta(as.numeric(dat_udca$entry.dt))
```

getWeilbull	<i>Estimate parameters of the Weibull survival distributions for each subgroup.</i>
-------------	---

Description

Estimate parameters of the Weibull survival distributions for each subgroup.

Usage

```
getWeilbull(dat)
```

Arguments

dat - data frame of three columns: 1) time: follow up time; 2) status: indicator, 0=no event, 1=event; 3) group, integer 1,...,n, for each subgroup of patients.

Value

shape and scale parameters of the Weibull distribution for each subgroup.

References

Hong Zhang, Jie Pu, Shibing Deng, Satrajit Roychoudhury, Haitao Chu and Douglas Robinson. "Study Duration Prediction for Clinical Trials with Time-to-Event Endpoints Using Mixture Distributions Accounting for Heterogeneous Population", arXiv:2401.00540.

Examples

```
# dat_udca already has time, status and group columns defined,  
getWeilbull(dat_udca)
```

Index

* **datasets**

dat_udca, 3

calcDuration, 2

dat_udca, 3

getBeta, 4

getWeilbull, 5