Package ‘survSNP’

February 20, 2015

Type Package
Title Power Calculations for SNP Studies with Censored Outcomes
Version 0.23.2
Date 2015-01-17
Author Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi
Maintainer Kouros Owzar <owzar@duke.edu>
Description This package can be used to conduct asymptotic and empirical power and sample size calculations for SNP association studies with right censored time to event outcomes
Depends R (>= 3.0.0), survival (>= 2.36-9), Rcpp (>= 0.9.10), lattice (>= 0.20-0), foreach (>= 1.3.2), xtable (>= 1.7-0)
SystemRequirements GNU GSL (>= 1.14)
Suggests RColorBrewer, latticeExtra, knitr
LinkingTo Rcpp
URL https://bitbucket.org/kowzar/survsnp
License GPL-3
LazyLoad yes
VignetteBuilder knitr
NeedsCompilation yes
Repository CRAN
Date/Publication 2015-01-18 06:50:29

R topics documented:
survSNP-package ................................................. 2
asypow .......................................................... 3
censbnd .......................................................... 4
hwe .............................................................. 5
sim.snp.expsurv.power ........................................ 6
sim.snp.expsurv.sctest ........................................ 8
surv.exp.gt.model ............................................. 9
survSNP.power.table ......................................... 10
Description

This package facilitates asymptotic and empirical power and sample size calculations for SNP association studies with censored time-to-event outcomes.

Details

Package: survSNP
Type: Package
Version: 0.23.2
Date: 2015-01-17
License: GPL-3
LazyLoad: yes

The functions `sim.snp.expsurv.power` and `survSNP.power.table` can be used to conduct power and sample size calculations. The package vignette serves as a tutorial for using this package. The technical details are provided in the reference cited below. It is highly recommended that the vignette along with this are reviewed before using this package. Currently, this package only supports additive risk models.

Author(s)

Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi

References


Examples

# See vignette for examples including details on the example
# considered below

```r
results<-sim.snp.expsurv.power(GHR=1.25, B=0, n=500, raf=0.1, erate=0.75, pilm=0.5,
lm=1, model="additive", test="additive", alpha=0.05)
results[,c("n","erate","alpha","pow0")]
```

```r
GHRs<-seq(1.05,1.5,by=0.05)
ns<-c(100,500,700)
```
asypow <- c(0.1, 0.3, 0.5)
erates <- c(0.5, 0.7, 0.9)
res <- survSNP.power.table(GHRS, ns, raf, erate, pilm=0.5, lm=1, model="additive", test="additive", alpha=0.05)

# Create key for illustration
KEY <- paste("q=", levels(factor(res$raf)), sep="")
KEY <- list(lines=list(col=1:length(KEY),lty=1:length(KEY)),
text=list(labels=paste("q=", levels(factor(res$raf)), sep="")),
column=3)

# Illustrate Power
print(xyplot(pow~ghr|factor(erate)*factor(n),group=factor(raf),
data=res, type="l", lty=KEY$lines$lty, col=KEY$lines$col,
key=KEY, xlab="Genotype hazard ratio", ylab="Power"))

# Illustrate Power (restricted to n=100)
print(xyplot(pow~ghr|factor(erate), group=factor(raf),
data=subset(res, n==ns[1]),
type="l", lty=KEY$lines$lty, col=KEY$lines$col,
key=KEY, xlab="Genotype hazard ratio", ylab="Power",
sub=paste("n=" , ns[1], ", alpha=" , round(unique(res$alpha), 2))))

asypow  Calculating the asymptotic power and variance

Description
This function calculates the asymptotic power and variance assuming that the survival distribution is a mixture of exponentials with rates and the censoring distribution is uniform on the interval (a,b).

Usage
asypow(n, theta, a, b, lambda0, q, p, alpha, z, exactvar)

Arguments
n Sample size
theta Effect size (log genotype hazard ratio (GHR))
a Censoring distribution parameter (assumed to be uniform on [a,b])
b Censoring distribution parameter (assumed to be uniform on [a,b])
lambda0 Baseline exponential hazard rate
q Relative risk allele frequency
Relative genotype frequency
Nominal two-sided type I error rate
Genotype scores (right now only additive scores AA=0, AB=1, BB=2 generate correct power)
Indicator for using the exact variance formula

Details
This function is called by `sim.snp.expsurv.power` to calculate the asymptotic variance (exact and approximate) formulas. It is not intended to be called directly by the user. To conduct power calculations, use `sim.snp.expsurv.power` or the convenience wrapper function `survSNP.power.table`.

Value
- `power`: Asymptotic power based on exact variance formula
- `power0`: Asymptotic power based on approximate variance formula
- `v1`: First term of asymptotic variance
- `v2`: Second term of asymptotic variance
- `v12`: Third term of the asymptotic variance (covariance)
- `vapprox`: Approximate asymptotic variance formula (=v1)
- `exact`: Exact asymptotic variance formula (=v1+v2+v12)
- `diff`: Difference between variances (=v2+v12)
- `ratio`: Ratio of variances (=v1/(v1+v2+v12))

Author(s)
Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi

References

censbnd

Calculate bound for censoring distribution.

Description
This function computes the bound

Usage
```
censbnd(lambda, p, crate, rootint = c(0.1, 1000))
```
Arguments

lambda Baseline exponential hazard rate
p Relative genotype frequency
crate Desired censoring rate
rootint Interval to be searched for the root

Details

The time to event distribution $T$ is assumed to be a mixture of exponentials with parameter $\lambda_0$, $\lambda_1$ and $\lambda_2$ with mixing proportion $p_0, p_1$ and $p_2$. Suppose that the censoring distribution is uniform on the interval $[0, b]$. This function calculates $b$ for a desired censoring rate. It is not intended to be called directly by the user. To conduct power calculations, use sim.snp.expsurv.power or the convenience wrapper function survSNP.power.table.

Value

This function returns a list from the uniroot function. The root component of this list in the

Author(s)

Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi

References


Examples

censbnd(0.1, hwe(0.1), 0.9)$root

hwe Relative genotypic frequencies under HWE

Description

Compute relative genotypic frequencies for a given relative allelic frequency

Usage

hwe(raf)

Arguments

raf Relative minor allele frequency for the B allele.
Details
For a bi-allelic SNP with genotypes AA, AB and BB, with a relative allele frequency \( q \) for the B allele, this function returns the corresponding relative genotypic frequencies. It is not intended to be called directly by the user. To conduct power calculations, use \( \text{sim.snp.expsurv.power} \) or the convenience wrapper function \( \text{survSNP.power.table} \).

Value
A vector of length three relative genotypic frequencies.

Author(s)
Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi

References

Examples
\begin{verbatim}
censbnd(0.1,hwe(0.1),0.9)
\end{verbatim}

\begin{verbatim}
sim.snp.expsurv.power  Asymptotic and Empirical Power
\end{verbatim}

Description
This function calculates asymptotic and empirical power for SNP association studies

Usage
\begin{verbatim}
sim.snp.expsurv.power(GHR, B, n, raf, erate, pilm, lm, model, test, alpha, exactvar = FALSE, interval = c(0, 10), rootint = c(0.1, 200))
\end{verbatim}

Arguments
\begin{itemize}
\item \text{GHR} \hspace{1cm} \text{Genotype Hazard Ratio}
\item \text{B} \hspace{1cm} \text{Number of simulation replicates (set to 0 if no empirical calculations are desired)}
\item \text{n} \hspace{1cm} \text{Sample size}
\item \text{raf} \hspace{1cm} \text{Relative risk allele frequency}
\item \text{erate} \hspace{1cm} \text{Event Rate}
\item \text{pilm} \hspace{1cm} \text{Probability that the time-to-event is greater than lm}
\item \text{lm} \hspace{1cm} \text{Landmark time used for powering the study}
\end{itemize}
model          True genetic risk model (choices are "additive", "recessive" or "dominant". For the asymptotic calculations only the "additive" should be used until further notice)

test           Hypothesized genetic risk model (choices are "additive", "recessive" or "dominant". For the asymptotic calculations only the "additive" should be used until further notice)

alpha          Nominal two-sided type I error rate

exactvar       Indicator for using the exact variance formula

interval       Interval to search for baseline hazard rate

rootint        Interval to search for censoring bound

Details

This function calculates asymptotic and empirical power for SNP association studies

Value

This function returns a data.frame with the following columns

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Number of simulation replicates</td>
</tr>
<tr>
<td>raf</td>
<td>Relative allelic frequency of the B allele</td>
</tr>
<tr>
<td>q0</td>
<td>Relative genotypic frequency for AA</td>
</tr>
<tr>
<td>q1</td>
<td>Relative genotypic frequency for AB</td>
</tr>
<tr>
<td>q2</td>
<td>Relative genotypic frequency for BB</td>
</tr>
<tr>
<td>lam0</td>
<td>Exponential hazard rate for P(T&gt;t</td>
</tr>
<tr>
<td>lam1</td>
<td>Exponential hazard rate for P(T&gt;t</td>
</tr>
<tr>
<td>lam2</td>
<td>Exponential hazard rate for P(T&gt;t</td>
</tr>
<tr>
<td>GHR</td>
<td>Genotype Hazard Ratio</td>
</tr>
<tr>
<td>pilm</td>
<td>Probability that the time to event in the population exceeds the landmark lm: P(T&gt;lm)</td>
</tr>
<tr>
<td>lm</td>
<td>Landmark (see pilm above)</td>
</tr>
<tr>
<td>model</td>
<td>The true genetic risk model</td>
</tr>
<tr>
<td>test</td>
<td>The hypothesized genetic risk model</td>
</tr>
<tr>
<td>a</td>
<td>Parameter for the censoring distribution (uniform on [a,b])</td>
</tr>
<tr>
<td>b</td>
<td>Parameter for the censoring distribution (uniform on [a,b])</td>
</tr>
<tr>
<td>erate</td>
<td>Event rate</td>
</tr>
<tr>
<td>n</td>
<td>Sample size</td>
</tr>
<tr>
<td>powB</td>
<td>Empirical Power</td>
</tr>
<tr>
<td>pow</td>
<td>Asymptotic Power based on the exact variance formula</td>
</tr>
<tr>
<td>pow0</td>
<td>Asymptotic Power based on the approximate variance formula</td>
</tr>
<tr>
<td>v1</td>
<td>First term of the variance (v1+v2+v12)</td>
</tr>
<tr>
<td>v2</td>
<td>Second term of the variance (v1+v2+v12)</td>
</tr>
<tr>
<td>v12</td>
<td>Third term (covariance) of the variance (v1+v2+v12)</td>
</tr>
</tbody>
</table>
**Author(s)**
Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi

**References**

**See Also**
survSNP.power.table

**Examples**
```r
# See vignette for examples including details on the example
# considered below

results<-sim.snp.expsurv.power(GHR=1.25, B=0, n=500, raf=0.1, erate=0.75, pilm=0.5, lm=1, model="additive", test="additive", alpha=0.05)
results[,c("n","erate","alpha","pow0")]
```

---

**sim.snp.expsurv.sctest**

*Simulation of Cox Score Statistic*

**Description**
This function simulates the asymptotic P-value for the Cox score statistic under a specified genetic risk model.

**Usage**
sim.snp.expsurv.sctest(n, gtprev, lam, a, b, ztest, diag = FALSE)

**Arguments**
- **n** Sample size
- **gtprev** Relative genotypic frequency
- **lam** Exponential hazard rates for conditional time to event survival functions
- **a** Lower bound for the uniform censoring bound
- **b** Upper bound for the uniform censoring bound
- **ztest** Assumed genetic model
- **diag** Set to TRUE if print out of diagnostics is desired
Details

This function simulates the genotypes 0, 1 or 2 from relative genotypic frequencies. Then conditional on genotype, it simulates the time to event from an exponential distribution with parameter $\lambda_{t|d}$. The censoring time is drawn from a uniform law on the interval $[a, b]$. It is not intended to be called directly by the user. To conduct power calculations, use `sim.snp.expsurv.power` or the convenience wrapper function `survSNP.power.table`.

Value

A vector of length 2 containing the observed event rate and the asymptotic $P$-value for the Cox score test

Author(s)

Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi

References


Description

This function calculates the exponential hazard rates for the conditional survival functions

Usage

```r
surv.exp.gt.model(pilm, lm, gtprev, GRR, zmodel, interval)
```

Arguments

- `pilm`: Probability that the time-to-event is greater than lm
- `lm`: Landmark time used for powering the study
- `gtprev`: Relative genotypic frequency
- `GRR`: Genotype Hazard Ratio
- `zmodel`: Genetic Risk Model (choices are "additive", "recessive" or "dominant")
- `interval`: Interval to search for baseline hazard rate

Details

Determines the time to event distributions in the three component mixture model discussed in the reference below. It is not intended to be called directly by the user. To conduct power calculations, use `sim.snp.expsurv.power` or the convenience wrapper function `survSNP.power.table`.
Value

A vector of length three containing the exponential hazard rates

Author(s)

Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi

References


survSNP.power.table  Table for Asymptotic and Empirical Power

Description

This function produces a table with the resulting empirical and asymptotic power over a given range of Genotype Hazard Ratios, sample sizes, etc.

Usage

survSNP.power.table(GHRs, ns, rafs, erates, pilm, lm, model, test, alpha, exactvar = FALSE, B = 0, para = FALSE)

Arguments

GHRs  A vector of Genotype Hazard Ratios
ns  A vector of sample sizes
rafs  A vector of relative allelic frequencies (for the risk allele)
erates  A vector of event rates
pilm  Probability that the time-to-event is greater than lm
lm  Landmark time used for powering the study
model  True genetic risk model (choices are "additive", "recessive" or "dominant". For the asymptotic calculations only the "additive" should be used until further notice)
test  Hypothesized genetic risk model (choices are "additive", "recessive" or "dominant". For the asymptotic calculations only the "additive" should be used until further notice)
alpha  Nominal two-sided type I error rate
exactvar  Indicator for using the exact variance formula
B  Number of simulation replicates (set to 0 if no empirical calculations are desired)
para  Indicator (TRUE/FALSE) to use a parallel backend for foreach.
**Details**

This version only supports additive models.

**Value**

See output of sim.snp.expsurv.power.

**Author(s)**

Kouros Owzar, Zhiguo Li, Nancy Cox, Sin-Ho Jung and Chanhee Yi

**References**


**Examples**

```r
grrs <- seq(1.5, 2, by=0.25)
ns <- c(100, 500, 1000)
rafs <- c(0.3, 0.5, 0.7)
erates <- c(0.5, 0.7, 0.9)
res <- survSNP.power.table(grrs, ns, rafs, erates, pilm=0.5, lm=1, model="additive",
test="additive", alpha=0.05)

# Create key for illustration
KEY <- paste("q=" , levels(factor(res$raf)), sep="")
KEY <- list(lines=list(col=1:length(KEY), lty=1:length(KEY)),
text=list(labels=paste("q=" , levels(factor(res$raf)), sep="")),
column=3)

# Illustrate Power
print(xyplot(pow~grr|factor(erate)*factor(n), group=factor(raf),
data=res, type="l", lty=KEY$lines$lty, col=KEY$lines$col,
key=KEY, xlab="Genotype Hazard Ratio", ylab="Power"))

# Illustrate Power (restricted to n=100)
print(xyplot(pow~grr|factor(erate), group=factor(raf),
data=subset(res, n==ns[1]),
type="l", lty=KEY$lines$lty, col=KEY$lines$col,
key=KEY, xlab="Genotype Hazard Ratio", ylab="Power",
sub=paste("n=" , ns[1], ", alpha=" , round(unique(res$alpha), 2))))
```
Index

asypow, 3

censbnd, 4

hwe, 5

sim.snp.expsurv.power, 2, 4–6, 9, 11
sim.snp.expsurv.sctest, 8
surv.exp.gt.model, 9
survSNP (survSNP-package), 2
survSNP-package, 2
survSNP.power.table, 2, 4–6, 8, 9, 10