

Package 'ldDesign'

April 17, 2009

Version 1.1-0

Title Design of experiments for detection of linkage disequilibrium

Author Rod Ball <rod.ball@ensisjv.com>

Maintainer Rod Ball <rod.ball@ensisjv.com>

Description R package for design of experiments for association studies for detection of linkage disequilibrium. Uses an existing deterministic power calculation for detection of linkage disequilibrium between a bi-allelic QTL and a bi-allelic marker, together with the Spiegelhalter and Smith Bayes factor to generate designs with power to detect effects with a given Bayes factor.

Depends

License GPL (>= 2)

URL <mailto:rod.ball@ensisjv.com> www.scionresearch.com/

Repository CRAN

Date/Publication 2007-04-11 08:22:01

R topics documented:

ld.design	2
ld.sim	3
luo.ld.populations	6
luo.ld.power	6
oneway.bf.alpha	8
SS.oneway.bf	10

Index	12
--------------	-----------

ld.design

*Functions for design of experiments to detect linkage disequilibrium***Description**

Find the sample size required to detect linkage disequilibrium with a given Bayes factor, with a given power, or find the power of experimental designs to detect linkage equilibrium with a given Bayes factor.

Usage

```
ld.design(p, q, D, h2, phi, Bf, power, nmin = 50, nmax = 1e+05, ninterp = 50,
          missclass.rate = 0, print.it = FALSE)
ld.power(n, p, q, D, h2, phi, Bf, missclass.rate = 0)
```

Arguments

n	ld.power: vector of sample sizes
p	Bi-allelic marker allele frequency
q	Bi-allelic QTL allele frequency
D	Linkage disequilibrium coefficient
h2	QTL ‘heritability’, i.e. proportion of total or phenotypic variance explained by the QTL
phi	Dominance ratio: phi = 0 denotes purely additive, phi = 1 denotes purely dominant allele effects
Bf	Bayes factor
power	ld.design: Power, or probability of detecting an effect with Bayes factor greater than Bf
nmin	ld.design: Lower bound for sample size
nmax	ld.design: Upper bound for sample size
ninterp	ld.design: Number of sample sizes to try
missclass.rate	Proportion of marker values which are missclassified, i.e. incorrect (to allow for genotyping errors)
print.it	If TRUE print results for sample sizes tried

Details

These functions implement the method described in Ball (2003) for obtaining the power of designs for detecting linkage disequilibrium with a given Bayes factor. The F values, (and hence significance levels) corresponding to the given Bayes factors, sample sizes, and marker genotype frequencies, are calculated using the method of Spiegelhalter and Smith (1982) (R functions [oneway.bf.alpha.required](#), [SS.oneway.bf](#)). The power is obtained using a corrected version of the classical deterministic power calculation from Luo (1988) (R function [luo.ld.power](#)).

Value

For `ld.power`, a matrix with columns:

<code>n</code>	Sample sizes
<code>power</code>	Power of the design with the given sample sizes

Additionally the return value has attributes indicating the linkage disequilibrium parameters used. For `ld.design` the sample size is returned.

Author(s)

Rod Ball (rod.ball@forestresearch.co.nz) www.forestresearch.co.nz

References

- Ball, R.D. 2003 Experimental designs for reliable detection of linkage disequilibrium in unstructured random population association studies.
- Luo, Z.W. 1988 Detecting linkage disequilibrium between a polymorphic marker locus and a trait locus in natural populations. *Heredity* 80, 198–208
- Spiegelhalter, D. and A.F.M. Smith 1982 Bayes factors for linear and log-linear models with vague prior information *J. Royal Statist Soc. B* 44: 377–387.

See Also

[luo.ld.power](#), [ld.sim](#), [oneway.bf.alpha](#), [oneway.bf.alpha.required](#), [SS.oneway.bf](#)

Examples

```
ld.power(n=seq(100,1000,by=100),p=0.5,q=0.5,D=0.1,h2=0.1,phi=0,Bf=20)
ld.design(p=0.5,q=0.5,D=0.1,h2=0.1,phi=0,Bf=20,power=0.9,print.it=TRUE,nmin=600,nmax=4000)
ld.design(p=0.5,q=0.5,D=0.1,h2=0.1,phi=0,Bf=20,power=0.9,print.it=FALSE,nmin=1700,nmax=1900)
```

ld.sim

Functions to simulate populations with a bi-allelic marker and QTL in linkage disequilibrium and test for association.

Description

For a bi-allelic marker and QTL, with given allele frequencies, linkage disequilibrium, and QTL heritability, multiple replicate populations with marker, QTL, and trait values are simulated and tested for a marker-trait association. Results can be used to estimate the power of an experimental design for detecting linkage disequilibrium.

Usage

```
ld.sim(nsim, n, p, q, D, h2, Vp, phi, missclass.rate = 0, sim.chunk = 100,
       method = 1, print.it = TRUE, data.only=FALSE)
ld.sim1(n, p, q, D, d, h, sig2.error, missclass.rate = 0, nreps = 1,
       method = 2, print.it = TRUE, data.only=FALSE)
```

Arguments

<code>nsim</code>	Number of replicate simulations to do
<code>n</code>	The sample size, i.e. number of individuals genotyped and tested for the trait of interest
<code>p</code>	Bi-allelic marker allele frequency
<code>q</code>	Bi-allelic QTL allele frequency
<code>D</code>	Linkage disequilibrium coefficient
<code>h2</code>	QTL ‘heritability’, i.e. proportion of total or phenotypic variance explained by the QTL
<code>Vp</code>	<code>ld.sim</code> : Total or phenotypic variance: an arbitrary value may be used
<code>phi</code>	<code>ld.sim</code> : Dominance ratio: $\phi = 0$ denotes purely additive gene action, $\phi = 1$ denotes completely dominant gene action
<code>d</code>	<code>ld.sim1</code> : Expected value for trait when QTL genotype is QQ,qq respectively is d,-d
<code>h</code>	<code>ld.sim1</code> : Expected value for trait when QTL genotype is Qq is h
<code>sig2.error</code>	<code>ld.sim1</code> : Error variance when QTL genotype known and modelled
<code>missclass.rate</code>	Proportion of marker values which are missclassified, i.e. incorrect
<code>sim.chunk</code>	<code>ld.sim</code> : Number of replicates to do in a ‘chunk’ in each call to <code>ld.sim1</code>
<code>nreps</code>	<code>ld.sim1</code> : Number of replicate simulations to do for the given set of marker genotypes
<code>method</code>	If <code>method=1</code> simulate random QTL genotypes conditional on marker values in <code>ld.sim1</code> ; if <code>method=2</code> simulate markers and QTL directly from table of joint probabilities. With <code>method=1</code> , a common set of marker values are used for each of the <code>nreps</code> replicates per call to <code>ld.sim1</code> , enabling MANOVA to be used.
<code>print.it</code>	if TRUE, print results
<code>data.only</code>	if TRUE, just return the simulated trait and marker genotype data

Details

Marker, QTL, and trait values are simulated according to the genetic model with normal errors. In `ld.sim`, QTL parameters `d`, `h` are determined from the parameters `h2`, `q`, `phi`, and `Vq`, and the main simulation done for each chunk of replicates by a call to `ld.sim1`. Marker-trait association is tested by a one-way analysis of variance of trait values in terms of marker classes. The proportion of results with P-value over a given threshold gives a stochastic estimate of the power calculated by [luo.ld.power](#).

Value

If `data.only=FALSE`, an array with 1 row per simulation run, and 4 columns with values for each run:

<code>MS.beta</code>	Between marker classes mean square
----------------------	------------------------------------

MS.within	Within marker classes mean square
F.value	F value
P.value	P value
marker	Marker genotype indicator with values {1,2,3}, corresponding to genotypes {MM,Mm,mm}
y	trait values
replicate	Replicate population indicator. Each blocks of rows with a given replicate number is a simulated population with the given parameters

Author(s)

Rod Ball (rod.ball@forestresearch.co.nz) www.forestresearch.co.nz

References

Luo, Z.W. 1988 Detecting linkage disequilibrium between a polymorphic marker locus and a trait locus in natural populations. *Heredity* 80, 198–208.

See Also

[luo.ld.power](#)

Examples

```
# Power from stochastic simulation for Luo's population 12
data(luo.ld.populations)
luo.pop12.sim <- ld.sim(nsim=3000,
                      n=luo.ld.populations[12,"n"],
                      p=luo.ld.populations[12,"p"],
                      q=luo.ld.populations[12,"q"],
                      D=luo.ld.populations[12,"D"],
                      h2=luo.ld.populations[12,"h2"],
                      phi=luo.ld.populations[12,"phi"],
                      Vp=100)

# power
table(luo.pop12.sim[,4] < 0.05)[2]/sum(table(luo.pop12.sim[,4] < 0.05))
# Cf power from deterministic calculation
luo.ld.power(n=luo.ld.populations[12,"n"],
            p=luo.ld.populations[12,"p"],
            q=luo.ld.populations[12,"q"],
            D=luo.ld.populations[12,"D"],
            h2=luo.ld.populations[12,"h2"],
            phi=luo.ld.populations[12,"phi"],
            Vp=100,
            alpha=0.05)
```

luo.ld.populations *Luo's Linkage disequilibrium example populations*

Description

Matrix with rows containing parameters (population size, allele frequencies, disequilibrium, dominance ratio) for example populations with a bi-allelic marker and QTL in linkage disequilibrium, from Luo (1998).

Usage

```
data(luo.ld.populations)
```

Format

The format is: num [1:12, 1:7] 1 2 3 4 5 6 7 8 9 10 ... - attr(*, "dimnames")=List of 2 ...: *NULL*... : chr [1:7] "pop." "n" "p" "q" ...

Source

Luo, Z.W. 1988 Detecting linkage disequilibrium between a polymorphic marker locus and a trait locus in natural populations. *Heredity* 80, 198–208.

Examples

```
data(luo.ld.populations)
luo.ld.populations
```

luo.ld.power *Classical deterministic power calculation for association studies to detect linkage disequilibrium*

Description

Classical deterministic power calculation for power to detect linkage disequilibrium between a bi-allelic QTL and a bi-allelic marker, at a given significance level in a population level association study.

Usage

```
luo.ld.power(n, p, q, D, h2, phi, Vp = 100, alpha, print.it = TRUE,
             missclass.rate = 0)
```

Arguments

<code>n</code>	The sample size, i.e. number of individuals genotyped and tested for the trait of interest
<code>p</code>	Bi-allelic marker allele frequency
<code>q</code>	Bi-allelic QTL allele frequency
<code>D</code>	Linkage disequilibrium coefficient
<code>h2</code>	QTL 'heritability', i.e. proportion of total or phenotypic variance explained by the QTL
<code>phi</code>	Dominance ratio: <code>phi = 0</code> denotes purely additive, <code>phi = 1</code> denotes purely dominant allele effects
<code>Vp</code>	Total or phenotypic variance: and arbitrary value may be used
<code>alpha</code>	Significance level for hypothesis tests
<code>print.it</code>	If TRUE print summary of results
<code>missclass.rate</code>	Proportion of marker values which are missclassified, i.e. incorrect

Details

This is based on the 'fixed model' power calculation from Luo (1998, *Heredity* 80, 198–208), with corrections described in Ball (2003). This function, combined with [oneway.bf.alpha](#), [oneway.bf.alpha.required](#), is used in Ball (2003) to design experiments for detecting linkage disequilibrium with a given power to achieve a given Bayes factor.

Value

Returns the power, or probability of detecting an effect, with the given parameters, at the given significance level.

Author(s)

Rod Ball (rod.ball@forestresearch.co.nz) www.forestresearch.co.nz

References

Ball, R.D. 2003 Experimental designs for reliable detection of linkage disequilibrium in unstructured random population association studies.

Luo, Z.W. 1988 Detecting linkage disequilibrium between a polymorphic marker locus and a trait locus in natural populations. *Heredity* 80, 198–208

See Also

[ld.sim](#), [oneway.bf.alpha](#), [oneway.bf.alpha.required](#), [SS.oneway.bf](#)

Examples

```

data(luo.ld.populations)
options(digits=3)
powers <- numeric(nrow(luo.ld.populations))
for(ii in 1:nrow(luo.ld.populations)){
  cat("ii=", ii, "\n")
  powers[ii] <- lu0.ld.power(n=luo.ld.populations[ii, "n"],
                           p=luo.ld.populations[ii, "p"],
                           q=luo.ld.populations[ii, "q"],
                           D=luo.ld.populations[ii, "D"],
                           h2=luo.ld.populations[ii, "h2"],
                           phi=luo.ld.populations[ii, "phi"],
                           Vp=100,
                           alpha=0.05)
}
cbind(luo.ld.populations, power=powers)

```

oneway.bf.alpha	<i>Correspondence between significance levels and Bayes factors for effects of marker genotype classes.</i>
-----------------	---

Description

Functions to calculate the correspondence between significance levels alpha and the Bayes factor, for association between a bi-allelic marker and QTL, for given sample sizes and marker genotype frequencies for bi-allelic marker.

Usage

```

oneway.bf.alpha(n, group.sizes = c(0.25, 0.5, 0.25) * n, alpha = 0.05)
oneway.bf.alpha.required(n, group.sizes = c(0.25, 0.5, 0.25) * n, Bf)

```

Arguments

n	Sample size, i.e. number of individuals genotyped and phenotyped for the trait
group.sizes	Number in each of the 3 possible marker genotype classes MM, Mm, mm
alpha	Significance level, i.e. threshold for ‘detection’
Bf	Bayes factor, used as threshold for detection

Details

These functions implement the correspondence between the significance levels and Bayes factors used in Ball (2003) to design experiments for detecting linkage disequilibrium with a given power to achieve a given Bayes factor. The function `SS.oneway.bf` is used to calculate the Bayes factor corresponding to a given F statistic (Spiegelhalter and Smith 1982). This is combined with a call to `qf`, for `oneway.bf.alpha` or calls to `pf` and interpolation for `oneway.bf.alpha.required`, to calculate the Bayes factor corresponding to a given alpha or alpha values for a given Bayes factor.


```

                                D=luo.ld.populations[ii,"D"],
                                h2=luo.ld.populations[ii,"h2"],
                                phi=luo.ld.populations[ii,"phi"],
                                Vp=100,
                                alpha=alphas[jj],
                                print.it=FALSE)
    }
    n.Bf20s[ii] <- approx(P.Bf20s,ns,xout=0.9)$y
    cat("n =",n.Bf20s[ii],"\n")
  }
  cbind(luo.ld.populations,powers,n.Bf20s)

```

SS.oneway.bf

Bayes factors for one-way analysis of variance models.

Description

Function to calculate the Bayes factor for a one-way analysis of variance layout with vague or improper priors.

Usage

```
SS.oneway.bf(group.sizes, Fstat)
```

Arguments

`group.sizes` Sizes of groups in the one-way layout
`Fstat` F statistic obtained

Details

The function the Bayes factor corresponding to a given F statistic in a one-way analysis of variance model is calculated using the method of Spiegelhalter and Smith 1982. With improper priors the marginal probabilities of the data under each of the models (corresponding to the NULL and alternative hypotheses) is indeterminate. This is resolved by updating each prior with a small imaginary training sample, which is equivalent to normalising the Bayes factor to be 1 for the small training sample. Spiegelhalter and Smith obtain a formula for the Bayes factor in terms of the classical F value.

Value

Returns the Bayes factor corresponding to the given design and observed value of F statistic.

Author(s)

Rod Ball (rod.ball@forestresearch.co.nz) www.forestresearch.co.nz

References

Spiegelhalter, D. and A.F.M. Smith 1982 Bayes factors for linear and log-linear models with vague prior information *J. Royal Statist Soc. B* 44: 377–387.

See Also

[oneway.bf.alpha](#), [oneway.bf.alpha.required](#)

Examples

```
# Bayes factors corresponding to P-values 0.05,0.01,0.001,0.0001 for n=200
SS.oneway.bf(group.sizes=c(50,100,50),Fstat=qf(0.95,2,197))
SS.oneway.bf(group.sizes=c(50,100,50),Fstat=qf(0.99,2,197))
SS.oneway.bf(group.sizes=c(50,100,50),Fstat=qf(0.999,2,197))
SS.oneway.bf(group.sizes=c(50,100,50),Fstat=qf(0.9999,2,197))
```

Index

*Topic **datasets**

luo.ld.populations, 5

*Topic **design**

ld.design, 1

ld.sim, 3

luo.ld.power, 6

*Topic **htest**

oneway.bf.alpha, 8

SS.oneway.bf, 9

*Topic **models**

oneway.bf.alpha, 8

SS.oneway.bf, 9

ld.design, 1

ld.power (*ld.design*), 1

ld.sim, 3, 3, 7

ld.sim1 (*ld.sim*), 3

luo.ld.populations, 5

luo.ld.power, 2–5, 6, 9

oneway.bf.alpha, 3, 7, 8, 10

oneway.bf.alpha.required, 2, 3, 7, 10

SS.oneway.bf, 2, 3, 7, 8, 9, 9