Package 'XHWE'

June 3, 2015

Type Package

Title X Chromosome Hardy-Weinberg Equilibrium

Version 1.0

Date 2015-06-02

Author Xiao-Ping You, Qi-Lei Zou, Jian-Long Li, Ji-Yuan Zhou

Maintainer Xiao-Ping You <13268259225@163.com>

Description Conduct the likelihood ratio tests for Hardy-Weinberg equilibrium at marker loci on the X chromosome.

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2015-06-03 11:47:55

R topics documented:

---

emc *Parameter Estimation by EM Algorithm When the Male and Female Allele Frequencies Are the Same*

Description

The maximum likelihood estimates of the combined allele frequencies and the inbreeding coefficient by EM algorithm when the male and female allele frequencies are the same.

Usage

```
emc(n1m, n0m, n2f, n1f, n0f, nm, nf, rho, dv, itertime)
```

Arguments

n1m	The number of males with allele 1.
n0m	The number of males with allele 2.
n2f	The number of females with genotype (1, 1).
n1f	The number of females with genotype (1, 2).
n0f	The number of females with genotype (2, 2).
nm	The number of all the males to be analyzed.
nf	The number of all the females to be analyzed.
rho	The initial value of the inbreeding coefficient.
dv	The convergence criterion for the absolute difference between the estimates of the log-likelihood function at two consecutive iterations. The default value is 1e-7.
itertime	The maximum number of iterations. The default value is 1000.

Details

The maximum likelihood estimates of the combined allele frequencies and the inbreeding coefficient by EM algorithm when the male and female allele frequencies are the same.

Value

rho.last The estimate of the inbreeding coefficient rho by EM algorithm.

pc.last The estimate of the combined allele frequency p by EM algorithm.

Author(s)

Xiao-Ping You, Qi-Lei Zou, Jian-Long Li, Ji-Yuan Zhou

References

You XP, Zou QL, Li JL, Zhou JY. 2015 Likelihood Ratio Tests for Hardy-Weinberg Equilibrium at Marker Loci on X Chromosome. (submitted)

Zheng G, Joo J, Zhang C, Geller NL. 2007 Testing Association for Markers on the X Chromosome. Genetic Epidemiology, 31: 834-843.

Dempster AP, Laird NM, Rubin DB. 1977 Maximum Likelihood from Incomplete Data via the EM algorithm. Journal of the Royal Statistical Society Series B(Methodological), 39: 1-38.

See Also

XHWE, ped, emf, result, Likelihoodfun.

emf *Parameter Estimation by EM Algorithm When the Male and Female Allele Frequencies Are Different*

Description

The maximum likelihood estimates of the male allele frequencies, the female allele frequencies and the inbreeding coefficient by EM algorithm when the male and female allele frequencies are different.

Usage

```
emf(n1m, n0m, n2f, n1f, n0f, nm, nf, rho, dv, itertime)
```

Arguments

n1m	The number of males with allele 1.
n0m	The number of males with allele 2.
n2f	The number of females with genotype (1, 1).
n1f	The number of females with genotype (1, 2).
n0f	The number of females with genotype (2, 2).
nm	The number of all the males to be analyzed.
nf	The number of all the females to be analyzed.
rho	The initial value of the inbreeding coefficient.
dv	The convergence criterion for the absolute difference between the estimates of the log-likelihood function at two consecutive iterations. The default value is 1e-7.
itertime	The maximum number of iterations. The default value is 1000.

Details

The maximum likelihood estimates of the male allele frequencies, the female allele frequencies and the inbreeding coefficient by EM algorithm when the male and female allele frequencies are different.

Value

rho.last The estimate of the inbreeding coefficient rho by EM algorithm.

pm.last The estimate of the male allele frequency pm by EM algorithm.

pf.last The estimate of the female allele frequency pf by EM algorithm.

Author(s)

Xiao-Ping You, Qi-Lei Zou, Jian-Long Li, Ji-Yuan Zhou

References

You XP, Zou QL, Li JL, Zhou JY. 2015 Likelihood Ratio Tests for Hardy-Weinberg Equilibrium at Marker Loci on X Chromosome. (submitted)

Zheng G, Joo J, Zhang C, Geller NL. 2007 Testing Association for Markers on the X Chromosome. Genetic Epidemiology, 31: 834-843.

Dempster AP, Laird NM, Rubin DB. 1977 Maximum Likelihood from Incomplete Data via the EM algorithm. Journal of the Royal Statistical Society Series B(Methodological), 39: 1-38.

See Also

XHWE, ped, emc, result, Likelihoodfun.

`Likelihoodfun` *Calculating log-Likelihood Function*

Description

Calculate the natural logarithm of the likelihood function.

Usage

```
Likelihoodfun(p1m, p2f, p1f, n1m, n0m, n2f, n1f, n0f)
```

Arguments

`p1m`	The male allele frequency of allele 1.
`p2f`	The female genotype frequency of genotype (1, 1).
`p1f`	The female genotype frequency of genotype (1, 2).
`n1m`	The number of males with allele 1.
`n0m`	The number of males with allele 2.
`n2f`	The number of females with genotype (1, 1).
`n1f`	The number of females with genotype (1, 2).
`n0f`	The number of females with genotype (2, 2).

Details

The values of p1m, p2f and p1f should be from 0 to 1, otherwise NaN is presented.

Value

The function returns the value of the natural logarithm of the likelihood function.

Author(s)

Xiao-Ping You, Qi-Lei Zou, Jian-Long Li, Ji-Yuan Zhou

References

You XP, Zou QL, Li JL, Zhou JY. 2015 Likelihood Ratio Tests for Hardy-Weinberg Equilibrium at Marker Loci on X Chromosome. (submitted)

Zheng G, Joo J, Zhang C, Geller NL. 2007 Testing Association for Markers on the X Chromosome. Genetic Epidemiology, 31: 834-843.

Dempster AP, Laird NM, Rubin DB. 1977 Maximum Likelihood from Incomplete Data via the EM algorithm. Journal of the Royal Statistical Society Series B(Methodological), 39: 1-38.

See Also

XHWE, ped, emf, result, emc.

`ped` *A Standard Linkage Pedigree File*

Description

The first 5 columns give the pedigree id, individual id, father id (0 if founder), mother id (0 if founder), and sex (1=male, 2=female, 0=unknown), in that order. Note that these fields need to be numeric. The 6th column is the affection status (1 = unaffected, 2 = affected). From the 7th column, every consecutive two columns are the alleles at a SNP marker. For males, these two alleles at a SNP marker are the same, because males have only one X chromosome.

Usage

```
data("ped")
```

Format

A data frame with 1000 observations on the following 10 variables.

`FamID` The pedigree id.

`IID` The individual id.

`FID` The father id.

`MID` The mother id.

`gender` The sex of the individual.

`Status` The affection status.

`SNP1_1` The allele at a SNP1 marker.

`SNP1_2` The allele at a SNP1 marker.

`SNP2_1` The allele at a SNP2 marker.

`SNP2_2` The allele at a SNP2 marker.

Details

The first 5 columns give the pedigree id, individual id, father id (0 if founder), mother id (0 if founder), and sex (1=male, 2=female, 0=unknown), in that order. Note that these fields need to be numeric. The 6th column is the affection status (1 = unaffected, 2 = affected). From the 7th column, every consecutive two columns are the alleles at a SNP marker. For males, these two alleles at a SNP marker are the same, because males have only one X chromosome.

Examples

```
data(ped)
```

`result` *Test Statistics, P Values and Parameter Estimation*

Description

The function outputs the values of all the test statistics (Tstat: LRT0, LRT1, LRT2, z0, z1, and z2) and the corresponding P values (Pvalue). Also, the code outputs the estimates of all the parameters under both the null and alternative hypotheses for each test statistic (Estimates_H1: the parameter estimates under H1; Estimates_H01: the parameter estimates under H01; Estimates_H02: the parameter estimates under H02; Estimates_H0: the parameter estimates under H0).

Usage

```
result(ped, loci, dv, start.rho, simuno, status_missing, allele_missing, header,
itertime, SNP_name)
```

Arguments

`ped`	The name of a standard linkage pedigree file or a matrix/dataframe containing pedigree relationship, genotype, and phenotype information, one row for each individual. The first 5 columns give the pedigree id, individual id, father id (0 if founder), mother id (0 if founder), and sex (1=male, 2=female, 0=unknown), in that order. Note that these fields need to be numeric. The 6th column is the affection status (1 = unaffected, 2 = affected). From the 7th column, every consecutive two columns are the alleles at a SNP marker. For males, these two alleles at a SNP marker are the same, because males have only one X chromosome.
`loci`	The name of a standard linkage loci file. Note that only SNP markers are suitable for the code. For each SNP locus, there are two alleles 1 and 2.
`dv`	The convergence criterion for the absolute difference between the estimates of the log-likelihood function at two consecutive iterations. The default value is 1e-7.
`start.rho`	The initial value of the inbreeding coefficient for iterations, which should be taken to be larger than 0. The default value is 0.02.
`simuno`	The number of bootstrap replications. The default value is 1000.

`status_missing`	The input variable "status_missing" is the missing value for the affection status in the data file, and the default value is 9. It can take NA, but cannot take 1, 2, or any other string values.
`allele_missing`	The input variable "allele_missing" represents the missing value for the allele. It may be 9 in some data files; or other numeric values; the default value is 0. It cannot be NA, 1, 2, or string values.
`header`	The header of input data. If ped contains variable names, then set header = TRUE (or T). The names of SNPs will be used as row names of the output statistics and P values. Otherwise set header = FALSE (or F). The default value is TRUE (or T).
`itertime`	The maximum number of iterations. The default value is 1000.
`SNP_name`	The name of SNP being analyzed.

Details

This code contains the likelihood ratio tests for Hardy-Weinberg equilibrium (HWE) at SNP markers on X chromosome. The code only uses all the founders with the genotypes available from the input pedigree file. The results may be different for different runs due to the parametric bootstrap techniques.

Value

test.testa A data frame that contains all the test statistics.

test.pvalue A data frame that contains all the P values.

test.para A data frame that contains all the parameter estimates.

Note

ped is required.

Author(s)

Xiao-Ping You, Qi-Lei Zou, Jian-Long Li, Ji-Yuan Zhou

References

You XP, Zou QL, Li JL, Zhou JY. 2015 Likelihood Ratio Tests for Hardy-Weinberg Equilibrium at Marker Loci on X Chromosome. (submitted)

Zheng G, Joo J, Zhang C, Geller NL. 2007 Testing Association for Markers on the X Chromosome. Genetic Epidemiology, 31: 834-843.

Dempster AP, Laird NM, Rubin DB. 1977 Maximum Likelihood from Incomplete Data via the EM algorithm. Journal of the Royal Statistical Society Series B(Methodological), 39: 1-38.

See Also

XHWE, ped, emf, emc, Likelihoodfun.

XHWE *Likelihood Ratio Tests for Hardy-Weinberg Equilibrium (HWE) at SNP Markers on X Chromosome*

Description

Likelihood ratio tests for Hardy-Weinberg equilibrium (HWE) at marker loci on X chromosome.

Usage

```
XHWE(ped, loci = NULL, header = T, status_missing = 9, allele_missing = 0,
start.rho = 0.02, simuno = 1000, dv = 1e-07, itertime = 1000, filename = "results.txt")
```

Arguments

ped	The name of a standard linkage pedigree file or a matrix/dataframe containing pedigree relationship, genotype, and phenotype information, one row for each individual. The first 5 columns give the pedigree id, individual id, father id (0 if founder), mother id (0 if founder), and sex (1=male, 2=female, 0=unknown), in that order. Note that these fields need to be numeric. The 6th column is the affection status (1 = unaffected, 2 = affected). From the 7th column, every consecutive two columns are the alleles at a SNP marker. For males, these two alleles at a SNP marker are the same, because males have only one X chromosome.
loci	The name of a standard linkage loci file. Note that only SNP markers are suitable for the code. For each SNP locus, there are two alleles 1 and 2.
header	The header of input data. If ped contains variable names, then set header = TRUE (or T). The names of SNPs will be used as row names of the output statistics and P values. Otherwise set header = FALSE (or F). The default value is TRUE (or T).
status_missing	The input variable "status_missing" is the missing value for the affection status in the data file, and the default value is 9. It can take NA, but cannot take 1, 2, or any other string values.
allele_missing	The input variable "allele_missing" represents the missing value for the allele. It may be 9 in some data files; or other numeric values; the default value is 0. It cannot be NA, 1, 2, or string values.
start.rho	The initial value of the inbreeding coefficient for iterations, which should be taken to be larger than 0. The default value is 0.02.
simuno	The number of bootstrap replications. The default value is 1000.
dv	The convergence criterion for the absolute difference between the estimates of the log-likelihood function at two consecutive iterations. The default value is 1e-7.
itertime	The maximum number of iterations. The default value is 1000.
filename	The name of the output file together with its physical path. The default filename is "results.txt".

Details

This code contains the likelihood ratio tests for Hardy-Weinberg equilibrium (HWE) at SNP markers on X chromosome. The code only uses all the founders with the genotypes available from the input pedigree file. The results may be a little different for different runs due to the parametric bootstrap techniques.

Value

The function outputs the values of all the test statistics (Tstat: LRT0, LRT1, LRT2, z0, z1, and z2) and the corresponding P values (Pvalue). Also, the code outputs the estimates of all the parameters under both the null and alternative hypotheses for each test statistic (Estimates_H1: the parameter estimates under H1; Estimates_H01: the parameter estimates under H01; Estimates_H02: the parameter estimates under H02; Estimates_H0: the parameter estimates under H0).

Note

Ped is required.

Author(s)

Xiao-Ping You, Qi-Lei Zou, Jian-Long Li, Ji-Yuan Zhou

References

You XP, Zou QL, Li JL, Zhou JY. 2015 Likelihood Ratio Tests for Hardy-Weinberg Equilibrium at Marker Loci on X Chromosome. (submitted)

Zheng G, Joo J, Zhang C, Geller NL. 2007 Testing Association for Markers on the X Chromosome. Genetic Epidemiology, 31: 834-843.

Dempster AP, Laird NM, Rubin DB. 1977 Maximum Likelihood from Incomplete Data via the EM algorithm. Journal of the Royal Statistical Society Series B(Methodological), 39: 1-38.

See Also

emc, ped, emf, result, Likelihoodfun.

Examples

```
data(ped)
XHWE(ped,loci=NULL,header=T,simuno=100,filename="results.txt")

##XHWE("ped.txt",loci=NULL,header=T,filename="results.txt")
##if the pedigrees are saved in a text file
```

Index