Package 'xlink'

August 20, 2019

Title Genetic Association Models for X-Chromosome SNPS on Continuous, Binary and Survival Outcomes

Version 1.0.1

Maintainer Yi Zhu <yizhu87@gmail.com>

URL https://github.com/qiuanzhu/xlink

BugReports https://github.com/qiuanzhu/xlink/issues

Description

The expression of X-chromosome undergoes three possible biological processes: X-chromosome inactivation (XCI), escape of the X-chromosome inactivation (XCI-E),and skewed X-chromosome
inactivation (XCI-S). To analyze the X-linked genetic association for phenotype such as continuous, binary, and time-to-
event outcomes with the actual process unknown, we propose a unified
approach of maximizing the likelihood or partial likelihood over all of the potential biological processes. The methods are described in Wei Xu, Meil-
ing Hao (2017) <doi:10.1002/gepi.22097>. And
also see Dongxiao Han, Meiling Hao, Lian-
qiang Qu, Wei Xu (2019) <doi:10.1177/0962280219859037>.

License GPL-2

Encoding UTF-8

LazyData true

Depends R (>= 3.1.0)

Imports survival (>= 2.41.3)

RoxygenNote 6.1.1

Suggests knitr, rmarkdown, testthat

VignetteBuilder knitr

NeedsCompilation no

Author Wei Xu [aut],
Meiling Hao [aut],
Yi Zhu [cre]

Repository CRAN

Date/Publication 2019-08-20 21:40:02 UTC

R topics documented:

`fit_all_models` *Model fitting results for each SNP considering as XCI, XCI-E and XCI-S type*

Description

`fit_all_models` returns model fitting results for each SNP understanding as XCI, XCI-E and XCI-S type respectively. Model comparison results is provided by using AIC as a criterion.

Usage

```
fit_all_models(resp, os, ostime, snp, gender, male, female, covars, model,
  data)
```

Arguments

`resp`	Response variable for continuous or binary model fitting.
`os`	Survival indicator, 1 for death, 0 for censoring.
`ostime`	Duration time of survival.
`snp`	Single SNP name.
`gender`	Gender variable.
`male`	Male indicator in gender variable.
`female`	Female indicator in gender variable.
`covars`	Covariates list.
`model`	Fitting model type. For 'linear', fitting linear model. For 'binary', fitting logistic regression model. For 'survival', fitting survival model.
`data`	Data set.

Value

It returns estimated parameters, confidence interval and P value for each variable. Baseline model and full model maximum likelihood estimation are provided.

See Also

lm{stats} for linear model, glm{stats} for logistic regression model, and coxph{survival} for survival model.

`fit_XCI_E_model` *Model fitting results for each SNP considering as XCI-E type*

Description

`fit_XCI_E_model` returns model fitting results for each SNP understanding as XCI-E type.

Usage

```
fit_XCI_E_model(resp, os, ostime, snp, gender, male, female, covars, model,
  data)
```

Arguments

`resp`	Response variable for continuous or binary model fitting.
`os`	Survival indicator, 1 for death, 0 for censoring.
`ostime`	Duration time of survival.
`snp`	Single SNP name.
`gender`	Gender variable.
`male`	Male indicator in gender variable.
`female`	Female indicator in gender variable.
`covars`	Covariates list.
`model`	Fitting model type. For 'linear', fitting linear model. For 'binary', fitting logistic regression model. For 'survival', fitting survival model.
`data`	Data set.

Value

It returns estimated parameters, confidence interval and P value for each variable. Baseline model and full model maximum likelihood estimation are provided.

See Also

lm{stats} for linear model, glm{stats} for logistic regression model, and coxph{survival} for survival model.

`fit_XCI_model` *Model fitting results for each SNP considering as XCI type*

Description

`fit_XCI_model` returns model fitting results for each SNP understanding as XCI type.

Usage

```
fit_XCI_model(resp, os, ostime, snp, gender, male, female, covars, model,
  data)
```

Arguments

`resp`	Response variable for continuous or binary model fitting.
`os`	Survival indicator, 1 for death, 0 for censoring.
`ostime`	Duration time of survival.
`snp`	Single SNP name.
`gender`	Gender variable.
`male`	Male indicator in gender variable.
`female`	Female indicator in gender variable.
`covars`	Covariates list.
`model`	Fitting model type. For 'linear', fitting linear model. For 'binary', fitting logistic regression model. For 'survival', fitting survival model.
`data`	Data set.

Value

It returns estimated parameters, confidence interval and P value for each variable. Baseline model and full model maximum likelihood estimation are provided.

See Also

`lm{stats}` for linear model, `glm{stats}` for logistic regression model, and `coxph{survival}` for survival model.

infor_table *Information table.*

Description

infor_table returns information table of estimated coefficients/hazard ratio, confidence interval and P value.

Usage

```
infor_table(x, snp, covar_n, MAF_value, model)
```

Arguments

x	An output from continuous/binary/survival model.
snp	Single SNP name.
covar_n	Covariate names.
MAF_value	A minimum value of minor allele frequency.
model	Model type.

Value

Information table. If linear or binary model is chosen, it returns estimated coefficients, confidence interval and P value. If survival model is chosen, it returns hazard ratio, confidence interval and P value.

MAF *Select SNP by MAF.*

Description

MAF returns SNPs with higher MAF than default value.

Usage

```
MAF(snp, gender, male, MAF_v, data)
```

Arguments

snp	SNP name.
gender	gender variable.
male	male information.
MAF_v	minimum MAF value.
data	a dataset.

Rdata *Simulation data for Genetic association models for X-chromosome SNPS*

Description

A simulated dataset containing 400 observations. The variables list as follows:

Usage

```
data(Rdata)
```

Format

A data frame with 400 rows and 17 variables.

Details

- ID. Identification number.
- OS. Survival indicator, 1 for death, 0 for censoring.
- OS_time. Duration time of survival.
- gender. Binary value 0,1 with P(x=1)=0.5 and hazard ratio is 1.
- Age. Uniform distribution in [20,80] and hazard ratio is 1.02.
- Smoking. Binary value 0,1 with P(x=1)=0.3 and hazard ratio is 1.2.
- Treatment. Binary value 0,1 with P(X=1)=0.3 and hazard ratio is 1.2.
- snp_1. True type in coxph model is 'XCI', minor allele frequency is 0.2, hazard ratio is 1.5.
- snp_2. True type in coxph model is 'XCI-E', minor allele frequency is 0.3, hazard ratio is 1.5.
- snp_3. True type in coxph model is 'XCI-S', minor allele frequency is 0.4, hazard ratio is 1.5.
- snp_4. True type in coxph model is 'XCI', minor allele frequency is 0.3, hazard ratio is 1.
- snp_5. True type in coxph model is 'XCI-E', minor allele frequency is 0.1, hazard ratio is 1.
- snp_6. True type in coxph model is 'XCI', minor allele frequency is 0.2, hazard ratio is 1.
- snp_7. True type in coxph model is 'XCI', minor allele frequency is 0.15, hazard ratio is 1.
- snp_8. True type in coxph model is 'XCI-E', minor allele frequency is 0.1, hazard ratio is 1.
- snp_9. True type in coxph model is 'XCI', minor allele frequency is 0.2, hazard ratio is 1.
- snp_10. True type in coxph model is 'XCI', minor allele frequency is 0.15, hazard ratio is 1.

`select_output` *Selected results table by P value*

Description

`select_output` returns selected SNP information by P value.

Usage

```
select_output(input, pv_thold = 1)
```

Arguments

`input`	Input results from xlink fit.
`pv_thold`	P value threshold for output.

Value

It returns estimated parameters, confidence interval, P value, MAF and Best model information.

See Also

`xlink_fit{xlink}` for input results.

Examples

```
Covars<-c("Age","Smoking","Treatment")
SNPs<-c("snp_1","snp_2","snp_3")
result<-xlink_fit(os="OS",ostime ="OS_time",snps=SNPs,gender ="gender",covars=Covars,
                  option =list(type="all",MAF_v=0.05), model="survival", data = Rdata)
select_output(input=result,pv_thold=10^-5)
```

`xlink` *xlink: A package for genetic association models for X-chromosome SNPS on continuous, binary and survival outcomes.*

Description

The expression of X-chromosome undergoes three possible biological processes: X-chromosome inactivation (XCI), escape of the X-chromosome inactivation (XCI-E),and skewed X-chromosome inactivation (XCI-S). To analyze the X-linked genetic association for phenotype such as continuous, binary, and time-to-event outcomes with the actual process unknown, we propose a unified approach of maximizing the likelihood or partial likelihood over all of the potential biological processes.

xlink functions

xlink_fit, select_output.

References

Xu, Wei, and Meiling Hao. 'A unified partial likelihood approach for X-chromosome association on time-to-event outcomes.' Genetic epidemiology 42.1 (2018): 80-94.

`xlink_fit` *Genetic association models for X-chromosome SNPs on continuous, binary and survival outcomes*

Description

`xlink_fit` returns model fitting results for each SNP with the covariates.

Usage

```
xlink_fit(resp = c(), os = c(), ostime = c(), snps = c(),
  gender = c(), covars = c(), option = c(type = c(), male = c(),
  female = c(), MAF_v = 0), model = c(), data)
```

Arguments

`resp`	Response variable for continuous or binary model fitting.
`os`	Survival indicator, 1 for death, 0 for censoring.
`ostime`	Duration time of survival.
`snps`	SNP name list for model fitting.
`gender`	Gender information must be included in the data. Default setting is male=1 and female=0. If not as default setting, please provide male and female information in the option.
`covars`	Covariates list if needed.
`option`	There are three options. First, type has default 'all', which provides model fitting results for each SNP understanding as 'XCI', 'XCI-E' and 'XCI-S' type respectively. If type is chosen as 'XCI' or 'XCI-E', all the SNPS consider as 'XCI' or 'XCI-E' type in corresponding model. Secondly, if gender is not as default gender setting (male=1,female=0), male and female information should be provided here.The third one, MAF_v is the low bound of the minimum allele frequency, the SNP MAF below this value will not be used in xlink_fit.
`model`	Fitting model. For 'linear', fitting linear model. For 'binary', fitting logistic regression model. For 'survival', fitting survival model.
`data`	Data set.

Value

It returns estimated parameters, confidence interval and P value for each variable in the chosen model. The baseline and full model maximum likelihood estimation are provided. If type is 'all', best model choice is provided by using AIC as an benchmark.

References

Xu, Wei, and Meiling Hao. 'A unified partial likelihood approach for X-chromosome association on time-to-event outcomes.' Genetic epidemiology 42.1 (2018): 80-94.

Han, D., Hao, M., Qu, L., & Xu, W. (2019). A novel model for the X-chromosome inactivation association on survival data. Statistical Methods in Medical Research.

See Also

lm{stats} for linear model, glm{stats} for logistic regression model, and coxph{survival} for survival model.

Examples

```
Covars<-c("Age","Smoking","Treatment")
SNPs<-c("snp_1","snp_2","snp_3")
xlink_fit(os="OS",ostime="OS_time",snps=SNPs,gender="gender",covars=Covars,
          option =list(MAF_v=0.05),model="survival",data = Rdata)
xlink_fit(resp="OS_time",snps=SNPs,gender="gender",option =list(type="XCI",MAF_v=0.05),
          model="linear",data = Rdata)
```

Index