Package ‘fabCI’

January 7, 2021

Title FAB Confidence Intervals
Version 0.2
Description Frequentist assisted by Bayes (FAB) confidence interval construction. See 'Adaptive multigroup confidence intervals with constant coverage' by Yu and Hoff <DOI:10.1093/biomet/asy009> and 'Exact adaptive confidence intervals for linear regression coefficients' by Hoff and Yu <DOI:10.1214/18-EJS1517>.
License GPL-3
Encoding UTF-8
LazyData true
Imports MASS
Date 2021-01-07
Author Peter Hoff and Chaoyu Yu
Maintainer Peter Hoff <peter.hoff@duke.edu>
RoxygenNote 5.0.1
NeedsCompilation no
Repository CRAN
Date/Publication 2021-01-07 15:10:02 UTC

R topics documented:
ebayes_est ................................. 2
fabregCI ................................. 3
fabCI ..................................... 3
fabtzCI ................................ 4
fabzCI .................................. 5
hhetmodel ................................ 6
hhommodel ................................ 7
multifabCI ............................... 8
multifabClhom ......................... 9
radon .................................. 10
sfabz .................................. 10
umauregCI .............................. 11
Description

Compute empirical Bayes estimates of the error variance and distribution of the regression coefficients.

Usage

ebayes_est(y, X, emu = FALSE, dof = min(50, round(0.5 * (dim(X)[1] - dim(X)[2]))))

Arguments

- y: a numeric vector of data
- X: a design matrix
- emu: (logical) estimate mean of coefficient (TRUE) or assume it is zero (FALSE)?
- dof: degrees of freedom to use for the t-quantiles (the remainder go to adaptive estimation of the prior)

Details

This function computes the adaptive FAB confidence interval for each coefficient in a linear regression model.

Value

A list (s,sigma2,tau2,tau2) where

1. s: an estimate of the error standard deviation
2. sigma2: an estimate of the error variance, independent of s
3. tau2: an estimate of the coefficient variance, independent of s
4. mu: an estimate of the coefficient mean, independent of s

Author(s)

Peter Hoff
**fabregCI**  

*FAB regression coefficient intervals*

**Description**

Compute the adaptive FAB t-intervals for the coefficients of a regression model.

**Usage**

```r
fabregCI(y, X, alpha = 0.05, dof = min(50, round(0.5 * (dim(X)[1] -
             dim(X)[2]))), verbose = TRUE)
```

**Arguments**

- `y`: a numeric vector of data
- `X`: a design matrix
- `alpha`: the type I error rate, so 1-alpha is the coverage rate
- `dof`: degrees of freedom to use for the t-quantiles (the remainder go to adaptive estimation of the prior)
- `verbose`: logical, print progress or not

**Details**

This function computes the adaptive FAB confidence interval for each coefficient in a linear regression model.

**Value**

A matrix where each row corresponds to the interval and OLS estimate of a coefficient.

**Author(s)**

Peter Hoff

---

**fabtCI**  

*FAB t-interval*

**Description**

Computation of a 1-alpha FAB t-interval

**Usage**

```r
fabtCI(y, psi = c(0, 100, 1, 2), alpha = 0.05)
```
Arguments

- **y**: A numeric vector with at least two non-missing values
- **psi**: A length-four vector of hyperparameters for the prior
- **alpha**: The type I error rate, so 1-alpha is the coverage rate

Details

A FAB interval is the “frequentist” interval procedure that is Bayes optimal: It minimizes the prior expected interval width among all interval procedures with exact 1-alpha frequentist coverage. This function computes the FAB t-interval for the mean of a normal population with an unknown variance, given a user-specified prior distribution determined by psi. The prior is that the population mean and variance are independently distributed as normal and inverse-gamma random variables. Referring to the elements of psi as mu, t2, s20, nu0, the prior is determined as follows:

1. mu is the prior expectation of the mean
2. t2 is the prior variance of the mean
3. the population variance is inverse-gamma(nu0/2,nu0 s20/2)

Author(s)

Peter Hoff

Examples

```r
y <- rnorm(10)
fabtCI(y, c(0, 10, 1, 5))
fabtCI(y, c(0, 1/10, 1, 5))
fabtCI(y, c(2, 10, 1, 5))
fabtCI(y, c(0, 1/10, 1, 5))
```

Description

Computation of a 1-alpha FAB t-interval using z-optimal spending function

Usage

```r
fabtzCI(y, s, dof, alpha = 0.05, psi = list(mu = 0, tau2 = 1e+05, sigma2 = 1))
```
Arguments

- **y**: a numeric scalar, a normally distributed statistic
- **s**: a numeric scalar, the standard error of y
- **dof**: positive integer, degrees of freedom for s
- **alpha**: the type I error rate, so 1-alpha is the coverage rate
- **psi**: a list of parameters for the spending function, including
  1. **mu**: the prior expectation of E[y]
  2. **tau2**: the prior variance of E[y]
  3. **sigma2**: the variance of y

Examples

```r
n<-10
y<-rnorm(n)
fabtzCI(mean(y),sqrt(var(y)/n),n-1)
t.test(y)$conf.int
```

Description

Computation of a 1-alpha FAB z-interval

Usage

```r
fabzCI(y, mu, t2, s2, alpha = 0.05)
```

Arguments

- **y**: a numeric scalar
- **mu**: a numeric scalar
- **t2**: a positive numeric scalar
- **s2**: a positive numeric scalar
- **alpha**: the type I error rate, so 1-alpha is the coverage rate

Details

A FAB interval is the "frequentist" interval procedure that is Bayes optimal: It minimizes the prior expected interval width among all interval procedures with exact 1-alpha frequentist coverage. This function computes the FAB z-interval for the mean of a normal population with an known variance, given a user-specified prior distribution determined by psi. The prior is that the population mean is normally distributed. Referring to the elements of psi as mu, t2, s2, the prior and population variance are determined as follows:
1. mu is the prior expectation of the mean
2. t2 is the prior variance of the mean
3. s2 is the population variance

Author(s)
Peter Hoff

Examples
y<-0
fabzCI(y,0,10,1)
fabzCI(y,0,1/10,1)
fabzCI(y,2,10,1)
fabzCI(y,0,1/10,1)

hhetmodel
Hierarchical heteroscedastic model estimates

Description
Estimate across-group heterogeneity of means and variances

Usage
hhetmodel(y, g)

Arguments
y a numeric vector of data
g a group membership vector, of the same length as y

Details
This function estimates parameters in a hierarchical model for normally distributed groups, where the across-group model for means is normal and the across group model for variances is inverse-gamma.

Value
A vector (mu,t2,s20,nu0), where
1. mu is the mean of the group means
2. t2 is the variance of the group means
3. the the distribution of group variances is inverse-gamma(nu0/2,nu0 s20/2)

Author(s)
Peter Hoff
**hhommodel**

**Hierarchical homoscedastic model estimates**

**Description**

Estimate across-group heterogeneity of means

**Usage**

```r
hhommodel(y, g, group, p1)
```

**Arguments**

- `y`: a numeric vector of data
- `g`: a group membership vector, of the same length as `y`
- `group`: the index of the group
- `p1`: number of groups used to pool sample variance

**Details**

This function estimates parameters in a hierarchical model for normally distributed groups, where the across-group model for means is normal and the variance is the same across groups.

**Value**

A vector \((s2, df, muw, t2w, s2w)\), where

1. \(s2\) is the pooled variance
2. \(df\) is the degree of freedom of the t-quantiles
3. \(muw\) is the estimate mean of the group means
4. \(t2w\) is the estimate variance of the group means
5. \(s2w\) is the estimate within-group variance

**Author(s)**

Chaoyu Yu
Description

Computation of 1-alpha FAB t-intervals for heteroscedastic multigroup data.

Usage

```r
multifabCI(y, g, alpha = 0.05)
```

Arguments

- `y`: a numeric vector of data
- `g`: a group membership vector, of the same length as `y`
- `alpha`: the type I error rate, so 1-alpha is the coverage rate

Details

For each group `j`, this function computes an estimate of the parameters in a hierarchical model for means and variances from data other than group `j`, and uses this information to construct a FAB t-interval for group `j`. These intervals have 1-alpha frequentist coverage, assuming within-group normality.

Author(s)

Peter Hoff

Examples

```r
## -- simulated data
p<-10 ; n<-10
y<-rnorm(n*p) ; g<-rep(1:p,n)
## -- more interesting data takes longer
# data(radon) ; y<-radon[,2] ; g<-radon[,1]
## -- FAB t-intervals
FCI<-multifabCI(y,g)
## -- UMAU t-intervals
ybar<-tapply(y,g,mean) ; ssd<-tapply(y,g,sd) ; n<-table(g)
qtn<-cbind( qt(.025,n-1), qt(.975,n-1) )
UCI<-sweep(sweep(qtn,1,ssd/sqrt(n),"*"),1,ybar,"+")
mean( (UCI[,2]-UCI[,1])/(FCI[,2]-FCI[,1]) , na.rm=TRUE)
```
multifabCIhom

Description

Computation of 1-alpha FAB t-intervals for homoscedastic multigroup data.

Usage

multifabCIhom(y, g, alpha = 0.05, prop = 0.5)

Arguments

y a numeric vector of data
g a group membership vector, of the same length as y
alpha the type I error rate, so 1-alpha is the coverage rate
prop the proportion of groups to obtain the sample variance estimate

Details

For each group j, this function computes an estimate of the parameters in a hierarchical model for means using data from other groups, and uses this information to construct a FAB t-interval for group j. These intervals have 1-alpha frequentist coverage, assuming within-group normality and that the within group variance is the same across groups.

Author(s)

Chaoyu Yu

Examples

```r
## -- simulate the data
mu = 0; sigma2 = 10; tau2 = 1; p =100;
theta = rnorm(p,mu,sqrt(tau2))
ns = round(runif(p,2,18))
Y=c()
for(i in 1:p){
d2 = rnorm(ns[i],theta[i],sqrt(sigma2))
d1 = rep(i,ns[i])
d = cbind(d1,d2)
Y = rbind(Y,d)}
y = Y[,2]
g = Y[,1]

## -- FAB t-intervals
FCI = multifabCIhom(y,g)

## -- UMAU t-intervals
```
radon                              

Minneapolis Radon Dataset

Radon levels in 919 homes from 85 Minnesota counties

Usage

data(radon)

Format

A numeric matrix

Source

http://www.stat.columbia.edu/~gelman/arm/software/

sfabz                              

Bayes-optimal spending function

Description

Compute Bayes optimal spending function

Usage

sfabz(theta, psi, alpha = 0.05)

Arguments

theta  value of theta being tested
psi     a list of parameters for the spending function, including
1. mu, the prior expectation of E[y]
2. tau2, the prior variance of E[y]
3. sigma2 the variance of y
alpha  level of test
umauregCI

Details
This function computes the value of s that minimizes the acceptance probability of a biased level-alpha test for a normal population with known variance, under a specified prior predictive distribution.

Author(s)
Peter Hoff

umauregCI  UMAU regression coefficient intervals

Description
Compute the usual t-intervals for the coefficients of a regression model

Usage
umauregCI(y, X, alpha = 0.05)

Arguments
y  a numeric vector of data
X  a design matrix
alpha  the type I error rate, so 1-alpha is the coverage rate

Details
This function computes the 'usual' uniformly most accurate unbiased confidence interval for each coefficient in a linear regression model.

Value
A matrix where each row corresponds to the interval and OLS estimate of a coefficient.

Author(s)
Peter Hoff
Index

* datasets
  - radon, 10

* htest
  - fabtCI, 3
  - fabzCI, 5
  - hhetmodel, 6
  - hhommodel, 7
  - multifabCI, 8
  - multifabCIhom, 9

  ebayes_est, 2

  fabregCI, 3
  fabtCI, 3
  fabtzCI, 4
  fabzCI, 5

  hhetmodel, 6
  hhommodel, 7

  multifabCI, 8
  multifabCIhom, 9

  radon, 10

  sfabz, 10

  umauregCI, 11