Package ‘epsiwal’

October 13, 2022

Version 0.1.0

Date 2019-06-28

License LGPL-3

Title Exact Post Selection Inference with Applications to the Lasso

BugReports https://github.com/shabbychef/epsiwal/issues

Description Implements the conditional estimation procedure of Lee, Sun, Sun and Taylor (2016) <doi:10.1214/15-AOS1371>. This procedure allows hypothesis testing on the mean of a normal random vector subject to linear constraints.

Depends R (>= 3.0.2)

Suggests testthat

URL https://github.com/shabbychef/epsiwal

Collate 'ci_connorm.r' 'epsiwal.r' 'pconnorm.r' 'ptruncnorm.r' 'utils.r'

RoxygenNote 6.1.1

NeedsCompilation no

Author Steven E. Pav [aut, cre] (<https://orcid.org/0000-0002-4197-6195>)

Maintainer Steven E. Pav <shabbychef@gmail.com>

Repository CRAN

Date/Publication 2019-07-02 12:10:10 UTC

R topics documented:

ci_connorm *ci_connorm* .

Description

Confidence intervals on normal mean, subject to linear constraints.

Usage

```
ci_connorm(y, A, b, eta, Sigma = NULL, p = c(level/2, 1 - (level/2)),
  level = 0.05, Sigma_eta = Sigma %*% eta)
```

Arguments

y	an n vector, assumed multivariate normal with mean μ and covariance Σ.
A	an $k \times n$ matrix of constraints.
b	a k vector of inequality limits.
eta	an n vector of the test contrast, η.
Sigma	an $n \times n$ matrix of the population covariance, Σ. Not needed if Sigma_eta is given.
p	a vector of probabilities for which we return equivalent $\eta^\top \mu$.
level	if p is not given, we set it by default to c(level/2,1-level/2).
Sigma_eta	an n vector of $\Sigma\eta$.

Details

Inverts the constrained normal inference procedure described by Lee *et al.*

Let y be multivariate normal with unknown mean μ and known covariance Σ. Conditional on $Ay \le b$ for conformable matrix A and vector b, and given constrast vector eta and level p, we compute $\eta^\top \mu$ such that the cumulative distribution of $\eta^\top y$ equals p.

Value

The values of $\eta^\top \mu$ which have the corresponding CDF.

Note

An error will be thrown if we do not observe $Ay \le b$.

Author(s)

Steven E. Pav <shabbychef@gmail.com>

References

Lee, J. D., Sun, D. L., Sun, Y. and Taylor, J. E. "Exact post-selection inference, with application to the Lasso." Ann. Statist. 44, no. 3 (2016): 907-927. doi:10.1214/15-AOS1371. https://arxiv.org/abs/1311.6238

See Also

the CDF function, `pconnorm`.

Examples

```
set.seed(1234)
n <- 10
y <- rnorm(n)
A <- matrix(rnorm(n*(n-3)),ncol=n)
b <- A%*%y + runif(nrow(A))
Sigma <- diag(runif(n))
mu <- rnorm(n)
eta <- rnorm(n)

pval <- pconnorm(y=y,A=A,b=b,eta=eta,mu=mu,Sigma=Sigma)
cival <- ci_connorm(y=y,A=A,b=b,eta=eta,Sigma=Sigma,p=pval)
stopifnot(abs(cival - sum(eta*mu)) < 1e-4)
```

`epsiwal` *Exact Post Selection Inference with Applications to the Lasso.*

Description

Exact Post Selection Inference with Applications to the Lasso.

Details

This simple package supports the simple procedure outlined in Lee *et al.* where one observes a normal random variable, then performs inference conditional on some linear inequalities.

Suppose y is multivariate normal with mean μ and covariance Σ. Conditional on $Ay \le b$, one can perform inference on $\eta^\top \mu$ by transforming y to a truncated normal. Similarly one can invert this procedure and find confidence intervals on $\eta^\top \mu$.

Legal Mumbo Jumbo

Note

This package is maintained as a hobby.

Author(s)

Steven E. Pav <shabbychef@gmail.com>

References

Lee, J. D., Sun, D. L., Sun, Y. and Taylor, J. E. "Exact post-selection inference, with application to the Lasso." Ann. Statist. 44, no. 3 (2016): 907-927. doi:10.1214/15-AOS1371. https://arxiv.org/abs/1311.6238

Pav, S. E. "Conditional inference on the asset with maximum Sharpe ratio." Arxiv e-print (2019). http://arxiv.org/abs/1906.00573

`epsiwal-NEWS` *News for package 'epsiwal':*

Description

News for package 'epsiwal'

epsiwal Initial Version 0.1.0 (2019-06-28)

- first CRAN release.

`pconnorm` *pconnorm .*

Description

CDF of the conditional normal variate.

Usage

```
pconnorm(y, A, b, eta, mu = NULL, Sigma = NULL, Sigma_eta = Sigma
  %*% eta, eta_mu = as.numeric(t(eta) %*% mu), lower.tail = TRUE,
  log.p = FALSE)
```

Arguments

`y`	an n vector, assumed multivariate normal with mean μ and covariance Σ.
`A`	an $k \times n$ matrix of constraints.
`b`	a k vector of inequality limits.
`eta`	an n vector of the test contrast, η.
`mu`	an n vector of the population mean, μ. Not needed if `eta_mu` is given.
`Sigma`	an $n \times n$ matrix of the population covariance, Σ. Not needed if `Sigma_eta` is given.

`Sigma_eta`	an n vector of $\Sigma\eta$.
`eta_mu`	the scalar $\eta^\top\mu$.
`lower.tail`	logical; if TRUE (default), probabilities are $P[X \le x]$ otherwise, $P[X > x]$.
`log.p`	logical; if TRUE, probabilities p are given as log(p).

Details

Computes the CDF of the truncated normal conditional on linear constraints, as described in section 5 of Lee *et al.*

Let y be multivariate normal with mean μ and covariance Σ. Conditional on $Ay \le b$ for conformable matrix A and vector b we compute the CDF of a truncated normal maximally aligned with η. Inference depends on the population parameters only via $\eta^\top\mu$ and $\Sigma\eta$, and only these need to be given.

The test statistic is aligned with y, meaning that an output p-value near one casts doubt on the null hypothesis that $\eta^\top\mu$ is less than the posited value.

Value

The CDF.

Note

An error will be thrown if we do not observe $Ay \le b$.

Author(s)

Steven E. Pav <shabbychef@gmail.com>

References

Lee, J. D., Sun, D. L., Sun, Y. and Taylor, J. E. "Exact post-selection inference, with application to the Lasso." Ann. Statist. 44, no. 3 (2016): 907-927. doi:10.1214/15-AOS1371. https://arxiv.org/abs/1311.6238

See Also

the confidence interval function, `ci_connorm`.

ptruncnorm *ptruncnorm .*

Description

Cumulative distribution of the truncated normal function.

Usage

```
ptruncnorm(q, mean = 0, sd = 1, a = -Inf, b = Inf,
  lower.tail = TRUE, log.p = FALSE)
```

Arguments

q	vector of quantiles.
mean	vector of means.
sd	vector of standard deviations.
a	vector of the left truncation value(s).
b	vector of the right truncation value(s).
lower.tail	logical; if TRUE (default), probabilities are $P[X \le x]$ otherwise, $P[X > x]$.
log.p	logical; if TRUE, probabilities p are given as log(p).

Value

The distribution function of the truncated normal.

Invalid arguments will result in return value NaN with a warning.

Note

Input are recycled as possible.

Author(s)

Steven E. Pav <shabbychef@gmail.com>

References

Hattaway, James T. "Parameter estimation and hypothesis testing for the truncated normal distribution with applications to introductory statistics grades." BYU Masters Thesis (2010). https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=3052&context=etd

Examples

```
y <- ptruncnorm(seq(-5,5,length.out=101), a=-1, b=2)
```

Index