Package: epsiwal (via r-universe)

June 10, 2026

Version 0.2.0

Date 2026-06-08

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. Also supports computation of the MLE of the mean subject to the same constraints.

Depends R (>= 3.0.2)

Suggests testthat

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

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

RoxygenNote 7.3.3

NeedsCompilation no

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

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

Repository https://cran.r-universe.dev

Date/Publication 2026-06-10 06:45:53 UTC

RemoteUrl https://github.com/cran/epsiwal

RemoteRef HEAD

RemoteSha 26f962dc151f874b32e7e7572fb95c9a6a998b55

Contents

`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, the MLE function, mle_connorm, the special case code for conditioning on the max, ci_connorm_max

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)
```

ci_connorm_max *ci_connorm_max* .

Description

Confidence intervals on normal mean, conditioning on the max.

Usage

```
ci_connorm_max(
  yk,
  yk1,
  sigma = 1,
  rho = 0,
  p = c(level/2, 1 - (level/2)),
  level = 0.05
)
```

Arguments

yk	the observed maximum value, y_k.
yk1	a vector of the other observed values, y_{k1}, or just the scalar second largest value.
sigma	the common standard deviation.
rho	the common correlation.
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).

Details

Computes the confidence interval of unknown mean of a normal vector conditional on the one element being the maximum.

Let y be multivariate normal with unknown mean μ and known covariance Σ. We assume that Σ is compound symmetric with common variance σ^2 and common correlation ρ.

Conditional on $y_k \geq y_i$ for all i, we compute the confidence interval of μ_k.

Value

The values of μ_k which have the corresponding CDF.

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, the MLE function, mle_connorm_max, the more general version, ci_connorm.

---

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 $Ay \leq b$, 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>

Maintainer: Steven E. Pav <shabbychef@gmail.com> (ORCID)

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). https://arxiv.org/abs/1906.00573

Pav, S. E. "Post selection estimation of Sharpe ratios." Arxiv e-print (2026). https://arxiv.org/abs/2606.01650

See Also

Useful links:

- https://github.com/shabbychef/epsiwal
- Report bugs at https://github.com/shabbychef/epsiwal/issues

---

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

Description

News for package ‘epsiwal’

epsiwal Initial Version 0.2.0 (2026-06-08)

- fix numerical stability issues in ptruncnorm and downstream utilities (CIs).
- add MLE estimator of Reid, Taylor, Tibshirani.
- add helper functions for the case of conditioning on the max of a vector.

epsiwal Initial Version 0.1.0 (2019-06-28)

- first CRAN release.

mle_connorm *mle_connorm .*

Description

Maximum likelihood estimate of normal mean, subject to linear constraints.

Usage

```
mle_connorm(y, A, b, eta, Sigma = NULL, 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.
Sigma_eta	an n vector of $\Sigma\eta$.
...	dots are passed to `uniroot`.

Details

Computes the maximum likelihood estimate of unknown mean of a normal vector conditional on linear constraints.

Let y be multivariate normal with unknown mean μ and known covariance Σ. Conditional on $Ay \leq b$ for conformable matrix A and vector b, and given constrast vector eta, we compute the maximum likelihood estimate of $\eta^\top \mu$.

Value

The maximum likelihood estimate of $\eta^\top \mu$.

Author(s)

Steven E. Pav <shabbychef@gmail.com>

References

Reid, S., Taylor, J. and Tibshirani, R. "Post-selection point and interval estimation of signal sizes in Gaussian samples." Can. J. Statistics. 45, no. 2 (2017): 128-148. doi:10.1002/cjs.11320. https://arxiv.org/abs/1405.3340

See Also

the confidence interval function, ci_connorm, the CDF function, pconnorm, the special case code for conditioning on the max, mle_connorm_max

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)

mval <- mle_connorm(y=y,A=A,b=b,eta=eta,Sigma=Sigma)
# try again, but control tolerance:
mval <- mle_connorm(y=y,A=A,b=b,eta=eta,Sigma=Sigma,tol=1e-8)
```

---

mle_connorm_max *mle_connorm_max .*

Description

Maximum likelihood estimate of normal mean, conditioning on the max.

Usage

```
mle_connorm_max(yk, yk1, sigma = 1, rho = 0, ...)
```

Arguments

yk	the observed maximum value, y_k.
yk1	a vector of the other observed values, y_{k1}, or just the scalar second largest value.
sigma	the common standard deviation.
rho	the common correlation.
...	dots are passed to uniroot.

Details

Computes the maximum likelihood estimate of unknown mean of a normal vector conditional on the one element being the maximum.

Let y be multivariate normal with unknown mean μ and known covariance Σ. We assume that Σ is compound symmetric with common variance σ^2 and common correlation ρ.

Conditional on $y_k \geq y_i$ for all i, we compute the maximum likelihood estimate of μ_k.

Value

The maximum likelihood estimate of μ_k.

Author(s)

Steven E. Pav <shabbychef@gmail.com>

References

Reid, S., Taylor, J. and Tibshirani, R. "Post-selection point and interval estimation of signal sizes in Gaussian samples." Can. J. Statistics. 45, no. 2 (2017): 128-148. doi:10.1002/cjs.11320. https://arxiv.org/abs/1405.3340

See Also

the confidence interval function, ci_connorm_max, the CDF function, pconnorm, the more general version, mle_connorm.

---

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 <= x] otherwise, P[X > x].
log.p	logical; if TRUE, probabilities p are returned 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, the MLE function, mle_connorm.

pconnorm_max *pconnorm_max .*

Description

CDF of the conditional normal variate, conditioning on the max.

Usage

```
pconnorm_max(
  yk,
  yk1,
  mu_k,
  sigma = 1,
  rho = 0,
  lower.tail = TRUE,
  log.p = FALSE
)
```

Arguments

yk	the observed maximum value, y_k.
yk1	a vector of the other observed values, y_{k1}, or just the scalar second largest value.
mu_k	the scalar mean of the maximal element μ_k.
sigma	the common standard deviation.
rho	the common correlation.
lower.tail	logical; if TRUE (default), probabilities are P[X <= x] otherwise, P[X > x].
log.p	logical; if TRUE, probabilities p are returned as log(p).

Details

Computes the CDF of the conditional maximum of a normal vector using the truncated normal from the polyhedral lemma. Let y be multivariate normal where the maximal observed element is known to have mean μ_k, and the vector has known covariance Σ. We assume that Σ is compound symmetric with common variance σ^2 and common correlation ρ.

Conditional on $y_k \geq y_i$ for all i, we compute the CDF of y_k

Value

The CDF.

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 general CDF function, `pconnorm`, the MLE function, `mle_connorm_max`, the confidence interval function, `ci_connorm_max`.

`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 <= x] otherwise, P[X > x].
`log.p`	logical; if TRUE, probabilities p are returned 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