Package ‘ciuupi’

January 15, 2019

Type Package

Title Confidence Intervals Utilizing Uncertain Prior Information

Version 1.1.0

Description Computes a confidence interval for a specified linear combination of the regression parameters in a linear regression model with iid normal errors with known variance when there is uncertain prior information that a distinct specified linear combination of the regression parameters takes a given value. This confidence interval, found by numerical constrained optimization, has the required minimum coverage and utilizes this uncertain prior information through desirable expected length properties.
This confidence interval has the following three practical applications.
Firstly, if the error variance has been accurately estimated from previous data then it may be treated as being effectively known. Secondly, for sufficiently large (dimension of the response vector) minus (dimension of regression parameter vector), greater than or equal to 30 (say),
if we replace the assumed known value of the error variance by its usual estimator in the formula for the confidence interval then the resulting interval has, to a very good approximation, the same coverage probability and expected length properties as when the error variance is known. Thirdly, some more complicated models can be approximated by the linear regression model with error variance known when certain unknown parameters are replaced by estimates. This confidence interval is described in Kabaila, P. and Mainzer, R. (2017) <arXiv:1708.09543>, and is a member of the family of confidence intervals proposed by Kabaila, P. and Giri, K. (2009) <doi:10.1016/j.jspi.2009.03.018>.

License GPL-2

Encoding UTF-8

LazyData true

Imports nloptr, statmod, functional, pracma, stats

RoxygenNote 6.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Rheanna Mainzer [aut, cre],
Paul Kabaila [aut]

Maintainer Rheanna Mainzer <rheanna.mainzer@unimelb.edu.au>

Repository CRAN

Date/Publication 2019-01-15 05:20:03 UTC

R topics documented:

`bsciuupi` *Compute the vector (b(1),...,b(5),s(0),...,s(5)) that specifies the CIUUPI*

Description

Compute the vector (b(1),...,b(5),s(0),...,s(5)) that specifies the confidence interval that utilizes uncertain prior information (CIUUPI).

Usage

```
bsciuupi(alpha, natural = 1, rho = NULL, a, c, x)
```

Arguments

`alpha`	The minimum coverage probability is 1 - alpha
`natural`	Equal to 1 (default) if the functions b and s are found by natural cubic spline interpolation or 0 if these functions are found by clamped cubic spline interpolation in the interval [-6,6]
`rho`	A known correlation
`a`	A vector used to specify the parameter of interest
`c`	A vector used to specify the parameter about which we have uncertain prior information
`x`	The n by p design matrix

Details

Suppose that
$$y = X\beta + \epsilon$$
where y is a random n-vector of responses, X is a known n by p matrix with linearly independent columns, β is an unknown parameter p-vector and ϵ is the random error with components that are iid normally distributed with zero mean and known variance. The parameter of interest is $\theta =$ a' β. The uncertain prior information is that $\tau =$ c' β - t = 0, where a and c are specified linearly independent vectors and t is a specified number. rho is the known correlation between the least squares estimators of θ and τ. The user must specify either a, c and x or rho. If a, c and x are specified then rho is computed.

The confidence interval for θ, with minimum coverage probability 1 - alpha, that utilizes the uncertain prior information that $\tau = 0$ belongs to a class of confidence intervals indexed by the functions b and s. The function b is an odd continuous function and the function s is an even continuous function. In addition, b(x)=0 and s(x) is equal to the $1 - \alpha/2$ quantile of the standard normal distribution for all |x| greater than or equal to 6. The values of these functions in the interval $[-6, 6]$ are specified by $b(1), b(2), \ldots, b(5)$ and $s(0), s(1), \ldots, s(5)$ as follows. By assumption, $b(0) = 0$ and $b(-i) = -b(i)$ and $s(-i) = s(i)$ for $i = 1, ..., 6$. The values of $b(x)$ and $s(x)$ for any x in the interval $[-6, 6]$ are found using cube spline interpolation for the given values of $b(i)$ and $s(i)$ for $i = -6, -5, ..., 0, 1, ..., 5, 6$.

The vector $(b(1), b(2), \ldots, b(5), s(0), s(1), \ldots, s(5))$ is found by numerical constrained optimization so that the confidence interval has minimum coverage probability 1 - alpha and utilizes the uncertain prior information through its desirable expected length properties. The optimization is performed using the slsqp function in the nloptr package.

Value

The vector $(b(1), b(2), \ldots, b(5), s(0), s(1), \ldots, s(5))$ that specifies the CIUUPI.

See Also

ciuupi

Examples

```
# Compute the vector (b(1),...,b(5),s(0),...,s(5)) that specifies the CIUUPI,
# for given alpha and rho: (may take a few minutes to run)

bsvec <- bsciuupi(0.05, rho = 0.4)


# The result (to 7 decimal places) is
bsvec <- c(0.129443483, 0.218926703, 0.125880945, 0.024672734, -0.001427343,
           1.792489585, 1.893870240, 2.081786492, 2.080407355,  1.986667246,
           1.958594824)
bsvec

# Compute the vector (b(1),...,b(5),s(0),...,s(5)) that specifies the CIUUPI,
# for given alpha, a, c and x
x1 <- c(-1, 1, -1, 1)
```

```
x2 <- c(-1, -1, 1, 1)
x <- cbind(rep(1, 4), x1, x2, x1*x2)
a <- c(0, 2, 0, -2)
c <- c(0, 0, 0, 1)

# The following may take a few minutes to run:

bsvec2 <- bsciuupi(0.05, a = a, c = c, x = x)


# The result (to 7 decimal places) is
bsvec2 <- c(-0.03639701, -0.18051953, -0.25111411, -0.15830362, -0.04479113,
            1.71997203, 1.79147968, 2.03881195, 2.19926399, 2.11845381,
            2.00482563)
bsvec2
```

---

bsspline *Evaluate the functions b and s at x*

Description

Evaluate the functions b and s, as specified by (b(1),b(2),...,b(5),s(0),s(1),...,s(5)), `alpha` and `natural`, at `x`.

Usage

```
bsspline(x, bsvec, alpha, natural = 1)
```

Arguments

`x`	A value or vector of values at which the functions b and s are to be evaluated
`bsvec`	The vector (b(1),b(2),...,b(5),s(0),s(1),...,s(5))
`alpha`	The minimum coverage probability is 1 - `alpha`
`natural`	Equal to 1 (default) for natural cubic spline interpolation or 0 for clamped cubic spline interpolation

Details

The function b is an odd continuous function and the function s is an even continuous function. In addition, b(x)=0 and s(x) is equal to the $1 - \alpha/2$ quantile of the standard normal distribution for all |x| greater than or equal to 6. The values of these functions in the interval $[-6, 6]$ are specified by the vector $(b(1), b(2), \ldots, b(5), s(0), s(1), \ldots, s(5))$ as follows. By assumption, $b(0) = 0$ and $b(-i) = -b(i)$ and $s(-i) = s(i)$ for $i = 1, ..., 6$. The values of $b(x)$ and $s(x)$ for any x in the interval $[-6, 6]$ are found using cube spline interpolation for the given values of $b(i)$ and $s(i)$ for $i = -6, -5, ..., 0, 1, ..., 5, 6$.

The vector (b(1),b(2),...,b(5),s(0),s(1),...,s(5)) that specifies the confidence interval that utilizes uncertain prior information (CIUUPI) is obtained using `bsciuupi`.

Value

A data frame containing x and the corresponding values of the functions b and s.

See Also

bsciuupi, ciuupi

Examples

```
alpha <- 0.05

# Find the vector (b(1),b(2),...,b(5),s(0),s(1),...,s(5)) that specifies the
# CIUUPI: (this may take a few minutes to run)

bsvec <- bsciuupi(alpha, rho = 0.4)


# The result (to 7 decimal places) is
bsvec <- c(0.129443483, 0.218926703, 0.125880945, 0.024672734, -0.001427343,
           1.792489585, 1.893870240, 2.081786492, 2.080407355,  1.986667246,
           1.958594824)

# Graph the functions b and s
x <- seq(0, 8, by = 0.1)
xseq <- seq(0, 6, by = 1)
bvec <- c(0, bsvec[1:5], 0)
quantile <- qnorm(1-(alpha)/2, 0, 1)
svec <- c(bsvec[6:11], quantile)
splineval <- bsspline(x, bsvec, alpha)

plot(x, splineval[, 2], type = "l", main = "b function",
ylab = " ", las = 1, lwd = 2, xaxs = "i", col = "blue")
points(xseq, bvec, pch = 19, col = "blue")
plot(x, splineval[, 3], type = "l", main = "s function",
ylab = " ", las = 1, lwd = 2, xaxs = "i",  col = "blue")
points(xseq, svec, pch = 19, col = "blue")
```

---

cistandard *Compute the standard confidence interval*

Description

Compute the standard confidence interval

Usage

```
cistandard(a, x, y, alpha, sig = NULL)
```

Arguments

a	A vector used to specify the parameter of interest
x	A known n by p matrix
y	A known n-vector of responses
alpha	1 - alpha is the nominal coverage probability of the confidence interval
sig	Standard deviation of the random error. If a value is not specified, sig will be estimated from the data.

Details

Suppose that

$$Y = X\beta + \epsilon$$

is a random n-vector of responses, X is a known n by p matrix with linearly independent columns, β is an unknown parameter p-vector and ϵ has a multivariate normal distribution with mean vector 0 and variance sig^2 times the n by n identity matrix. Then cistandard will compute the standard confidence interval for a' β.

In the example below we use the data set described in Table 7.5 of Box et al. (1963). A description of the parameter of interest is given in Dicsussion 5.8, p.3426 of Kabaila and Giri (2009).

Value

The standard confidence interval

References

Box, G.E.P., Connor, L.R., Cousins, W.R., Davies, O.L., Hinsworth, F.R., Sillitto, G.P. (1963) The Design and Analysis of Industrial Experiments, 2nd edition, reprinted. Oliver and Boyd, London.

Kabaila, P. and Giri, K. (2009) Confidence intervals in regression utilizing prior information. Journal of Statistical Planning and Inference, 139, 3419 - 3429.

Examples

```
y <- c(87.2, 88.4, 86.7, 89.2)
x1 <- c(-1, 1, -1, 1)
x2 <- c(-1, -1, 1, 1)
x <- cbind(rep(1, 4), x1, x2, x1*x2)
a <- c(0, 2, 0, -2)

# Calculate the standard 95% confidence interval when sigma = 0.8
res <- cistandard(a, x, y, 0.05, sig = 0.8)
res
```

`ciuupi`	*Compute the confidence interval that utilizes the uncertain prior information*

Description

Compute the confidence interval that utilizes the uncertain prior information

Usage

```
ciuupi(alpha, a, c, x, bsvec, t, y, natural = 1, sig = NULL)
```

Arguments

`alpha`	1 - `alpha` is the minimum coverage probability of the confidence interval
`a`	A vector used to specify the parameter of interest
`c`	A vector used to specify the parameter about which we have uncertain prior information
`x`	The n by p design matrix
`bsvec`	The vector (b(1),...,b(5),s(0),...,s(5)) that specifies the CIUUPI
`t`	A number used to specify the parameter about which we have uncertain prior information
`y`	The n-vector of observed responses
`natural`	Equal to 1 (default) if b and s functions are obtained by natural cubic spline interpolation or 0 if obtained by clamped cubic spline interpolation
`sig`	Standard deviation of the random error. If a value is not specified then `sig` is estimated from the data.

Details

Suppose that

$$y = X\beta + \epsilon$$

where y is a random n-vector of responses, X is a known n by p matrix with linearly independent columns, β is an unknown parameter p-vector and ϵ with components that are iid normally distributed with zero mean and known variance. Then `ciuupi` will compute a confidence interval for $\theta =$`a`' β that utilizes the uncertain prior information that `c`' β - `t` = 0, where `a` and `c` are specified linearly independent vectors and `t` is a specified number.

In the example below we use the data set described in Table 7.5 of Box et al. (1963). A description of the parameter of interest and the parameter about which we have uncertain prior information is given in Dicsussion 5.8, p.3426 of Kabaila and Giri (2009).

Value

The confidence interval that utilizes uncertain prior information

References

Box, G.E.P., Connor, L.R., Cousins, W.R., Davies, O.L., Hinsworth, F.R., Sillitto, G.P. (1963) The Design and Analysis of Industrial Experiments, 2nd edition, reprinted. Oliver and Boyd, London.

Kabaila, P. and Giri, K. (2009) Confidence intervals in regression utilizing prior information. Journal of Statistical Planning and Inference, 139, 3419 - 3429.

Examples

```
# Specify alpha, a, c, x
alpha <- 0.05
a <- c(0, 2, 0, -2)
c <- c(0, 0, 0, 1)
x1 <- c(-1, 1, -1, 1)
x2 <- c(-1, -1, 1, 1)
x <- cbind(rep(1, 4), x1, x2, x1*x2)

# Find the vector (b(1),b(2),...,b(5),s(0),s(1),...,s(5)) that specifies the
# CIUUPI: (this may take a few minutes to run)

bsvec <- bsciuupi(alpha, a = a, c = c, x = x)


# The result (to 7 decimal places) is
bsvec <- c(-0.03639701, -0.18051953, -0.25111411, -0.15830362, -0.04479113,
           1.71997203, 1.79147968, 2.03881195, 2.19926399, 2.11845381,
           2.00482563)

# Specify t and y
t <- 0
y <- c(87.2, 88.4, 86.7, 89.2)

# Find the CIUUPI
res <- ciuupi(alpha, a, c, x, bsvec, t, y, natural = 1, sig = 0.8)
res
```

`cpciuupi` *Compute the coverage probability of the CIUUPI*

Description

Evaluate the coverage probability of the confidence interval that utilizes uncertain prior information (CIUUPI) at `gam`.

Usage

```
cpciuupi(gam, bsvec, alpha, natural = 1, rho = NULL, a, c, x)
```

Arguments

`gam`	A value of gamma or vector of gamma values at which the coverage probability function is evaluated
`bsvec`	The vector (b(1),...,b(5),s(0),...,s(5)) that specifies the CIUUPI
`alpha`	The nominal coverage probability is 1 - `alpha`
`natural`	Equal to 1 (default) if the b and s functions are obtained by natural cubic spline interpolation or 0 if obtained by clamped cubic spline interpolation
`rho`	A known correlation
`a`	A vector used to specify the parameter of interest
`c`	A vector used to specify the parameter about which we have uncertain prior information
`x`	The n by p design matrix

Details

Suppose that

$$y = X\beta + \epsilon$$

where y is a random n-vector of responses, X is a known n by p matrix with linearly independent columns, β is an unknown parameter p-vector and ϵ is the random error with components that are iid normally distributed with zero mean and known variance. The parameter of interest is $\theta =$ `a`' β. The uncertain prior information is that $\tau =$ `c`' β - `t` = 0, where `a` and `c` are specified linearly independent vectors and `t` is a specified number. `rho` is the known correlation between the least squares estimators of θ and τ. The user must specify either `a`, `c` and `x` or `rho`. If `a`, `c` and `x` are specified then `rho` is computed.

The CIUUPI is specified by the vector (b(1),...,b(5),s(0),...,s(5)), `alpha` and `natural`

Value

The value(s) of the coverage probability of the CIUUPI at `gam`.

See Also

`ciuupi`, `bsciuupi`

Examples

```
alpha <- 0.05

# Find the vector (b(1),b(2),...,b(5),s(0),s(1),...,s(5)) that specifies the
# CIUUPI: (this may take a few minutes to run)

bsvec <- bsciuupi(alpha, rho = 0.4)


# The result (to 7 decimal places) is
bsvec <- c(0.129443483, 0.218926703, 0.125880945, 0.024672734, -0.001427343,
           1.792489585, 1.893870240, 2.081786492, 2.080407355,  1.986667246,
```

```
            1.958594824)

# Graph the coverage probability function
gam <- seq(0, 10, by = 0.1)
cp <- cpciuupi(gam, bsvec, alpha, rho = 0.4)
plot(gam, cp, type = "l", lwd = 2, ylab = "", las = 1, xaxs = "i",
main = "Coverage Probability", col = "blue",
xlab = expression(paste("|", gamma, "|")), ylim = c(0.94999, 0.95001))
abline(h = 1-alpha, lty = 2)
```

---

selciuupi *Compute the scaled expected length of the CIUUPI*

Description

Evaluate the scaled expected length of the confidence interval that utilizes uncertain prior information (CIUUPI) at gam.

Usage

```
selciuupi(gam, bsvec, alpha, natural = 1, rho = NULL, a, c, x)
```

Arguments

gam	A value of gamma or vector of gamma values at which the scaled expected length function is evaluated
bsvec	The vector (b(1),...,b(5),s(0),...,s(5)) that specifies the CIUUPI
alpha	The minimum coverage probability is 1 - alpha
natural	Equal to 1 (default) if the functions b and s are obtained by natural cubic spline interpolation or 0 if obtained by clamped cubic spline interpolation
rho	A known correlation
a	A vector used to specify the parameter of interest
c	A vector used to specify the parameter about which we have uncertain prior information
x	The n by p design matrix

Details

Suppose that

$$y = X\beta + \epsilon$$

where y is a random n-vector of responses, X is a known n by p matrix with linearly independent columns, β is an unknown parameter p-vector and ϵ is the random error with components that are iid normally distributed with zero mean and known variance. The parameter of interest is θ = a' β. The uncertain prior information is that τ = c' β - t = 0, where a and c are specified linearly independent vectors and t is a specified number. rho is the known correlation between the least squares estimators of θ and τ. The user must specify either a, c and x or rho. If a, c and x are specified then rho is computed.

The CIUUPI is specified by the vector (b(1),...,b(5),s(0),...,s(5)), alpha and natural

The scaled expected length is defined as the expected length of the CIUUPI divided by the expected length of the standard confidence interval with the same minimum coverage probability.

Value

The value(s) of the scaled expected length at gam.

See Also

ciuupi, bsciuupi

Examples

```
alpha <- 0.05

# Find the vector (b(1),b(2),...,b(5),s(0),s(1),...,s(5)) that specifies the
# CIUUPI: (this may take a few minutes to run)

bsvec <- bsciuupi(alpha, rho = 0.4)


# The result (to 7 decimal places) is
bsvec <- c(0.129443483, 0.218926703, 0.125880945, 0.024672734, -0.001427343,
           1.792489585, 1.893870240, 2.081786492, 2.080407355,  1.986667246,
           1.958594824)

# Graph the scaled expected length function
gam <- seq(0, 8, by = 0.1)
sel <- selciuupi(gam, bsvec, alpha, rho = 0.4)
plot(gam, sel, type = "l", lwd = 2, ylab = "", las = 1, xaxs = "i",
main = "Scaled Expected Length", col = "blue",
xlab = expression(paste("|", gamma, "|")))
abline(h = 1, lty = 2)
```

Index