Package ‘rQCC’

July 5, 2020

Title Robust Quality Control Chart

Version 1.20.7

Date 2020-07-05

Author Chanseok Park [aut, cre] (<https://orcid.org/0000-0002-2208-3498>),
Min Wang [ctb] (<https://orcid.org/0000-0002-9233-7844>)

Maintainer Chanseok Park <statpnu@gmail.com>

Depends R (>= 3.2.3)

Description Constructs robust quality control chart based on the median or Hodges-Lehmann estimator (location) and the median absolute deviation (MAD) or Shamos estimator (scale). These estimators are all unbiased with a sample of finite size. For more details, see Park, Kim and Wang (2020) <doi:10.1080/03610918.2019.1699114>. In addition, using this package, the conventional quality control charts such as X-bar, S and R charts are also easily constructed. This work was partially supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (NRF-2017R1A2B4004169).

License GPL-2 | GPL-3

URL https://AppliedStat.GitHub.io/R/

BugReports https://github.com/AppliedStat/R/issues

Encoding UTF-8

LazyData yes

NeedsCompilation no

Repository CRAN

Date/Publication 2020-07-05 04:40:02 UTC

R topics documented:

`empirical.variance` *Empirical variances of robust estimators*

Description

Provides the empirical variances of the median, Hodges-Lehmann (HL1, HL2, HL3) median absolute deviation (MAD) and Shamos estimators.

Usage

```
evar (n, method = c("median", "HL1", "HL2", "HL3", "mad", "shamos"))
```

Arguments

`n`	sample size ($n \geq 1$).
`method`	a character string specifying the estimator, must be one of `"median"` (default), `"HL1"`, `"HL2"`, `"HL3"`, `"mad"`, and `"shamos"`.

Details

The empirical variances for $n = 1, 2, \ldots, 100$ are obtained using the extensive Monte Carlo simulation with 1E07 replicates. For the case of $n > 100$, they are obtained using the method of Hayes (2014).

`evar` computes the empirical variance of a specific estimator (one of `"median"`, `"HL1"`, `"HL2"`, `"HL3"`, `"mad"`, and `"shamos"`) when a random sample is from the standard normal distribution.

Value

It returns a numeric value.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Hayes, K. (2014). Finite-sample bias-correction factors for the median absolute deviation. *Communications in Statistics: Simulation and Computation*, **43**, 2205–2212.

See Also

`RE`{rQCC} for the relative efficiency.
`n.times.eVar.of.HL1`{rQCC} for the empirical variance of the HL1 estimator (times n).
`n.times.eVar.of.HL2`{rQCC} for the empirical variance of the HL2 estimator (times n).
`n.times.eVar.of.HL3`{rQCC} for the empirical variance of the HL3 estimator (times n).
`n.times.eVar.of.mad`{rQCC} for the empirical variance of the MAD estimator (times n).
`n.times.eVar.of.median`{rQCC} for the empirical variance of the median estimator (times n).
`n.times.eVar.of.shamos`{rQCC} for the empirical variance of the Shamos estimator (times n).

Examples

```
# Empirical variance of the Hodges-Lehmann estimator (HL2) under the standard normal distribution.
evar (n=10, method="HL2")
```

`factors.for.control.chart` *Factors for constructing control charts*

Description

Factors for constructing control charts.

Usage

```
factors.cc(n, factor=c( "A","A1","A2","A3", "B1","B2","B3","B4","B5","B6", "c2","c4",
                "d2","d3","D1","D2","D3","D4", "E1", "E2", "E3"), sigma.factor=3)
```

Arguments

`n`	sample size ($n \geq 1$).
`factor`	a character string specifying the foctor. `"c2"`, `"c4"`, `"d2"`, `"d3"` for control chart lines; `"A"`, `"A1"`, `"A2"`, `"A3"` for averages for computing control limits; `"B1"`, `"B2"`, `"B3"`, `"B4"`, `"B5"`, `"B6"` for standard devations; `"D1"`, `"D2"`, `"D3"`, `"D4"` for ranges; and `"E1"`, `"E2"`, `"E3"` for individuals.
`sigma.factor`	a factor for the standard deviation (σ). For example, the American Standard uses "*3*sigma*" limits (0.27% false alarm rate), while the British Standard uses "*3.09*sigma*" limits (0.20% false alarm rate).

Details

The values of the factors are used for constructing various control charts.

For example, the conventional $\bar{X}$ chart with the sample standard deviation is given by

$$\bar{\bar{X}} \pm A_3 \bar{S}.$$

For more details, refer to `vignette("factors.cc",package="rQCC")`.

Value

It returns a numeric value.

Author(s)

Chanseok Park

References

ASTM (2010). *Manual on Presentation of Data and Control Chart Analysis* (STP 15-D), 8th edition. American Society for Testing and Materials, West Conshohocken, PA.

ASTM (1951). *Manual on Quality Control of Materials* (STP 15-C), American Society for Testing and Materials, Philadelphia, PA.

See Also

`c4.factor`{rQCC} for c_4 factor for the finite-sample unbiasing factor to estimate the standard deviation (σ) under the normal distribution using various estimators such as the sample standard deviation, the sample range, the median absolute deviation (MAD), and the Shamos estimate.

Examples

```
## A3 is used for constructing the conventional X-bar chart
#  with the sample standard deviation.
factors.cc(n=10, factor="A3")


## Unbiasing factor for the standard deviation
#           using the sample standard deviation.
factors.cc(n=10, factor="c4")
#  The above is the same as below:
c4.factor(n=10, method="sd")


## Unbiasing factor for the standard deviation
#           using the sample range.
factors.cc(n=10, factor="d2")
#  The above is the same as below:
c4.factor(n=10, method="range")
```

```
## Table B2 in Supplement B of ASTM (1951).
char = c("A","A1","A2","c2", "B1","B2","B3","B4", "d2","d3","D1","D2","D3","D4")
nn = 2L:25L

res=NULL
for(n in nn){tmp=NULL;for(ch in char) tmp=c(tmp,factors.cc(n,ch));res=rbind(res,tmp)}
rownames(res) = paste0("n=",nn)
round(res,4)


## Table 49 in Chapter 3 of ASTM (2010).
char = c("A","A2","A3","c4", "B3","B4","B5","B6", "d2","d3","D1","D2","D3","D4")
nn = 2L:25L

res=NULL
for(n in nn){tmp=NULL;for(ch in char) tmp=c(tmp,factors.cc(n,ch));res=rbind(res,tmp)}
rownames(res) = paste0("n=",nn)
round(res,4)


## Table 50 in Chapter 3 of ASTM (2010).
char = c("E2", "E3")
nn = 2L:25L

res=NULL
for(n in nn){tmp=NULL;for(ch in char) tmp=c(tmp,factors.cc(n,ch));res=rbind(res,tmp)}
rownames(res) = paste0("n=",nn)
round(res,3)


## vignette ##
if (interactive()) vignette("factors.cc", package="rQCC")
```

`finite.sample.breakdown` *Finite-sample breakdown point*

Description

Calculates the finite-sample breakdown point of the mean, median, Hodges-Lehmann estimators (HL1, HL2, HL3), standard deviation, range, MAD (median absolute deviation) and Shamos estimators. Note that for the case of the mean, standard deviation and range, the finite-sample breakdown points are always zero.

Usage

```
finite.breakdown(n,
      method=c("mean","median","HL1","HL2","HL3","sd","range","mad","shamos") )
```

Arguments

n	sample size ($n \geq 1$).
method	a character string specifying the estimator, must be one of "mean" (default), "median", "HL1", "HL2", "HL3", "sd", "range", "mad", and "shamos".

Details

`finite.breakdown` gives the finite-sample breakdown point of the specified estimator.

The Hodges-Lehmann (HL1) is defined as

$$\mathrm{HL1} = \operatorname*{median}_{i<j}\left(\frac{X_i + X_j}{2}\right)$$

where $i, j = 1, 2, \ldots, n$.

The Hodges-Lehmann (HL2) is defined as

$$\mathrm{HL2} = \operatorname*{median}_{i\leq j}\left(\frac{X_i + X_j}{2}\right).$$

The Hodges-Lehmann (HL3) is defined as

$$\mathrm{HL3} = \operatorname*{median}_{\forall(i,j)}\left(\frac{X_i + X_j}{2}\right).$$

Value

It returns a numeric value.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Hodges, Jr., J. L. (1967). Efficiency in normal samples and tolerance of extreme values for some estimates of location. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. **1**, 163–186. University of California Press, Berkeley.

Hampel, F. R., Ronchetti, E., Rousseeuw, P. J., and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*, Subsection 2.2a. John Wiley & Sons, New York.

See Also

`HL`{rQCC} for the Hodges-Lehmann estimate.

Examples

```
# finite-sample breakdown point of the Hodges-Lehmann (HL1) with size n=10.
finite.breakdown(n=10, method="HL2")
```

Hodges-Lehmann *Hodges-Lehmann estimate*

Description

Calculates the Hodges-Lehmann estimate.

Usage

```
HL(x, method = c("HL1", "HL2", "HL3"), na.rm = FALSE)
```

Arguments

x	a numeric vector of observations.
method	a character string specifying the estimator, must be one of "HL1" (default), "HL2" and "HL3".
na.rm	a logical value indicating whether NA values should be stripped before the computation proceeds.

Details

HL computes the Hodges-Lehmann estimates (one of "HL1", "HL2", "HL3").

The Hodges-Lehmann (HL1) is defined as

$$\mathrm{HL1} = \underset{i<j}{\mathrm{median}} \left(\frac{X_i + X_j}{2}\right)$$

where $i, j = 1, 2, \ldots, n$.

The Hodges-Lehmann (HL2) is defined as

$$\mathrm{HL2} = \underset{i\le j}{\mathrm{median}} \left(\frac{X_i + X_j}{2}\right).$$

The Hodges-Lehmann (HL3) is defined as

$$\mathrm{HL3} = \underset{\forall(i,j)}{\mathrm{median}} \left(\frac{X_i + X_j}{2}\right).$$

Value

It returns a numeric value.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Hodges, J. L. and E. L. Lehmann (1963). Estimates of location based on rank tests. *Annals of Mathematical Statistics*, **34**, 598–611.

See Also

mean{base} for calculating sample mean and median{stats} for calculating sample median.

`finite.breakdown`{rQCC} for calculating the finite-sample breakdown point.

Examples

```
x = c(0:10, 50)
HL(x, method="HL2")
```

`MAD` *Median absolute deviation (MAD)*

Description

Calculates the unbiased median absolute deviation (MAD) estimate and the unbiased squared MAD under the normal distribution which are adjusted by the Fisher-consistency and finite-sample unbiasing factors.

Usage

```
mad.unbiased(x, center = median(x), constant=1.4826, na.rm = FALSE)

mad2.unbiased(x, center = median(x), constant=1.4826, na.rm = FALSE)
```

Arguments

`x`	a numeric vector of observations.
`center`	pptionally, the center: defaults to the median.
`constant`	correction factor for the Fisher-consistency under the normal distribution
`na.rm`	a logical value indicating whether NA values should be stripped before the computation proceeds.

Details

The unbiased MAD (`mad.unbiased`) is defined as the mad{stats} divided by $c_5(n)$, where $c_5(n)$ is the finite-sample unbiasing factor. Note that $c_5(n)$ notation is used in Park et. al (2020), and $c_5(n)$ is calculated using the function `c4.factor`{rQCC} with `method="mad"` option. The default value (`constant=1.4826`) ensures the Fisher-consistency under the normal distribution. Note that the original MAD was proposed by Hampel (1974).

The unbiased squared MAD (`mad2.unbiased`) is defined as the squared mad{stats} divided by $w_5(n)$ where $w_5(n)$ is the finite-sample unbiasing factor. Note that $w_5(n)$ notation is used in Park et. al (2020), and $w_5(n)$ is calculated using the function `w4.factor`{rQCC} with `method="mad2"` option. The default value (`constant=1.4826`) ensures the Fisher-consistency under the normal distribution. Note that the square of the conventional MAD is Fisher-consistent for the variance (σ^2) under the normal distribution, but it is not unbiased with a sample of finite size.

Value

They return a numeric value.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, **69**, 383–393.

See Also

`c4.factor`{rQCC} for finite-sample unbiasing factor for the standard deviation under the normal distribution.

`w4.factor`{rQCC} for finite-sample unbiasing factor for the variance under the normal distribution.

`shamos`{rQCC} for robust Fisher-consistent estimator of the standard deviation under the normal distribution.

`shamos.unbiased`{rQCC} for robust finite-sample unbiased estimator of the standard deviation under the normal distribution.

mad{stats} for calculating the sample MAD.

`finite.breakdown`{rQCC} for calculating the finite-sample breakdown point.

Examples

```
x = c(0:10, 50)

# Fisher-consistent MAD, but not unbiased with a finite sample.
mad(x)
```

```
# Unbiased MAD.
mad.unbiased(x)

# Fisher-consistent squared MAD, but not unbiased.
mad(x)^2

# Unbiased squared MAD.
mad2.unbiased(x)
```

`n.times.eBias` *Empirical biases (times n)*

Description

n times the empirical biases of the median absolute deviation (MAD) and Shamos estimators under the standard normal distribution, where n is the sample size and n is from 1 to 100. For the MAD estimator, mad{stats} is used. For the Shamos estimator, the Fisher-consistent Shamos estimator, `shamos`{rQCC}, is used.

These estimators are not unbiased with a finite sample. The empirical biases are obtained using the extensive Monte Carlo simulation with 1E07 replicates.

Usage

```
n.times.eBias.of.mad

n.times.eBias.of.shamos
```

Value

They return a vector of 100 values.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Shamos, M. I. (1976). Geometry and statistics: Problems at the interface. In Traub, J. F., editor, *Algorithms and Complexity: New Directions and Recent Results*, pages 251–280. Academic Press, New York.

Lèvy-Leduc, C., Boistard, H., Moulines, E., Taqqu, M. S., and Reisen, V. A. (2011). Large sample behaviour of some well-known robust estimators under long-range dependence. *Statistics*, **45**, 59–71.

`n.times.eVar` *Empirical variances (times n)*

Description

n times the empirical variances of the Hodges-Lehmann (HL1, HL2, HL3), the median, the median absolute deviation (MAD), and the Shamos estimators under the standard normal distribution, where n is the sample size and n is from 1 to 100.

For the MAD estimates, mad{stats} is used. For the Hodges-Lehmann, `HL`{rQCC} is used. For the Shamos, the Fisher-consistent Shamos estimator, `shamos`{rQCC}, is used.

The empirical variances are obtained using the Monte Carlo simulation with 1E07 replicates.

Usage

```
n.times.eVar.of.HL1

n.times.eVar.of.HL2

n.times.eVar.of.HL3

n.times.eVar.of.median

n.times.eVar.of.mad

n.times.eVar.of.shamos
```

Value

They return a vector of 100 values.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Hodges, J. L. and E. L. Lehmann (1963). Estimates of location based on rank tests. *Annals of Mathematical Statistics*, **34**, 598–611.

Shamos, M. I. (1976). Geometry and statistics: Problems at the interface. In Traub, J. F., editor, *Algorithms and Complexity: New Directions and Recent Results*, pages 251–280. Academic Press, New York.

Lèvy-Leduc, C., Boistard, H., Moulines, E., Taqqu, M. S., and Reisen, V. A. (2011). Large sample behaviour of some well-known robust estimators under long-range dependence. *Statistics*, **45**, 59–71.

`plot.rcc` *Plot robust quality control chart*

Description

Plot method for an object of class "rcc".

Usage

```
## S3 method for class 'rcc'
plot(x, digits=getOption("digits")-2, cex.text=0.7, x.text=1, ...)
```

Arguments

x	an object of class "rcc".
digits	number of significant digits to use, see print.
cex.text	magnification to be used for the text labels (LCL, CL, UCL).
x.text	x-coordinate where the text labels (LCL, CL, UCL) should be written.
...	additional parameters to plot.

Author(s)

Chanseok Park

See Also

rcc{rQCC}, plot

`relative.efficiency` *Relative efficiency (RE)*

Description

Calculates the relative efficiency value of the mean, median and Hodges-Lehmann (HL1, HL2, HL3) estimators with respect to the sample mean and that of the standard deviation, range, median absolute deviation (MAD) and Shamos estimators with respect to the sample standard deviation. Note that the relative efficiency value of the mean and the standard deviation is always one by definition.

For the case of the sample mean, standard deviation and range, it is possible to derive their variances in analytic form, but, for the other case, it may be impossible. In this case, the variances with $n = 1, 2, \ldots, 100$ are obtained using the extensive Monte Carlo simulation with 1E07 replicates. For $n > 100$, the variances are approximated based on the method of Hayes (2014).

To obtain the relative efficiency value of the unbiased scale estimators, use `correction=TRUE` option. Note that the location estimators ("mean", "median", "HL1", "HL2", "HL3") are unbiased.

Usage

```
RE (n, method = c("mean", "median","HL1","HL2","HL3", "sd", "range","mad","shamos"),
       correction = FALSE)
```

Arguments

n	sample size ($n \geq 1$).
correction	a finite-sample bias correction. TRUE adjusts a finite-sample bias correction using A3.factor function.
method	a character string specifying the estimator, must be one of "mean" (default), "median", "HL1", "HL2", "HL3", "range", "mad", and "shamos".

Details

RE calculates the relative efficiency value of a location estimator ("median", "HL1", "HL2", "HL3") with respect to the sample mean and that of a scale estimator ("range", "mad", "shamos") with respect to the sample standard deviation.

Note that the relative efficiency (RE) of $\hat{\theta}_2$ with respect to $\hat{\theta}_1$ is defined as

$$\mathrm{RE}(\hat{\theta}_2|\hat{\theta}_1) = \frac{\mathrm{Var}(\hat{\theta}_1)}{\mathrm{Var}(\hat{\theta}_2)}.$$

Value

It returns a numeric value.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Hayes, K. (2014). Finite-sample bias-correction factors for the median absolute deviation. *Communications in Statistics: Simulation and Computation*, **43**, 2205–2212.

See Also

n.times.eVar.of.HL1{rQCC} for the empirical variance of the HL1 estimator (times n).
n.times.eVar.of.HL2{rQCC} for the empirical variance of the HL2 estimator (times n).
n.times.eVar.of.HL3{rQCC} for the empirical variance of the HL3 estimator (times n).
n.times.eVar.of.mad{rQCC} for the empirical variance of the MAD estimator (times n).
n.times.eVar.of.median{rQCC} for the empirical variance of the median estimator (times n).
n.times.eVar.of.shamos{rQCC} for the empirical variance of the Shamos estimator (times n).

Examples

```
# The RE of the Hodges-Lehmann (HL2) estimator
#  with respect to the sample standard deviation under the normal distribution.
RE(n=5, method="HL2")

# RE of the Shamos estimator
# with respect to the sample standard deviation under the normal distribution.
RE(n=5, method="shamos")

# RE of the unbiased Shamos estimator
# with respect to the unbiased sample standard deviation under the normal distribution.
RE(n=5, method="shamos", correction=TRUE)

# RE of the range (maximum minus minimum)
# with respect to the sample standard deviation under the normal distribution.
RE(n=6, method="range")

# RE of the unbiased range ( (maximum - minimum) / d2 )
# with respect to the unibased sample standard deviation under the normal distribution.
RE(n=6, method="range", correction=TRUE)
```

`robust.control.chart` *Robust quality control chart*

Description

Constructs a robust control chart.

Usage

```
rcc (x, location = c("mean","median","HL1","HL2","HL3"),
    scale = c("sd", "range","mad","shamos"), type = c("Xbar","S","R"), sigma.factor = 3)
```

Arguments

`x`	a numeric vector of observations.
`location`	a character string specifying the location estimator, must be one of `"mean"` (default), `"median"`, `"HL1"`, `"HL2"` and `"HL3"`.
`scale`	a character string specifying the scale estimator, must be one of `"sd"` (default), `"range"`, `"mad"` and `"shamos"`.
`type`	a character string specifying the type of control chart.
`sigma.factor`	a factor for the standard deviation (σ). For example, the American Standard uses "*3*sigma*" limits (0.27% false alarm rate), while the British Standard uses "*3.09*sigma*" limits (0.20% false alarm rate).

Details

`rcc` constructs a robust X-bar control chart. Using various robust location and scale estimates, one can construct a robust X-bar chart. The location and scale estimates used in this fuction are all unbiased. In addition, one can also construct the conventional $\bar{X}$ chart with S and $\bar{X}$ chart with R. For more details, see the vignette.

Value

`rcc` returns an object of class "rcc". The function `summary` is used to obtain and print a summary of the results and the function `plot` is used to plot the control chart.

Author(s)

Chanseok Park

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

ASTM (2010). *Manual on Presentation of Data and Control Chart Analysis* (STP 15-D), 8th edition. American Society for Testing and Materials, West Conshohocken, PA.

Ryan (2000). *Statistical Methods For Quality Improvement*, 2nd edition. John Wiley & Sons, New York, NY.

Examples

```
##############
# X-bar chart #
##############

# ==========
# Example 1a
# ----------
# The conventional X-bar chart with the standard deviation.
# Refer to Example 3 in Section 3.31 of ASTM (2010).

# The data below are from Table 29 in Section 3.31 of ASTM (2010).
x1 = c(0.5005, 0.5000, 0.5008, 0.5000, 0.5005, 0.5000)
x2 = c(0.4998, 0.4997, 0.4998, 0.4994, 0.4999, 0.4998)
x3 = c(0.4995, 0.4995, 0.4995, 0.4995, 0.4995, 0.4996)
x4 = c(0.4998, 0.5005, 0.5005, 0.5002, 0.5003, 0.5004)
x5 = c(0.5000, 0.5005, 0.5008, 0.5007, 0.5008, 0.5010)
x6 = c(0.5008, 0.5009, 0.5010, 0.5005, 0.5006, 0.5009)
x7 = c(0.5000, 0.5001, 0.5002, 0.4995, 0.4996, 0.4997)
x8 = c(0.4993, 0.4994, 0.4999, 0.4996, 0.4996, 0.4997)
x9 = c(0.4995, 0.4995, 0.4997, 0.4992, 0.4995, 0.4992)
x10= c(0.4994, 0.4998, 0.5000, 0.4990, 0.5000, 0.5000)
data1 = rbind(x1, x2, x3, x4, x5, x6, x7, x8, x9, x10)
```

```
# Print LCL, CL, UCL.
# The mean and standard deviation are used.
result = rcc(data1, loc="mean", scale="sd", type="Xbar")
print(result)

# Note: X-bar chart is a default with the mean and sd
#       so the below is the same as the above.
rcc(data1)

# Summary of a control chart
summary(result)
RE(n=6, method="sd", correction=TRUE)

# The above limits are also calculated as
A3 = factors.cc(n=6, "A3")
xbarbar = mean(data1)
sbar = mean( apply(data1, 1, sd) )
c(xbarbar-A3*sbar, xbarbar, xbarbar+A3*sbar)

# Plot a control chart
plot(result, cex.text=0.8, x.text=4.1)
abline(v=5.5, lty=1, lwd=2, col="gold")
text( c(3,8), c(0.5005, 0.5005), labels=c("Group 1", "Group 2") )

# ==========
# Example 1b
# ----------
# The conventional X-bar chart with the range.
# Refer to Example 5 in Section 3.31 of ASTM (2010).
# The data are the same as in Example 1a.

# Print LCL, CL, UCL.
# The range is used for the scale estimate.
result = rcc(data1, loc="mean", scale="range")
print(result)

# Summary of a control chart
# Note: the RE is calculated based on the unbiased estimators.
summary(result)
RE(n=6, method="range", correction=TRUE)

# The above limits are also calculated as
A2 = factors.cc(n=6, "A2")
xbarbar = mean(data1)
Rbar = mean( apply(data1, 1, function(x) {diff(range(x))}) )
c(xbarbar-A2*Rbar, xbarbar, xbarbar+A2*Rbar)

# Plot a control chart
plot(result, cex.text=0.8, x.text=4.2)
abline(v=5.5, lty=1, lwd=2, col="gold")
text( c(3,8), c(0.5005, 0.5005), labels=c("Group 1", "Group 2") )

# ==========
```

```
# Example 1c
# ----------
# The median-MAD chart.
# Refer to Table 4.2 in Section 4.7 of Ryan (2000).
# Data: 20 subgroups with 4 observations each.
tmp = c(
72, 84, 79, 49, 56, 87, 33, 42, 55, 73, 22, 60, 44, 80, 54, 74,
97, 26, 48, 58, 83, 89, 91, 62, 47, 66, 53, 58, 88, 50, 84, 69,
57, 47, 41, 46, 13, 10, 30, 32, 26, 39, 52, 48, 46, 27, 63, 34,
49, 62, 78, 87, 71, 63, 82, 55, 71, 58, 69, 70, 67, 69, 70, 94,
55, 63, 72, 49, 49, 51, 55, 76, 72, 80, 61, 59, 61, 74, 62, 57 )
data2 = matrix(tmp, ncol=4, byrow=TRUE)

# Print LCL, CL, UCL.
# The median (location) and MAD (scale) are used.
result = rcc(data2, loc="median", scale="mad")
print(result)

# Summary of a control chart
summary(result)
# Note: the RE is calculated based on the unbiased estimators.
RE(n=4, method="median", correction=TRUE)
RE(n=4, method="mad", correction=TRUE)

# Plot a control chart
plot(result)

# ==========
# Example 1d
# ----------
# The HL2-Shamos chart.
# The data are the same as in Example 1c.

# Print LCL, CL, UCL.
# The HL2 (location) and Shamos (scale) are used.
result = rcc(data2, loc="HL2", scale="shamos")
print(result)

# Summary of a control chart
summary(result)
# Note: the RE is calculated based on the unbiased estimators.
RE(n=4, method="HL2", correction=TRUE)
RE(n=4, method="shamos", correction=TRUE)

# Plot a control chart
plot(result)


############
# S chart  #
############

# ==========
```

```
# Example 2a
# ----------
# The conventional S chart with the standard deviation.
# Refer to Example 3 in Section 3.31 of ASTM (2010).
# The data are the same as in Example 1a.

# Print LCL, CL, UCL.
# The standard deviaion (default) is used for the scale estimate.
result = rcc(data1, type="S")
print(result)

# Summary of a control chart
# Note: the RE is calculated based on the unbiased estimators.
summary(result)

# The above limits are also calculated as
B3 = factors.cc(n=6, "B3")
B4 = factors.cc(n=6, "B4")
sbar = mean( apply(data1, 1, sd) )
c(B3*sbar, sbar, B4*sbar)

# Plot a control chart
plot(result, cex.text=0.8, x.text=4.1)
abline(v=5.5, lty=1, lwd=2, col="gold")
text( c(3,8), c(0.0005, 0.0005), labels=c("Group 1", "Group 2") )


# ==========
# Example 2b
# ----------
# The S-type chart with the MAD.
# The data are the same as in Example 2a.

# Print LCL, CL, UCL.
# The mad (scale) are used.
result = rcc(data1, scale="mad", type="S")
print(result)

# Summary of a control chart
# Note: the RE is calculated based on the unbiased estimators.
summary(result)

# Plot a control chart
plot(result, cex.text=0.8, x.text=4.1)
abline(v=5.5, lty=1, lwd=2, col="gold")
text( c(3,8), c(0.00045, 0.00045), labels=c("Group 1", "Group 2") )


############
# R chart  #
############

# ==========
```

```
# Example 3a
# ----------
# The conventional R chart with the range.
# Refer to Example 5 in Section 3.31 of ASTM (2010).
# The data are the same as in Example 1a.

# Print LCL, CL, UCL.
# The range is used for the scale estimate.
# Unlike the S chart, scale="range" is not a default.
# Thus, for the conventional R chart, use the option (scale="range") as below.
result = rcc(data1, scale="range", type="R")
print(result)

# Summary of a control chart
# Note: the RE is calculated based on the unbiased estimators.
summary(result)

# The above limits are also calculated as
D3 = factors.cc(n=6, "D3")
D4 = factors.cc(n=6, "D4")
Rbar = mean( apply(data1, 1, function(x) {diff(range(x))}) )
c(D3*Rbar, Rbar, D4*Rbar)

# Plot a control chart
plot(result, cex.text=0.8, x.text=4.1)
abline(v=5.5, lty=1, lwd=2, col="gold")
text( c(3,8), c(0.0014, 0.0014), labels=c("Group 1", "Group 2") )


# ==========
# Example 3b
# ----------
# The R-type chart with the Shamos.
# Refer to Example 5 in Section 3.31 of ASTM (2010).
# The data are the same as in Example 3a.

# Print LCL, CL, UCL.
# The mad (scale) are used.
result = rcc(data1, scale="shamos", type="R")
print(result)

# Summary of a control chart
# Note: the RE is calculated based on the unbiased estimators.
summary(result)

# Plot a control chart
plot(result, cex.text=0.8, x.text=4.1)
abline(v=5.5, lty=1, lwd=2, col="gold")
text( c(3,8), c(0.0014, 0.0014), labels=c("Group 1", "Group 2") )


############
# vignette #
```

```
############
if (interactive()) vignette("rcc", package="rQCC")
```

`Shamos` *Shamos estimate*

Description

Calculates the conventional Shamos, unbiased Shamos and unbiased squared Shamos estimates. The conventional Shamos is calculated by `shamos` which is Fisher-consistent under the normal distribution. Note that it is not unbiased with a sample of finite size. The unbiased Shamos estimate under the normal distribution is calculated by `shamos.unbiased` with a finite-sample unbiasing factor. The unbiased *squared* Shamos estimate under the normal distribution is calculated by `shamos2.unbiased` with a finite-sample unbiasing factor.

Usage

```
shamos(x, constant=1.048358, na.rm = FALSE,  IncludeEqual=FALSE)

shamos.unbiased(x, constant=1.048358, na.rm = FALSE,  IncludeEqual=FALSE)

shamos2.unbiased(x, constant=1.048358, na.rm = FALSE, IncludeEqual=FALSE)
```

Arguments

`x`	a numeric vector of observations.
`constant`	Correction factor for the Fisher-consistency under the normal distribution
`na.rm`	a logical value indicating whether NA values should be stripped before the computation proceeds.
`IncludeEqual`	`FALSE` (default) calculates median of $\vert X_i - X_j\vert$ with $i < j$, while `TRUE` calculates median of $\vert X_i - X_j\vert$ with $i \le j$.

Details

The Shamos estimator is defined as

$$\text{Shamos} = \texttt{constant} \times \underset{i<j}{\text{median}} \big(|X_i - X_j|\big)$$

where $i, j = 1, 2, \ldots, n$. The default value (`constant=1.048358`) ensures the Fisher-consistency under the normal distribution. Note that `constant` $= 1/\{\sqrt{2}\,\Phi^{-1}(3/4)\} \approx 1.048358$.

The unbiased Shamos is defined as

$$\text{Shamos} = \texttt{constant} \times \underset{i<j}{\text{median}} \big(|X_i - X_j|\big)/c_6(n)$$

for $i, j = 1, 2, \ldots, n$, where $c_6(n)$ is the finite-sample unbiasing factor. Note that $c_6(n)$ notation is used in Park et. al (2020), and $c_6(n)$ is calculated using the function `c4.factor`{rQCC} with `method="shamos"` option.

The unbiased squared Shamos is defined as the squared `shamos`{rQCC} divided by $w_6(n)$ where $w_6(n)$ is the finite-sample unbiasing factor. Note that $w_6(n)$ notation is used in Park et. al (2020), and $w_6(n)$ is calculated using the function `w4.factor`{rQCC} with `method="shamos2"` option. Note that the square of the conventional Shamos estimator is Fisher-consistent for the variance (σ^2) under the normal distribution, but it is not unbiased with a sample of finite size.

Value

They return a numeric value.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Shamos, M. I. (1976). Geometry and statistics: Problems at the interface. In Traub, J. F., editor, *Algorithms and Complexity: New Directions and Recent Results*, pages 251–280. Academic Press, New York.

Lèvy-Leduc, C., Boistard, H., Moulines, E., Taqqu, M. S., and Reisen, V. A. (2011). Large sample behaviour of some well-known robust estimators under long-range dependence. *Statistics*, **45**, 59–71.

See Also

`mad.unbiased`{rQCC} for calculating the *unbiased* sample MAD.

mad{stats} for calculating the Fisher-consistent sample MAD.

`c4.factor`{rQCC} for finite-sample unbiasing factor for the standard deviation (σ) under the normal distribution.

`w4.factor`{rQCC} for finite-sample unbiasing factor for the squared Shamos estimator of the variance (σ^2) under the normal distribution.

`finite.breakdown`{rQCC} for calculating the finite-sample breakdown point.

Examples

```
x = c(0:10, 50)

# Fisher-consistent Shamos, but not unbiased with a finite sample.
shamos(x)

# Unbiased Shamos.
shamos.unbiased(x)

# Fisher-consistent squared Shamos, but not unbiased with a finite sample.
shamos(x)^2
```

```
# Unbiased squared Shamos.
shamos2.unbiased(x)
```

`unbiasing.factor` *Finite-sample unbiasing factor*

Description

Finite-sample unbiasing factor for estimating the standard deviation (σ) and the variance (σ^2) under the normal distribution.

Usage

```
c4.factor(n, method=c("sd","range", "mad","shamos"))
w4.factor(n, method=c("mad2","shamos2"))
```

Arguments

`n`	sample size ($n \geq 1$).
`method`	a character string specifying the estimator, must be one of `"sd"` (default), `"range"`, `"mad"`, `"shamos"` for `c4.factor`, and one of `"mad2"` (default), `"shamos2"` for `w4.factor`.

Details

The conventional sample standard deviation, range, median absolute deviation (MAD) and Shamos estimators are Fisher-consistent under the normal distribution, but they are not unbiased with a sample of finite size.

Using the sample standard deviation, an unbiased estimator of the standard deviation (σ) is calculated by
`sd(x)/c4.factor(length(x),method="sd")`

Using the range (maximum minus minimum), an unbiased estimator of σ is calculated by
`diff(range(x))/c4.factor(length(x),method="range")`

Using the median absolute deviation (mad{stats}), an unbiased estimator of σ is calculated by
`mad(x)/c4.factor(length(x),method="mad")`

Using the Shamos estimator (`shamos`{rQCC}), an unbiased estimator of σ is calculated by
`shamos(x)/c4.factor(length(x),method="shamos")`

Note that the formula for the unbiasing factor $c_4(n)$ is given by

$$c_4(n) = \sqrt{\frac{2}{n-1}} \cdot \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}.$$

The *squared* MAD and *squared* Shamos are Fisher-consistent for the variance (σ^2) under the normal distribution, but they are not unbiased with a sample of finite size.

An unbiased estimator of the variance (σ^2) is obtained using the finite-sample unbiasing factor (`w4.factor`).

Using the squared MAD, an unbiased estimator of σ^2 is calculated by
`mad(x)^2/w4.factor(length(x),method="mad2")`

Using the squared Shamos estimator, an unbiased estimator of σ^2 is calculated by
`shamos(x)^2/w4.factor(length(x),method="shamos2")`

The finite-sample unbiasing factors for the median absolute deviation (MAD) and Shamos estimators are obtained for $n = 1, 2, \ldots, 100$ using the extensive Monte Carlo simulation with 1E07 replicates. For the case of $n > 100$, they are obtained using the method of Hayes (2014).

Value

It returns a numeric value.

Author(s)

Chanseok Park and Min Wang

References

Park, C., H. Kim, and M. Wang (2020). Investigation of finite-sample properties of robust location and scale estimators. *Communications in Statistics - Simulation and Computation*, To appear. https://doi.org/10.1080/03610918.2019.1699114

Shamos, M. I. (1976). Geometry and statistics: Problems at the interface. In Traub, J. F., editor, *Algorithms and Complexity: New Directions and Recent Results*, 251–280. Academic Press, New York.

Hayes, K. (2014). Finite-sample bias-correction factors for the median absolute deviation. *Communications in Statistics: Simulation and Computation*, **43**, 2205–2212.

See Also

mad{stats} for the Fisher-consistent median absolute deviation (MAD) estimator of the standard deviation (σ) under the normal distribution.

`mad.unbiased`{rQCC} for finite-sample unbiased median absolute deviation (MAD) estimator of the standard deviation (σ) under the normal distribution.

`shamos`{rQCC} for the Fisher-consistent Shamos estimator of the standard deviation (σ) under the normal distribution.

`shamos.unbiased`{rQCC} for finite-sample unbiased Shamos estimator of the standard deviation (σ) under the normal distribution.

`n.times.eBias.of.mad`{rQCC} for the values of the empirical bias of the median absolute deviation (MAD) estimator under the standard normal distribution.

`n.times.eBias.of.shamos`{rQCC} for the values of the empirical bias of the Shamos estimator under the standard normal distribution.

mad2.unbiased{rQCC} for finite-sample unbiased squared MAD estimator of the variance (σ^2) under the normal distribution.

shamos2.unbiased{rQCC} for finite-sample unbiased squared Shamos estimator of the variance (σ^2) under the normal distribution.

n.times.evar.of.mad{rQCC} for the values of the empirical variance of the median absolute deviation (MAD) estimator under the standard normal distribution.

n.times.evar.of.shamos{rQCC} for the values of the empirical variance of the Shamos estimator under the standard normal distribution.

Examples

```
# unbiasing factor for estimating the standard deviation
c4.factor(n=10, method="sd")
c4.factor(n=10, method="mad")
c4.factor(n=10, method="shamos")

# Note: d2 notation is widely used for the bias-correction of the range.
d2 = c4.factor(n=10, method="range")
d2


# unbiasing factor for estimating the variance
w4.factor(n=10, "mad2")
w4.factor(n=10, "shamos2")
```

Index