Package 'ConfIntVariance'

January 20, 2025

Type Package

Title Confidence Interval for the Univariate Population Variance without Normality Assumption

Version 1.0.2

Date 2019-03-06

Author Mathias Fuchs

Maintainer Mathias Fuchs<mathias@mathiasfuchs.de>

Description

Surrounds the usual sample variance of a univariate numeric sample with a confidence interval for the population variance. This has been done so far only under the assumption that the underlying distribution is normal. Under the hood, this package implements the unique least-variance unbiased estimator of the variance of the sample variance, in a formula that is equivalent to estimating kurtosis and square of the population variance in an unbiased way and combining them according to the classical formula into an estimator of the variance of the sample variance. Both the sample variance and the estimator of its variance are U-statistics. By the theory of U-statistic, the resulting estimator is unique. See Fuchs, Krautenbacher (2016) <doi:10.1080/15598608.2016.1158675> and the references therein for an overview of unbiased estimation of variances of U-statistics.

License GPL-3

NeedsCompilation no

Repository CRAN

Date/Publication 2019-03-10 10:02:39 UTC

Contents

ConfIntVariance-package

Confidence Interval for the Univariate Population Variance without Normality Assumption

Description

Surrounds the usual sample variance of a univariate numeric sample with a confidence interval for the population variance. This has been done so far only under the assumption that the underlying distribution is normal. Under the hood, this package implements the unique least-variance unbiased estimator of the variance of the sample variance, in a formula that is equivalent to estimating kurtosis and square of the population variance in an unbiased way and combining them according to the classical formula into an estimator of the variance of the sample variance. Both the sample variance and the estimator of its variance are U-statistics. By the theory of U-statistic, the resulting estimator is unique. See Fuchs, Krautenbacher (2016) <doi:10.1080/15598608.2016.1158675> and the references therein for an overview of unbiased estimation of variances of U-statistics.

Details

The DESCRIPTION file:

Package:	ConfIntVariance
Type:	Package
Title:	Confidence Interval for the Univariate Population Variance without Normality Assumption
Version:	1.0.2
Date:	2019-03-06
Author:	Mathias Fuchs
Maintainer:	Mathias Fuchs<mathias@mathiasfuchs.de>
Description:	Surrounds the usual sample variance of a univariate numeric sample with a confidence interval for the populati
License:	GPL-3

Index of help topics:

ConfIntVariance-package	Confidence Interval for the Univariate Population Variance without Normality Assumption
varwci	varwci

A package providing one function varwci which is short for "variance with confidence interval."

Author(s)

Mathias Fuchs Maintainer: Mathias Fuchs<mathias@mathiasfuchs.de>

References

www.mathiasfuchs.de/b3.html

Examples

```
##
## Example: throwing a dice
##

                                        # True quantities that do not depend on n
trueMeanOfDice <- mean(1:6)


                                        # The true variance of the dice
                                        # This is the quantity that we
                                        # want to estimate by embracing
                                        # with a confidence interval
                                        # instead of just estimating
                                        # with a point estimator as is
                                        # done in the function var
trueVarianceOfDice <- mean((1:6)^2) - trueMeanOfDice^2
trueFourthCentralMomentOfDice <- mean(((1:6)-trueMeanOfDice)^4)

                                    # this requires some scribbling with paper and pencil
                                        # (or a study of Hoeffding 1948)
trueVarianceOfSampleVarianceOfDice <- function(n)
(trueFourthCentralMomentOfDice - trueVarianceOfDice^2 * (n-3)/(n-1))/n

##
## Simulation study: compute the coverage probability of
## the confidence interval by computing the probability
## that it contains the true value.
## We want that probability to be equal to the confidence level 0.95,
## not more and not less. (If it was higher, the test would be too conservative).
##

                                        # number of times we draw a
                                        # sample and compute a confidence interval
N <- 1e4
trueValueCovered <- sapply(
    1:N,
    function(i) {
                                        # throw a dice 100 times
        x <- sample(6, 100, replace=TRUE)
                                        # compute our confidence interval
        ci <- varwci(x)
                                        # We know that the true variance
                                        # of the dice is
                                        # 35/12 = 2.916666...
                                        # Record the boolean whether the
                                        # confidence interval contains
                                        # the true value.
```

```
        (35/12 > ci[1] && 35/12 < ci[2])
    }
)

                                        # Result of simulation study:
                                        # Will be close to 0.95.
print(mean(trueValueCovered))
```

---

varwci *varwci*

Description

Surround the univariate variance estimator of the function var with a confidence interval, not assuming normality

Usage

```
varwci(x, conf.level=0.95)
```

Arguments

x	A one-dimensional numeric vector
conf.level	The confidence level for the confidence interval. Defaults to 0.95

Value

Returns a vector with two entries: the lower and the upper bound of the confidence interval, and the following attributes:

point.estimator The usual sample variance at the center of the interval

conf.level The confidence level used

var.SampleVariance The estimated variance of the sample variance

Warning

On very small sample sizes, the result is NA because there is insufficient information on the variance estimation

Note

The underlying theory is that of U-statistics. See Hoeffding 1948.

Author(s)

Mathias Fuchs

References

http://dx.doi.org/10.1080/15598608.2016.1158675 and https://mathiasfuchs.de/b3.html

Examples

```
##
## Example: throwing a dice
##

# throw a dice 100 times
s <- sample(6, 100, replace=TRUE)

# the standard point estimator for the variance
print(var(s))

# contains the true value 2.9166 with a probability of 95 percent.
print(varwci(s))

##
## Check the coverage probability of the confidence interval
##

                                        # True quantities that do not depend on n
trueMeanOfDice <- mean(1:6)
trueVarianceOfDice <- mean((1:6)^2) - trueMeanOfDice^2

## see package description for more details
                                        # number of times we draw a
                                        # sample and compute a confidence interval
N <- 1e4
trueValueCovered <- rep(NA, N)
for (i in 1:N) {
    if (i %% 1e3 == 0) print(i)
                                        # throw a dice 100 times
    x <- sample(6, 100, replace=TRUE)
                                        # compute our confidence interval
    ci <- varwci(x)
                                        # We know that the true variance
                                        # of the dice is 91/6 - 49/4 = 2.916666...
                                    # did the confidence interval contain the correct value?
    trueValueCovered[i] <- (trueVarianceOfDice > ci[1] && trueVarianceOfDice < ci[2])
}

                                      # Result of simulation study: should be close to 0.95
print(mean(trueValueCovered))
```

Index