Package ‘OneTwoSamples’

January 20, 2025

Type Package

Title Deal with One and Two (Normal) Samples

Version 1.1-0

Date 2023-03-21

Author Frederic Bertrand [cre] (<https://orcid.org/0000-0002-0837-8281>),
Ying-Ying Zhang (Robert) [aut]

Maintainer Frederic Bertrand <frederic.bertrand@utt.fr>

Depends R (>= 3.3), methods

Description We introduce an R function
one_two_sample() which can deal with one and two (normal)
samples, Ying-Ying Zhang, Yi Wei (2012) <doi:10.2991/asshm-13.2013.29>.
For one normal sample x, the function reports
descriptive statistics, plot, interval estimation and test of
hypothesis of x. For two normal samples x and y, the function
reports descriptive statistics, plot, interval estimation and
test of hypothesis of x and y, respectively. It also reports
interval estimation and test of hypothesis of mu1-mu2 (the
difference of the means of x and y) and sigma1^2 / sigma2^2
(the ratio of the variances of x and y), tests whether x and y
are from the same population, finds the correlation coefficient
of x and y if x and y have the same length.

LazyLoad yes

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2023-03-28 16:30:02 UTC

Contents

`OneTwoSamples-package` *Deal with One and Two (Normal) Samples*

Description

In this package, we introduce an R function one_two_sample() which can deal with one and two (normal) samples, Ying-Ying Zhang, Yi Wei (2012), doi:10.2991/asshm13.2013.29. For one normal sample x, the function reports descriptive statistics, plot, interval estimation and test of hypothesis of x. For two normal samples x and y, the function reports descriptive statistics, plot, interval estimation and test of hypothesis of x and y, respectively. It also reports interval estimation and test of hypothesis of mu1-mu2 (the difference of the means of x and y) and sigma1^2/sigma2^2 (the ratio of the variances of x and y), tests whether x and y are from the same population, finds the correlation coefficient of x and y if x and y have the same length.

Details

Package:	OneTwoSamples
Type:	Package
Version:	1.1-0
Date:	2023-03-22
License:	GPL (>= 2)

The most important functions are: one_two_sample() and one_sample().

Author(s)

Ying-Ying Zhang (Robert)

Maintainer: Frederic Bertrand <frederic.bertrand@utt.fr>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
library("OneTwoSamples")
```

`data_outline` *Compute various descriptive statistics*

Description

Compute various descriptive statistics of x, such as mean, median, skewness, and kurtosis, etc.

Usage

```
data_outline(x)
```

Arguments

`x`	A numeric vector.

Value

A data.frame with variables:

`N`	The length.
`Mean`	The mean.
`Var`	The variance.
`std_dev`	Standard deviation.
`Median`	The median.
`std_mean`	The standard error of the sample mean.
`CV`	The coefficient of variation.
`CSS`	The corrected sum of squares.
`USS`	The uncorrected sum of squares.
`R`	The extreme difference.
`R1`	The half extreme difference, or the difference of upper quartile and lower quartile.
`Skewness`	The coefficient of skewness.
`Kurtosis`	The coefficient of kurtosis.
`row.names`	1.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
data_outline(x)
```

---

detail — *Show details of an object*

Description

Show details of an object.

Usage

```
detail(x)
```

Arguments

x	Any R object to be tested.

Value

A list with components:

x	The argument x.
isS4	Logical, indicates whether x is an S4 object.
isObject	Logical, indicates whether x is an object, i.e., with a class attribute.
class	The class of x.
attributes	The attributes of x. Usually result$attributes is also a list.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

See Also

isS4, is.object, class, attributes

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
t = t.test(x); t
detail(t)
```

---

interval_estimate1 *Two sided interval estimation of* mu *of one normal sample*

Description

Compute the two sided interval estimation of mu of one normal sample when the population variance is known or unknown.

Usage

```
interval_estimate1(x, sigma = -1, alpha = 0.05)
```

Arguments

x	A numeric vector.
sigma	The standard deviation of the population. sigma>=0 indicates it is known, sigma<0 indicates it is unknown. Default to unknown standard deviation.
alpha	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

mean	The sample mean.
df	The degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
interval_estimate1(x, sigma = 0.2)
interval_estimate1(x)
```

---

interval_estimate2 *Two sided interval estimation of* mu1-mu2 *of two normal samples*

Description

Compute the two sided interval estimation of mu1-mu2 of two normal samples when the population variances are known, unknown equal, or unknown unequal.

Usage

```
interval_estimate2(x, y, sigma = c(-1, -1), var.equal = FALSE, alpha = 0.05)
```

Arguments

x	A numeric vector.
y	A numeric vector.
sigma	A numeric vector of length 2, which contains the standard deviations of two populations. When the standard deviations are known, input it, then the function computes the interval endpoints using normal population; when the standard deviations are unknown, ignore it, now we need to consider whether the two populations have equal variances. See var.equal below.
var.equal	A logical variable indicating whether to treat the two variances as being equal. If TRUE then the pooled variance is used to estimate the variance otherwise the Welch (or Satterthwaite) approximation to the degrees of freedom is used.
alpha	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

mean	The difference of sample means xb-yb.
df	The degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
y=rnorm(20, mean = 2, sd = 0.3); y

interval_estimate2(x, y, sigma = c(0.2, 0.3))
interval_estimate2(x, y, var.equal = TRUE)
interval_estimate2(x, y)
```

`interval_estimate3` *Two sided interval estimation of* `mu` *of one non-normal sample with large sample size*

Description

Compute the two sided interval estimation of `mu` of one non-normal sample with large sample size when the population variance is known or unknown.

Usage

```
interval_estimate3(x, sigma = -1, alpha = 0.05)
```

Arguments

`x`	A numeric vector.
`sigma`	The standard deviation of the population. `sigma>=0` indicates it is known, `sigma<0` indicates it is unknown. Default to unknown standard deviation.
`alpha`	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

`mean`	The sample mean.
`a`	The confidence lower limit.
`b`	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) `<robertzhangyying@qq.com>`

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x = rexp(50, 1/2); x
interval_estimate3(x)
```

---

interval_estimate4 *Two sided or one sided interval estimation of* mu *of one normal sample*

Description

Compute the two sided or one sided interval estimation of mu of one normal sample when the population variance is known or unknown.

Usage

```
interval_estimate4(x, sigma = -1, side = 0, alpha = 0.05)
```

Arguments

x	A numeric vector.
sigma	The standard deviation of the population. sigma>=0 indicates it is known, sigma<0 indicates it is unknown. Default to unknown standard deviation.
side	A parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input side = -1; when computing the one sided lower limit, input side = 1; when computing the two sided limits, input side = 0 (default).
alpha	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

mean	The sample mean.
df	The degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
interval_estimate4(x, sigma = 0.2, side = -1)
interval_estimate4(x, side = 1)
```

---

interval_estimate5 *Two sided or one sided interval estimation of* mu1-mu2 *of two normal samples*

Description

Compute the two sided or one sided interval estimation of mu1-mu2 of two normal samples when the population variances are known, unknown equal, or unknown unequal.

Usage

```
interval_estimate5(x, y, sigma = c(-1, -1), var.equal = FALSE, side = 0, alpha = 0.05)
```

Arguments

x	A numeric vector.
y	A numeric vector.
sigma	A numeric vector of length 2, which contains the standard deviations of two populations. When the standard deviations are known, input it, then the function computes the interval endpoints using normal population; when the standard deviations are unknown, ignore it, now we need to consider whether the two populations have equal variances. See var.equal below.
var.equal	A logical variable indicating whether to treat the two variances as being equal. If TRUE then the pooled variance is used to estimate the variance otherwise the Welch (or Satterthwaite) approximation to the degrees of freedom is used.
side	A parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input side = -1; when computing the one sided lower limit, input side = 1; when computing the two sided limits, input side = 0 (default).
alpha	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

mean	The difference of sample means xb-yb.
df	The degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
y=rnorm(20, mean = 2, sd = 0.3); y

interval_estimate5(x, y, sigma = c(0.2, 0.3), side = -1)
interval_estimate5(x, y, var.equal = TRUE)
interval_estimate5(x, y)
```

---

interval_var1 *Two sided interval estimation of* sigma^2 *of one normal sample*

Description

Compute the two sided interval estimation of sigma^2 of one normal sample when the population mean is known or unknown.

Usage

```
interval_var1(x, mu = Inf, alpha = 0.05)
```

Arguments

x	A numeric vector.
mu	The population mean. When it is known, input it, and the function computes the interval endpoints using a chi-square distribution with degree of freedom n. When it is unknown, ignore it, and the function computes the interval endpoints using a chi-square distribution with degree of freedom n-1.
alpha	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

`var`	The estimate of the population variance. When the population mean `mu` is known, `var` = `mean((x-mu)^2)`. When `mu` is unknown, `var` = `var(x)`.
`df`	The degree of freedom.
`a`	The confidence lower limit.
`b`	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) `<robertzhangyying@qq.com>`

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
interval_var1(x, mu = 1)
interval_var1(x)
```

`interval_var2` *Two sided interval estimation of* `sigma1^2 / sigma2^2` *of two normal samples*

Description

Compute the two sided interval estimation of `sigma1^2 / sigma2^2` of two normal samples when the population means are known or unknown.

Usage

```
interval_var2(x, y, mu = c(Inf, Inf), alpha = 0.05)
```

Arguments

`x`	A numeric vector.
`y`	A numeric vector.
`mu`	The population means. When it is known, input it, and the function computes the interval endpoints using an F distribution with degree of freedom `(n1, n2)`. When it is unknown, ignore it, and the function computes the interval endpoints using an F distribution with degree of freedom `(n1-1, n2-1)`.
`alpha`	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

rate	The estimate of the ratio of population variances, rate = Sx2/Sy2. When the population means mu is known, Sx2 = 1/n1*sum((x-mu[1])^2) and Sy2 = 1/n2*sum((y-mu[2])^2. When mu is unknown, Sx2 = var(x) and Sy2 = var(y).
df1	The first degree of freedom.
df2	The second degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
y=rnorm(20, mean = 2, sd = 0.3); y
interval_var2(x, y, mu = c(1,2))
interval_var2(x, y)
```

---

interval_var3 *Two sided or one sided interval estimation of* sigma^2 *of one normal sample*

Description

Compute the two sided or one sided interval estimation of sigma^2 of one normal sample when the population mean is known or unknown.

Usage

```
interval_var3(x, mu = Inf, side = 0, alpha = 0.05)
```

Arguments

`x`	A numeric vector.
`mu`	The population mean. When it is known, input it, and the function computes the interval endpoints using a chi-square distribution with degree of freedom `n`. When it is unknown, ignore it, and the function computes the interval endpoints using a chi-square distribution with degree of freedom `n-1`.
`side`	A parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input `side = -1`; when computing the one sided lower limit, input `side = 1`; when computing the two sided limits, input `side = 0` (default).
`alpha`	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

`var`	The estimate of the population variance. When the population mean `mu` is known, `var = mean((x-mu)^2)`. When `mu` is unknown, `var = var(x)`.
`df`	The degree of freedom.
`a`	The confidence lower limit.
`b`	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
interval_var3(x, mu = 1, side = -1)
interval_var3(x)
```

interval_var4 *Two sided or one sided interval estimation of* sigma1^2 / sigma2^2 *of two normal samples*

Description

Compute the two sided or one sided interval estimation of sigma1^2 / sigma2^2 of two normal samples when the population means are known or unknown.

Usage

```
interval_var4(x, y, mu = c(Inf, Inf), side = 0, alpha = 0.05)
```

Arguments

x	A numeric vector.
y	A numeric vector.
mu	The population means. When it is known, input it, and the function computes the interval endpoints using an F distribution with degree of freedom (n1, n2). When it is unknown, ignore it, and the function computes the interval endpoints using an F distribution with degree of freedom (n1-1, n2-1).
side	A parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input side = -1; when computing the one sided lower limit, input side = 1; when computing the two sided limits, input side = 0 (default).
alpha	The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A data.frame with variables:

rate	The estimate of the ratio of population variances, rate = Sx2/Sy2. When the population means mu is known, Sx2 = 1/n1*sum((x-mu[1])^2) and Sy2 = 1/n2*sum((y-mu[2])^2. When mu is unknown, Sx2 = var(x) and Sy2 = var(y).
df1	The first degree of freedom.
df2	The second degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
y=rnorm(20, mean = 2, sd = 0.3); y
interval_var4(x, y, mu = c(1,2), side = -1)
interval_var4(x, y)
```

`mean_test1` *Two sided or one sided test of hypothesis of* `mu` *of one normal sample*

Description

Compute the two sided or one sided test of hypothesis of `mu` of one normal sample when the population variance is known or unknown.

Usage

```
mean_test1(x, mu = 0, sigma = -1, side = 0)
```

Arguments

`x`	A numeric vector.
`mu`	`mu` is `mu0` in the null hypothesis. Default is 0, i.e., `H0: mu = 0`.
`sigma`	The standard deviation of the population. `sigma>=0` indicates it is known, `sigma<0` indicates it is unknown. Default to unknown standard deviation.
`side`	A parameter used to control two sided or one sided test of hypothesis. When inputting `side = 0` (default), the function computes two sided test of hypothesis, and `H1: mu != mu0`; when inputting `side = -1` (or a number < 0), the function computes one sided test of hypothesis, and `H1: mu < mu0`; when inputting `side = 1` (or a number > 0), the function computes one sided test of hypothesis, and `H1: mu > mu0`.

Value

A data.frame with variables:

`mean`	The sample mean.
`df`	The degree of freedom.
`statistic`	The statistic, when `sigma>=0`, `statistic = Z = (xb-mu)/(sigma/sqrt(n))`; when `sigma<0`, `statistic = T = (xb-mu)/(sd(x)/sqrt(n))`.
`p_value`	The P value.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
mean_test1(x, mu = 1, sigma = 0.2, side = 1)
mean_test1(x, mu = 1)
```

---

mean_test2 — *Two sided or one sided test of hypothesis of* `mu1` *and* `mu2` *of two normal samples*

Description

Compute the two sided or one sided test of hypothesis of `mu1` and `mu2` of two normal samples when the population variances are known, unknown equal, or unknown unequal.

Usage

```
mean_test2(x, y, sigma = c(-1, -1), var.equal = FALSE, side = 0)
```

Arguments

`x`	A numeric vector.
`y`	A numeric vector.
`sigma`	A numeric vector of length 2, which contains the standard deviations of two populations. When the standard deviations are known, input it, then the function computes the interval endpoints using normal population; when the standard deviations are unknown, ignore it, now we need to consider whether the two populations have equal variances. See `var.equal` below.
`var.equal`	A logical variable indicating whether to treat the two variances as being equal. If `TRUE` then the pooled variance is used to estimate the variance otherwise the Welch (or Satterthwaite) approximation to the degrees of freedom is used.
`side`	A parameter used to control two sided or one sided test of hypothesis. When inputting `side = 0` (default), the function computes two sided test of hypothesis, and `H1: mu1 != mu2`; when inputting `side = -1` (or a number < 0), the function computes one sided test of hypothesis, and `H1: mu1 < mu2`; when inputting `side = 1` (or a number > 0), the function computes one sided test of hypothesis, and `H1: mu1 > mu2`.

Value

A data.frame with variables:

`mean`	The difference of sample means xb-yb.
`df`	The degree of freedom.
`statistic`	The statistic, when `all(sigma>=0)`, `statistic` = `Z`; otherwise, `statistic` = `T`.
`p_value`	The P value.

Author(s)

Ying-Ying Zhang (Robert) `<robertzhangyying@qq.com>`

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
y=rnorm(20, mean = 2, sd = 0.3); y
mean_test2(x, y, sigma = c(0.2, 0.3), side = 1)
mean_test2(x, y, var.equal = TRUE, side = 1)
mean_test2(x, y, side = 1)
```

`one_sample` *Deal with one (normal) sample*

Description

Deal with one sample `x`, especially normal. Report descriptive statistics, plot, interval estimation and test of hypothesis of `x`.

Usage

```
one_sample(x, mu = Inf, sigma = -1, side = 0, alpha = 0.05)
```

Arguments

`x`	A numeric vector.
`mu`	`mu` plays two roles. In two sided or one sided interval estimation (or test of hypothesis) of `sigma^2` of one normal sample, `mu` is the population mean. When it is known, input it, and the function computes the interval endpoints (or chi-square statistic) using a chi-square distribution with degree of freedom `n`. When it is unknown, ignore it (the default), and the function computes the interval endpoints (or chi-square statistic) using a chi-square distribution with degree of freedom `n-1`.

In two sided or one sided test of hypothesis of mu of one normal sample, mu is mu0 in the null hypothesis, and mu0 = if (mu < Inf) mu else 0.

sigma
: sigma plays two roles.

In two sided or one sided interval estimation (or test of hypothesis) of mu of one normal sample, sigma is the standard deviation of the population. sigma>=0 indicates it is known, and the function computes the interval endpoints (or Z statistic) using a standard normal distribution. sigma<0 indicates it is unknown, and the function computes the interval endpoints (or T statistic) using a t distribution with degree of freedom n-1. Default to unknown standard deviation.

In two sided or one sided test of hypothesis of sigma^2 of one normal sample, sigma is sigma0 in the null hypothesis. Default is 1, i.e., H0: sigma^2 = 1.

side
: side plays two roles and is used in four places.

In two sided or one sided interval estimation of mu of one normal sample, side is a parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input side = -1; when computing the one sided lower limit, input side = 1; when computing the two sided limits, input side = 0 (default).

In two sided or one sided interval estimation of sigma^2 of one normal sample, side is a parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input side = -1; when computing the one sided lower limit, input side = 1; when computing the two sided limits, input side = 0 (default).

In two sided or one sided test of hypothesis of mu of one normal sample, side is a parameter used to control two sided or one sided test of hypothesis. When inputting side = 0 (default), the function computes two sided test of hypothesis, and H1: mu != mu0; when inputting side = -1 (or a number < 0), the function computes one sided test of hypothesis, and H1: mu < mu0; when inputting side = 1 (or a number > 0), the function computes one sided test of hypothesis, and H1: mu > mu0.

In two sided or one sided test of hypothesis of sigma^2 of one normal sample, side is a parameter used to control two sided or one sided test of hypothesis. When inputting side = 0 (default), the function computes two sided test of hypothesis, and H1: sigma^2 != sigma0^2; when inputting side = -1 (or a number < 0), the function computes one sided test of hypothesis, and H1: sigma^2 < sigma0^2; when inputting side = 1 (or a number > 0), the function computes one sided test of hypothesis, and H1: sigma^2 > sigma0^2.

alpha
: The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A list with the following components:

mu_interval
: It contains the results of interval estimation of mu.

mu_hypothesis
: It contains the results of test of hypothesis of mu.

sigma_interval
: It contains the results of interval estimation of sigma.

sigma_hypothesis
: It contains the results of test of hypothesis of sigma.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
one_sample(x, mu = 1, sigma = 0.2, side = 1)
one_sample(x, sigma = 0.2, side = 1)
one_sample(x, mu = 1, side = 1)
one_sample(x)
```

---

one_two_sample *Deal with one and two (normal) samples*

Description

Deal with one and two (normal) samples. For one normal sample x, the function reports descriptive statistics, plot, interval estimation and test of hypothesis of x. For two normal samples x and y, the function reports descriptive statistics, plot, interval estimation and test of hypothesis of x and y, respectively. It also reports interval estimation and test of hypothesis of mu1-mu2 (the difference of the means of x and y) and sigma1^2/sigma2^2 (the ratio of the variances of x and y), tests whether x and y are from the same population, finds the correlation coefficient of x and y if x and y have the same length.

Usage

```
one_two_sample(x, y = NULL, mu = c(Inf, Inf), sigma = c(-1, -1),
                var.equal = FALSE, ratio = 1, side = 0, alpha = 0.05)
```

Arguments

`x`	A numeric vector.
`y`	A numeric vector.
`mu`	If `y = NULL`, i.e., there is only one sample. See the argument `mu` in `one_sample`. For two normal samples `x` and `y`, `mu` plays one role: the population means. However, `mu` is used in two places: one is the two sided or one sided interval estimation of `sigma1^2 / sigma2^2` of two normal samples, another is the two sided or one sided test of hypothesis of `sigma1^2` and `sigma2^2` of two normal samples. When `mu` is known, input it, and the function computes the interval endpoints (or the F value) using an F distribution with degree of freedom `(n1, n2)`. When it is unknown, ignore it, and the function computes the interval endpoints (or the F value) using an F distribution with degree of freedom `(n1-1, n2-1)`.

sigma
: If y = NULL, i.e., there is only one sample. See the argument sigma in one_sample. For two normal samples x and y, sigma plays one role: the population standard deviations. However, sigma is used in two places: one is the two sided or one sided interval estimation of mu1-mu2 of two normal samples, another is the two sided or one sided test of hypothesis of mu1 and mu2 of two normal samples. When the standard deviations are known, input it, then the function computes the interval endpoints using normal population; when the standard deviations are unknown, ignore it, now we need to consider whether the two populations have equal variances. See var.equal below.

var.equal
: A logical variable indicating whether to treat the two variances as being equal. If TRUE then the pooled variance is used to estimate the variance otherwise the Welch (or Satterthwaite) approximation to the degrees of freedom is used.

ratio
: The hypothesized ratio of the population variances of x and y. It is used in var.test(x, y, ratio = ratio, ...), i.e., when computing the interval estimation and test of hypothesis of sigma1^2 / sigma2^2 when mu1 or mu2 is unknown.

side
: If y = NULL, i.e., there is only one sample. See the argument side in one_sample. For two normal samples x and y, sigma is used in four places: interval estimation of mu1-mu2, test of hypothesis of mu1 and mu2, interval estimation of sigma1^2 / sigma2^2, test of hypothesis of sigma1^2 and sigma2^2. In interval estimation of mu1-mu2 or sigma1^2 / sigma2^2, side is a parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input side = -1 (or a number < 0); when computing the one sided lower limit, input side = 1 (or a number > 0); when computing the two sided limits, input side = 0 (default). In test of hypothesis of mu1 and mu2 or sigma1^2 and sigma2^2, side is a parameter used to control two sided or one sided test of hypothesis. When inputting side = 0 (default), the function computes two sided test of hypothesis, and H1: mu1 != mu2 or H1: sigma1^2 != sigma2^2; when inputting side = -1 (or a number < 0), the function computes one sided test of hypothesis, and H1: mu1 < mu2 or H1: sigma1^2 < sigma2^2; when inputting side = 1 (or a number > 0), the function computes one sided test of hypothesis, and H1: mu1 > mu2 or H1: sigma1^2 > sigma2^2.

alpha
: The significance level, a real number in [0, 1]. Default to 0.05. 1-alpha is the degree of confidence.

Value

A list with the following components:

one_sample_x
: It contains the results by one_sample(x, ...).

one_sample_y
: It contains the results by one_sample(y, ...).

mu1_mu2_interval
: It contains the results of interval estimation of mu1-mu2.

mu1_mu2_hypothesis
: It contains the results of test of hypothesis of mu1-mu2.

sigma_ratio_interval
: It contains the results of interval estimation of sigma1^2 / sigma2^2.

sigma_ratio_hypothesis
: It contains the results of test of hypothesis of sigma1^2 / sigma2^2.

res.ks
: It contains the results of ks.test(x,y).

res.binom
: It contains the results of binom.test(sum(x<y), length(x)).

res.wilcox
: It contains the results of wilcox.test(x, y, ...).

cor.pearson
: It contains the results of cor.test(x, y, method = "pearson", ...).

cor.kendall
: It contains the results of cor.test(x, y, method = "kendall", ...).

cor.spearman
: It contains the results of cor.test(x, y, method = "spearman", ...).

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
## One sample
x=rnorm(10, mean = 1, sd = 0.2); x

## one_sample(x, ...) == one_two_sample(x, ...)
one_sample(x, mu = 1, sigma = 0.2, side = 1)
one_two_sample(x, mu = 1, sigma = 0.2, side = 1)

one_sample(x, sigma = 0.2, side = 1)
one_two_sample(x, sigma = 0.2, side = 1)

one_sample(x, mu = 1, side = 1)
one_two_sample(x, mu = 1, side = 1)

one_sample(x)
one_two_sample(x)

## Two samples
set.seed(1)
x=rnorm(10, mean = 1, sd = 0.2); x
y=rnorm(20, mean = 2, sd = 0.3); y
y2=rnorm(20, mean = 2, sd = 0.2); y2

## sigma1, sigma2 known; mu1, mu2 known
one_two_sample(x, y, sigma = c(0.2, 0.3), mu = c(1, 2))

## sigma1 = sigma2 unknown; mu1, mu2 known
one_two_sample(x, y2, var.equal = TRUE, mu = c(1, 2))

## sigma1 != sigma2 unknown; mu1, mu2 known
one_two_sample(x, y, mu = c(1, 2))
```

```
## sigma1, sigma2 known; mu1, mu2 unknown
one_two_sample(x, y, sigma = c(0.2, 0.3))

## sigma1 = sigma2 unknown; mu1, mu2 unknown
one_two_sample(x, y2, var.equal = TRUE)

## sigma1 != sigma2 unknown; mu1, mu2 unknown
one_two_sample(x, y)
```

---

p_value *Compute the P value*

Description

Compute the P value of a cumulative distribution function (cdf).

Usage

```
p_value(cdf, x, paramet = numeric(0), side = 0)
```

Arguments

cdf	The cumulative distribution function. For normal distribution, `cdf = pnorm`.
x	A given value to compute the P value.
paramet	The parameter of the corresponding distribution. For normal distribution, `paramet = c(mu, sigma)`.
side	A parameter indicating whether to compute one sided or two sided P value. When inputting `side = -1` (or a number < 0), the function computes a left side P value; when inputting `side = 1` (or a number > 0), the function computes a right side P value; when inputting `side = 0` (default), the function computes a two sided P value.

Value

The P value.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
p_value(pnorm, x = 0, side = 1)
p_value(pt, x = 0, paramet = 5, side = 1)
```

---

var_test1 — *Two sided or one sided test of hypothesis of* `sigma^2` *of one normal sample*

Description

Compute the two sided or one sided test of hypothesis of `sigma^2` of one normal sample when the population mean is known or unknown.

Usage

```
var_test1(x, sigma2 = 1, mu = Inf, side = 0)
```

Arguments

`x`	A numeric vector.
`sigma2`	`sigma2` is `sigma0^2` in the null hypothesis. Default is 1, i.e., `H0: sigma^2 = 1`.
`mu`	The population mean. `mu < Inf` indicates it is known, `mu == Inf` indicates it is unknown. Default to unknown population mean.
`side`	A parameter used to control two sided or one sided test of hypothesis. When inputting `side = 0` (default), the function computes two sided test of hypothesis, and `H1: sigma^2 != sigma0^2`; when inputting `side = -1` (or a number < 0), the function computes one sided test of hypothesis, and `H1: sigma^2 < sigma0^2`; when inputting `side = 1` (or a number > 0), the function computes one sided test of hypothesis, and `H1: sigma^2 > sigma0^2`.

Value

A data.frame with variables:

`var`	The estimate of the population variance. When the population mean `mu` is known, `var = mean((x-mu)^2)`. When `mu` is unknown, `var = var(x)`.
`df`	The degree of freedom.
`chisq2`	The chisquare statistic.
`p_value`	The P value.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
var_test1(x, sigma2 = 0.2^2, mu = 1, side = 1)
var_test1(x, sigma2 = 0.2^2, side = 1)
```

---

var_test2 — *Two sided or one sided test of hypothesis of* `sigma1^2` *and* `sigma2^2` *of two normal samples*

Description

Compute the two sided or one sided test of hypothesis of `sigma1^2` and `sigma2^2` of two normal samples when the population means are known or unknown.

Usage

```
var_test2(x, y, mu = c(Inf, Inf), side = 0)
```

Arguments

- `x` A numeric vector.
- `y` A numeric vector.
- `mu` The population means. When it is known, input it, and the function computes the F value using an F distribution with degree of freedom `(n1, n2)`. When it is unknown, ignore it, and the function computes the F value using an F distribution with degree of freedom `(n1-1, n2-1)`.
- `side` A parameter used to control two sided or one sided test of hypothesis. When inputting `side = 0` (default), the function computes two sided test of hypothesis, and `H1: sigma1^2 != sigma2^2`; when inputting `side = -1` (or a number < 0), the function computes one sided test of hypothesis, and `H1: sigma1^2 < sigma2^2`; when inputting `side = 1` (or a number > 0), the function computes one sided test of hypothesis, and `H1: sigma1^2 > sigma2^2`.

Value

A data.frame with variables:

- `rate` The estimate of the ratio of population variances, `rate = Sx2/Sy2`. When the population means `mu` is known, `Sx2 = 1/n1*sum((x-mu[1])^2)` and `Sy2 = 1/n2*sum((y-mu[2])^2`. When `mu` is unknown, `Sx2 = var(x)` and `Sy2 = var(y)`.
- `df1` The first degree of freedom.

df2	The second degree of freedom.
F	The F statistic.
p_value	The P value.

Author(s)

Ying-Ying Zhang (Robert) <robertzhangyying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion, doi:10.2991/asshm-13.2013.29.

Examples

```
x=rnorm(10, mean = 1, sd = 0.2); x
y=rnorm(20, mean = 2, sd = 0.3); y
var_test2(x, y, mu = c(1, 2), side = 1)
var_test2(x, y, side = 1)
```

Index