

Package ‘OneTwoSamples’

February 19, 2015

Type Package

Title Deal with one and two (normal) samples

Version 1.0-3

Date 2013-05-29

Author Ying-Ying Zhang (Robert)

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

Description In this package, we introduce an R function `one_two_sample()` which can 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 $\mu_1 - \mu_2$ (the difference of the means of x and y) and σ_1^2 / σ_2^2 (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.

License GPL (≥ 2)

Repository CRAN

NeedsCompilation no

Date/Publication 2013-05-29 11:55:34

R topics documented:

OneTwoSamples-package	2
data_outline	3
detail	4
interval_estimate1	5
interval_estimate2	6
interval_estimate3	7
interval_estimate4	8

interval_estimate5	9
interval_var1	10
interval_var2	11
interval_var3	12
interval_var4	13
mean_test1	14
mean_test2	15
one_sample	17
one_two_sample	19
p_value	21
var_test1	22
var_test2	23

Index	25
--------------	-----------

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. 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 $\mu_1 - \mu_2$ (the difference of the means of x and y) and σ_1^2 / σ_2^2 (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: 1.0-3
 Date: 2013-05-29
 License: GPL (>= 2)

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

Author(s)

Ying-Ying Zhang (Robert)

Maintainer: Ying-Ying Zhang (Robert) <robertzhangying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion.

Examples

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

data_outline	<i>Compute various descriptive statistics</i>
--------------	---

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) <robertzhangying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion.

Examples

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

detail	<i>Show details of an object</i>
--------	----------------------------------

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.

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. $\text{sigma} \geq 0$ indicates it is known, $\text{sigma} < 0$ indicates it is unknown. Default to unknown standard deviation.
alpha	The significance level, a real number in $[0, 1]$. Default to 0.05. $1 - \text{alpha}$ 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.

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 $\mu_1 - \mu_2$ of two normal samples*

Description

Compute the two sided interval estimation of $\mu_1 - \mu_2$ 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 $\bar{x} - \bar{y}$.
df	The degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

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

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion.

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. $\text{sigma} \geq 0$ indicates it is known, $\text{sigma} < 0$ indicates it is unknown. Default to unknown standard deviation.
alpha	The significance level, a real number in [0, 1]. Default to 0.05. $1 - \text{alpha}$ 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) <robertzhangying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R function.

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. $\text{sigma} \geq 0$ indicates it is known, $\text{sigma} < 0$ 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 $\text{side} = -1$; when computing the one sided lower limit, input $\text{side} = 1$; when computing the two sided limits, input $\text{side} = 0$ (default).
alpha	The significance level, a real number in $[0, 1]$. Default to 0.05. $1 - \text{alpha}$ 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) <robertzhangying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion.

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	<i>Two sided or one sided interval estimation of $\mu_1 - \mu_2$ of two normal samples</i>
--------------------	---

Description

Compute the two sided or one sided interval estimation of $\mu_1 - \mu_2$ 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 $x_b - y_b$.
df	The degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

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

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R function.

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	<i>Two sided interval estimation of σ^2 of one normal sample</i>
---------------	--

Description

Compute the two sided interval estimation of σ^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, $\text{var} = \text{mean}((x-\mu)^2)$. When mu is unknown, $\text{var} = \text{var}(x)$.
df	The degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

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

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R function.

Examples

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

interval_var2	<i>Two sided interval estimation of σ_1^2 / σ_2^2 of two normal samples</i>
---------------	--

Description

Compute the two sided interval estimation of σ_1^2 / σ_2^2 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.

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	<i>Two sided or one sided interval estimation of σ^2 of one normal sample</i>
---------------	---

Description

Compute the two sided or one sided interval estimation of σ^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 μ is known, $\text{var} = \text{mean}((x-\mu)^2)$. When μ is unknown, $\text{var} = \text{var}(x)$.
df	The degree of freedom.
a	The confidence lower limit.
b	The confidence upper limit.

Author(s)

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

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R function.

Examples

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

interval_var4	<i>Two sided or one sided interval estimation of σ_1^2 / σ_2^2 of two normal samples</i>
---------------	---

Description

Compute the two sided or one sided interval estimation of σ_1^2 / σ_2^2 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 μ is known, $Sx2 = 1/n1*\sum((x-\mu[1])^2)$ and $Sy2 = 1/n2*\sum((y-\mu[2])^2)$. When μ 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) <robertzhangying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion.

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 μ 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 μ_0 in the null hypothesis. Default is 0, i.e., $H_0: \mu = 0$.
sigma	The standard deviation of the population. $\text{sigma} \geq 0$ indicates it is known, $\text{sigma} < 0$ 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 $\text{side} = 0$ (default), the function computes two sided test of hypothesis, and $H_1: \mu \neq \mu_0$; when inputting $\text{side} = -1$ (or a number < 0), the function computes one sided test of hypothesis, and $H_1: \mu < \mu_0$; when inputting $\text{side} = 1$ (or a number > 0), the function computes one sided test of hypothesis, and $H_1: \mu > \mu_0$.

Value

A data.frame with variables:

mean	The sample mean.
df	The degree of freedom.
statistic	The statistic, when $\text{sigma} \geq 0$, $\text{statistic} = Z = (\bar{x} - \mu) / (\text{sigma} / \sqrt{n})$; when $\text{sigma} < 0$, $\text{statistic} = T = (\bar{x} - \mu) / (\text{sd}(x) / \sqrt{n})$.
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.

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	<i>Two sided or one sided test of hypothesis of μ_1 and μ_2 of two normal samples</i>
------------	---

Description

Compute the two sided or one sided test of hypothesis of μ_1 and μ_2 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 <code>var.equal</code> 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 <code>side = 0</code> (default), the function computes two sided test of hypothesis, and $H_1: \mu_1 \neq \mu_2$; when inputting <code>side = -1</code> (or a number < 0), the function computes one sided test of hypothesis, and $H_1: \mu_1 < \mu_2$; when inputting <code>side = 1</code> (or a number > 0), the function computes one sided test of hypothesis, and $H_1: \mu_1 > \mu_2$.

Value

A data.frame with variables:

mean	The difference of sample means $x_b - y_b$.
df	The degree of freedom.
statistic	The statistic, when <code>all(sigma>=0)</code> , <code>statistic = Z</code> ; otherwise, <code>statistic = T</code> .
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.

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

<code>x</code>	A numeric vector.
<code>mu</code>	<p><code>mu</code> plays two roles.</p> <p>In two sided or one sided interval estimation (or test of hypothesis) of σ^2 of one normal sample, <code>mu</code> 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 $n-1$.</p> <p>In two sided or one sided test of hypothesis of μ of one normal sample, <code>mu</code> is μ_0 in the null hypothesis, and $\mu_0 = \text{if } (\mu < \text{Inf}) \mu \text{ else } 0$.</p>
<code>sigma</code>	<p><code>sigma</code> plays two roles.</p> <p>In two sided or one sided interval estimation (or test of hypothesis) of μ of one normal sample, <code>sigma</code> is the standard deviation of the population. <code>sigma >= 0</code> indicates it is known, and the function computes the interval endpoints (or Z statistic) using a standard normal distribution. <code>sigma < 0</code> 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.</p> <p>In two sided or one sided test of hypothesis of σ^2 of one normal sample, <code>sigma</code> is σ_0 in the null hypothesis. Default is 1, i.e., $H_0: \sigma^2 = 1$.</p>
<code>side</code>	<p><code>side</code> plays two roles and is used in four places.</p> <p>In two sided or one sided interval estimation of μ of one normal sample, <code>side</code> is a parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input <code>side = -1</code>; when computing the one sided lower limit, input <code>side = 1</code>; when computing the two sided limits, input <code>side = 0</code> (default).</p> <p>In two sided or one sided interval estimation of σ^2 of one normal sample, <code>side</code> is a parameter used to control whether to compute two sided or one sided interval estimation. When computing the one sided upper limit, input <code>side = -1</code>; when computing the one sided lower limit, input <code>side = 1</code>; when computing the two sided limits, input <code>side = 0</code> (default).</p>

In two sided or one sided test of hypothesis of μ 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 $H_1: \mu \neq \mu_0$; when inputting `side = -1` (or a number < 0), the function computes one sided test of hypothesis, and $H_1: \mu < \mu_0$; when inputting `side = 1` (or a number > 0), the function computes one sided test of hypothesis, and $H_1: \mu > \mu_0$.

In two sided or one sided test of hypothesis of σ^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 $H_1: \sigma^2 \neq \sigma_0^2$; when inputting `side = -1` (or a number < 0), the function computes one sided test of hypothesis, and $H_1: \sigma^2 < \sigma_0^2$; when inputting `side = 1` (or a number > 0), the function computes one sided test of hypothesis, and $H_1: \sigma^2 > \sigma_0^2$.

`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:

<code>mu_interval</code>	It contains the results of interval estimation of μ .
<code>mu_hypothesis</code>	It contains the results of test of hypothesis of μ .
<code>sigma_interval</code>	It contains the results of interval estimation of σ .
<code>sigma_hypothesis</code>	It contains the results of test of hypothesis of σ .

Author(s)

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

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion.

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	<i>Deal with one and two (normal) samples</i>
----------------	---

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 $\mu_1 - \mu_2$ (the difference of the means of x and y) and σ_1^2 / σ_2^2 (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

<code>x</code>	A numeric vector.
<code>y</code>	A numeric vector.
<code>mu</code>	If $y = \text{NULL}$, i.e., there is only one sample. See the argument <code>mu</code> in one_sample . For two normal samples x and y , <code>mu</code> plays one role: the population means. However, <code>mu</code> is used in two places: one is the two sided or one sided interval estimation of σ_1^2 / σ_2^2 of two normal samples, another is the two sided or one sided test of hypothesis of σ_1^2 and σ_2^2 of two normal samples. When <code>mu</code> is known, input it, and the function computes the interval endpoints (or the F value) using an F distribution with degree of freedom (n_1, n_2). 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 ($n_1 - 1, n_2 - 1$).
<code>sigma</code>	If $y = \text{NULL}$, i.e., there is only one sample. See the argument <code>sigma</code> in one_sample . For two normal samples x and y , <code>sigma</code> plays one role: the population standard deviations. However, <code>sigma</code> is used in two places: one is the two sided or one sided interval estimation of $\mu_1 - \mu_2$ of two normal samples, another is the two sided or one sided test of hypothesis of μ_1 and μ_2 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 <code>var.equal</code> below.
<code>var.equal</code>	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.
<code>ratio</code>	The hypothesized ratio of the population variances of x and y . It is used in <code>var.test(x, y, ratio = ratio, ...)</code> , i.e., when computing the interval

estimation and test of hypothesis of σ_1^2 / σ_2^2 when μ_1 or μ_2 is unknown.

side	If $y = \text{NULL}$, i.e., there is only one sample. See the argument side in one_sample . For two normal samples x and y , σ is used in four places: interval estimation of $\mu_1 - \mu_2$, test of hypothesis of μ_1 and μ_2 , interval estimation of σ_1^2 / σ_2^2 , test of hypothesis of σ_1^2 and σ_2^2 . In interval estimation of $\mu_1 - \mu_2$ or σ_1^2 / σ_2^2 , 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 μ_1 and μ_2 or σ_1^2 and σ_2^2 , 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 $H_1: \mu_1 \neq \mu_2$ or $H_1: \sigma_1^2 \neq \sigma_2^2$; when inputting side = -1 (or a number < 0), the function computes one sided test of hypothesis, and $H_1: \mu_1 < \mu_2$ or $H_1: \sigma_1^2 < \sigma_2^2$; when inputting side = 1 (or a number > 0), the function computes one sided test of hypothesis, and $H_1: \mu_1 > \mu_2$ or $H_1: \sigma_1^2 > \sigma_2^2$.
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 <code>one_sample(x, ...)</code> .
one_sample_y	It contains the results by <code>one_sample(y, ...)</code> .
mu1_mu2_interval	It contains the results of interval estimation of $\mu_1 - \mu_2$.
mu1_mu2_hypothesis	It contains the results of test of hypothesis of $\mu_1 - \mu_2$.
sigma_ratio_interval	It contains the results of interval estimation of σ_1^2 / σ_2^2 .
sigma_ratio_hypothesis	It contains the results of test of hypothesis of σ_1^2 / σ_2^2 .
res.ks	It contains the results of <code>ks.test(x,y)</code> .
res.binom	It contains the results of <code>binom.test(sum(x<y), length(x))</code> .
res.wilcox	It contains the results of <code>wilcox.test(x, y, ...)</code> .
cor.pearson	It contains the results of <code>cor.test(x, y, method = "pearson", ...)</code> .
cor.kendall	It contains the results of <code>cor.test(x, y, method = "kendall", ...)</code> .
cor.spearman	It contains the results of <code>cor.test(x, y, method = "spearman", ...)</code> .

Author(s)

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

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R function.

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)
```

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) <robertzhangying@qq.com>

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion.

Examples

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

var_test1	<i>Two sided or one sided test of hypothesis of sigma^2 of one normal sample</i>
-----------	--

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 σ_0^2 in the null hypothesis. Default is 1, i.e., $H_0: \sigma^2 = 1$.
mu	The population mean. $\mu < \text{Inf}$ indicates it is known, $\mu == \text{Inf}$ 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 $\text{side} = 0$ (default), the function computes two sided test of hypothesis, and $H_1: \sigma^2 \neq \sigma_0^2$; when inputting $\text{side} = -1$ (or a number < 0), the function computes one sided test of hypothesis, and $H_1: \sigma^2 < \sigma_0^2$; when inputting $\text{side} = 1$ (or a number > 0), the function computes one sided test of hypothesis, and $H_1: \sigma^2 > \sigma_0^2$.

Value

A data.frame with variables:

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

Author(s)

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

References

Zhang, Y. Y., Wei, Y. (2013), One and two samples using only an R funtion.

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	<i>Two sided or one sided test of hypothesis of σ_1^2 and σ_2^2 of two normal samples</i>
-----------	---

Description

Compute the two sided or one sided test of hypothesis of σ_1^2 and σ_2^2 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 $H_1: \sigma_1^2 \neq \sigma_2^2$; when inputting side = -1 (or a number < 0), the function computes one sided test of hypothesis, and $H_1: \sigma_1^2 < \sigma_2^2$; when inputting side = 1 (or a number > 0), the function computes one sided test of hypothesis, and $H_1: \sigma_1^2 > \sigma_2^2$.

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.

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

*Topic **package**

OneTwoSamples-package, [2](#)

attributes, [4](#)

class, [4](#)

data_outline, [3](#)

detail, [4](#)

interval_estimate1, [5](#)

interval_estimate2, [6](#)

interval_estimate3, [7](#)

interval_estimate4, [8](#)

interval_estimate5, [9](#)

interval_var1, [10](#)

interval_var2, [11](#)

interval_var3, [12](#)

interval_var4, [13](#)

is.object, [4](#)

isS4, [4](#)

mean_test1, [14](#)

mean_test2, [15](#)

one_sample, [17](#), [19](#), [20](#)

one_two_sample, [19](#)

OneTwoSamples (OneTwoSamples-package), [2](#)

OneTwoSamples-package, [2](#)

p_value, [21](#)

var_test1, [22](#)

var_test2, [23](#)