Package ‘RSSampling’

March 19, 2018

Type Package
Title Ranked Set Sampling
Version 1.0
Author Busra Sevinc, Bekir Cetintav, Melek Esemen, Selma Gurler
Maintainer Busra Sevinc <busra.sevincc@gmail.com>
Imports LearnBayes, stats
Description Ranked set sampling (RSS) is introduced as an advanced method for data collection which is substantial for the statistical and methodological analysis in scientific studies by McIntyre (1952) (reprinted in 2005) <doi:10.1198/000313005X54180>. This package introduces the first package that implements the RSS and its modified versions for sampling. With ‘RSSampling’, the researchers can sample with basic RSS and the modified versions, namely, Median RSS, Extreme RSS, Percentile RSS, Balanced groups RSS, Double RSS, L-RSS, Truncation-based RSS, Robust extreme RSS. The ‘RSSampling’ also allows imperfect ranking using an auxiliary variable (concomitant) which is widely used in the real life applications. Applicants can also use this package for parametric and nonparametric inference such as mean, median and variance estimation, regression analysis and some distribution-free tests where the samples are obtained via basic RSS.

License GPL-2
Encoding UTF-8
LazyData true
NeedsCompilation no
Repository CRAN
Date/Publication 2018-03-19 14:51:59 UTC

R topics documented:

con.Mrss .......................................................... 2
con.Rrss .......................................................... 4
con.rss .......................................................... 5
Drss .............................................................. 6
meanRSS ........................................................ 8
Mrss ............................................................. 9
con.Mrss

Description

The Mrss function samples from a target population by using modified ranked set sampling methods. Ranking procedure of X is done by using the concomitant variable Y.

Usage

con.Mrss(X,Y,m,r=1,type="r",sets=FALSE,concomitant=FALSE,p)

Arguments

- **X**: A vector of target population
- **Y**: A vector of concomitant variable from target population
- **m**: Size of units in each set
- **r**: Number of cycles. (By default = 1)
- **type**: type of the modified RSS method. "r" for traditional RSS, "p" for Percentile RSS, "m" for Median RSS, "bg" for Balanced Groups RSS, "e" for Extreme RSS. (By default = "r")
- **sets**: logical; if TRUE, ranked set samples are given with ranked sets (see rankedsets)
- **concomitant**: logical; if TRUE, ranked set sample of concomitant variable is given
- **p**: Value of percentile for Percentile RSS method

Details

X and Y must be vectors and also they should be in same length. Value of percentile (p) must be between 0 and 1.
**Value**

- **corr.coef** the correlation coefficient between X and Y
- **var.of.interest** the sets of X, which are ranked by Y
- **concomitant.var.** the ranked sets of Y
- **sample.x** the obtained ranked set sample of X
- **sample.y** the obtained ranked set sample of Y

**References**


**See Also**

- `mrss`, `rrss`, `drss`, `con.rrss`

**Examples**

```r
library("LearnBayes")
mu=c(1,12,2)
Sigma <- matrix(c(1,2,0,2,5,0.5,0,0.5,3), 3, 3)
x <- rmnorm(10000, mu, Sigma)
xx=as.numeric(x[,1])
xy=as.numeric(x[,3])

## Selecting modified ranked set samples
con.Mrss(xx, xy, m=5, r=3, type="r", concomitant=TRUE, sets=TRUE)  
con.Mrss(xx, xy, m=4, r=7, type="m", concomitant=TRUE, sets=TRUE)  
con.Mrss(xx, xy, m=5, r=2, type="e", concomitant=TRUE, sets=TRUE)  
con.Mrss(xx, xy, m=8, r=3, type="p", concomitant=TRUE, sets=TRUE, p=0.25)  
con.Mrss(xx, xy, m=6, r=5, type="bg", concomitant=TRUE, sets=TRUE)
```
**con.Rrss**

**Selecting a robust ranked set sample with a concomitant variable**

**Description**

The `con.Rrss` function samples from a target population by using robust ranked set sampling methods. Ranking procedure of X is done by using the concomitant variable Y.

**Usage**

```r
con.Rrss(x, y, m, r=1, type="l", sets=FALSE, concomitant=FALSE, alpha)
```

**Arguments**

- **x**: A vector of target population
- **y**: A vector of concomitant variable from target population
- **m**: Size of units in each set
- **r**: Number of cycles. (By default =1)
- **type**: type of the modified RSS method. "l" for L-RSS, "tb" for truncation-based RSS, "re" for robust extreme RSS. (By default ="l")
- **sets**: logical; if TRUE, ranked set sample is given with ranked sets (see `rankedsets`)
- **concomitant**: logical; if TRUE, ranked set sample of concomitant variable is given
- **alpha**: Coefficient of the method

**Details**

X and Y must be vectors and also they should be in same length. Coefficient of the method must be between 0 and 0.5.

**Value**

- **corr.coef**: the correlation coefficient between X and Y
- **var.of.interest**: the sets of X, which are ranked by Y
- **concomitant.var.**: the ranked sets of Y
- **sample.x**: the obtained ranked set sample of X
- **sample.y**: the obtained ranked set sample of Y
References


See Also

mrss, Rrss, Drss, con.Mrss

Examples

```r
library("LearnBayes")
mu <- c(1, 12, 2)
Sigma <- matrix(c(1, 2, 0, 2, 5, 0.5, 0, 0.5, 3), 3, 3)
x <- rmnorm(10000, mu, Sigma)
x <- as.numeric(x[, 1])
y <- as.numeric(x[, 3])

## Selecting robust ranked set samples
con.Rrss(x, y, m = 8, r = 4, type = "r", sets = TRUE, concomitant = TRUE, alpha = 0.3)
con.Rrss(x, y, m = 5, r = 2, type = "re", sets = TRUE, concomitant = TRUE, alpha = 0.2)
con.Rrss(x, y, m = 6, r = 3, type = "tb", sets = TRUE, concomitant = TRUE, alpha = 0.25)
```

Description

The `con.rss` function samples from a target population by using ranked set sampling method. Ranking procedure of X is done by using concomitant variable Y.

Usage

```r
con.rss(X, Y, m = 1, r = 1, sets = FALSE, concomitant = FALSE)
```

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>A vector of interested variable from target population</td>
</tr>
<tr>
<td>Y</td>
<td>A vector of concomitant variable from target population</td>
</tr>
<tr>
<td>m</td>
<td>Size of units in each set</td>
</tr>
<tr>
<td>r</td>
<td>Number of cycles. (Default by = 1)</td>
</tr>
<tr>
<td>sets</td>
<td>logical; if TRUE, ranked set sample is given with ranked sets (see <code>rankedsets</code>)</td>
</tr>
<tr>
<td>concomitant</td>
<td>logical; if TRUE, ranked set sample of concomitant variable is given</td>
</tr>
</tbody>
</table>
Details

X and Y must be vectors and also they should be in same length.

Value

corr.coef the correlation coefficient between X and Y
var.of.interest the sets of X, which are ranked by Y
concomitant.var. the ranked sets of Y
sample.x the obtained ranked set sample of X
sample.y the obtained ranked set sample of Y

References


See Also

rss

Examples

library("LearnBayes")
mu=c(1,12,2)
Sigma <- matrix(c(1,2,0,2,5,0.5,0,0.5,3), 3, 3)
x <- rmnorm(10000, mu, Sigma)
xx=as.numeric(x[,1])
xy=as.numeric(x[,3])
con.rss(xx, xy, m=3, r=4, sets=TRUE, concomitant=TRUE)

Drss Selecting double (classical or modified) ranked set sample

Description

The Drss function samples from a target population by using multi-stage ranked set sampling methods.
**Usage**

```r
Drss(x, m, r=1, type="d", sets=FALSE, p)
```

**Arguments**

- `x`: A vector of target population
- `m`: Size of units in each set
- `r`: Number of cycles. (By default = 1)
- `sets`: logical; if TRUE, ranked set samples are given with ranked sets (see `rankedsets`)
- `type`: type of the modified RSS method. "d" for double RSS, "dm" for double median RSS, "dp" for double percentile RSS, "de" for double extreme RSS. (By default = "d")
- `p`: Value of percentile for double percentile RSS method

**Details**

Target population X must be a vector. Value of percentile (p) must be between 0 and 1.

**Value**

- `sets`: the ranked sets where ranked set sample is chosen from
- `sample`: the obtained ranked set sample of X

**References**


**See Also**

`mrss`, `rrss`, `con.mrss`, `con.rrss`

**Examples**

```r
data=rnorm(10000)
## Selecting a double ranked set sample
Drss(data, m=4, r=3, sets=TRUE)
## Selecting a double median ranked set sample
Drss(data, m=4, r=3, type="dm", sets=TRUE)
```
## Selecting a double extreme ranked set sample

Drss(data, m=4, r=3, type="de", sets=TRUE)

## Selecting a double percentile ranked set sample

Drss(data, m=4, r=3, type="dm", sets=TRUE, p=0.6)

---

### Description

The `meanRSS` function estimates the population mean based on ranked set sampling. Also, it calculates confidence interval, p-value and z-statistics for hypothesis testing.

### Usage

```r
meanRSS(X, m, r, alpha=0.05, alternative="two.sided", mu_0)
```

### Arguments

- `X` is an obtained ranked set sample
- `m` is the size of units in each set
- `r` is the number of cycles
- `alpha` is the alpha value for the confidence interval. (By default = 0.05)
- `alternative` is a character string, one of "greater", "less" or "two.sided". For one sample test, alternative refers to the true mean of the parent population in relation to the hypothesized value `mu_0`
- `mu_0` is the initial value for mean in hypothesis testing formula

### Details

An obtained ranked set sample `X` must be `m` by `r` matrix.

### Value

- `mean` the estimated population mean based on ranked set sampling
- `CI` is a confidence interval for the true mean
- `z.test` the z-statistic for the test
- `p.value` the p-value for the test

### References

Examples

library("LearnBayes")
mu<-c(1,12,2)
Sigma <- matrix(c(1,2,0,2,5,0.5,0,0.5,3), 3, 3)
x <- rmnorm(10000, mu, Sigma)
xx=as.numeric(x[,1])
xy=as.numeric(x[,2])
samplerss=con.Mrss(xx,xy,m=4,r=8,type="r",sets=FALSE,concomitant=FALSE)$sample.x

## mean estimation, confidence interval and hypothesis testing for ranked set sample
meanRSS(samplerss,m=4,r=8,mu_0=1)

Description

The Mrss function samples from a target population by using modified ranked set sampling methods.

Usage

Mrss(X,m,r=1,type="r",sets=FALSE,p)

Arguments

X A vector of target population
m Size of units in each set
r Number of cycles. (By default = 1)
sets logical; if TRUE, ranked set samples are given with ranked sets (see rankedsets)
type type of the modified RSS method. "r" for traditional RSS, "p" for Percentile RSS, "m" for Median RSS, "bg" for Balanced Groups RSS, "e" for Extreme RSS. (By default = "r")
p Value of percentile for Percentile RSS method

Details

Target population X must be a vector.

Value

sets the ranked sets where ranked set sample is chosen from
sample the obtained ranked set sample of X
References


See Also

con.mrss, Rrss, Drss

Examples

data=rnorm(10000,1,1)
## Selecting a median ranked set sample
Mrss(data,m=4,r=5,sets=TRUE,type="m")
## Selecting an extreme ranked set sample
Mrss(data,m=3,r=5,sets=TRUE,type="e")
## Selecting a percentile ranked set sample
Mrss(data,m=4,r=3,sets=TRUE,type="p",p=0.2)
## Selecting a balanced groups ranked set sample
Mrss(data,m=6,r=2,sets=TRUE,type="bg")

mwwutestrss(Mann-Whitney-Wilcoxon test with RSS)

Description

In this function, we introduce the RSS version of the Mann-Whitney-Wilcoxon (MWW) test.

Usage

mwwutestrss(X,Y,m,r,l,n,delta0=0, alpha=0.05, lambda=0.5, alternative="two.sided")
Arguments

- **X**: First obtained ranked set sample
- **Y**: Second obtained ranked set sample
- **m**: Set size which was used while sampling X
- **r**: Cycles size which was used while sampling X
- **l**: Set size which was used while sampling Y
- **n**: Cycles size which was used while sampling Y
- **delta0**: The median value of difference in the null hypothesis. (By Default = 0)
- **alpha**: The significance level (by default = 0.05).
- **lambda**: Constant in the variance formula of the test statistic, see Chen et. al.(2003)
- **alternative**: Character string defining the alternative hypothesis, one of "two.sided", "less" or "greater" (by default = "two.sided")

Details

The test statistics and an approximate confidence intervals are constructed by using the normal approximation. Also note that, we assume that the ranking mechanism in the RSS is consistent. For more details please refer to Chen et. al.(2003, pg. 115-124).

There should be two datasets to compare as "X" and "Y", respectively.

Value

- **medianX**: median value of the first sample
- **medianY**: median value of the second sample
- **MWW.test.mwwUrss**: The value of the Mann-Whitney-Wilcoxon test statistic
- **C.I.**: the confidence interval of the Mann-Whitney-Wilcoxon test statistic
- **z.test**: the z statistic for test
- **p.value**: the p value for the test

References


Examples

```r
library("LearnBayes")
mu=c(1,1.2,2)
Sigma <- matrix(c(1,2,0,2,5,0.5,0,0.5,3), 3, 3)
x <- rmmnorm(10000, mu, Sigma)
x=x=as.numeric(x[,1])
y=as.numeric(x[,2])
samplerss=con.rss(x,y,m=3,r=12,concomitant=TRUE)
sample.x=as.numeric(samplerss$sample.x)
```
The `obsno.mrss` function gives the observation numbers to sample from a target population by using modified ranked set sampling methods. Ranking is done using the concomitant variable `Y`.

**Usage**

```r
obsno.mrss(Y, m, r=1, type="r", p)
```

**Arguments**

- `Y` A vector of concomitant variable from target population
- `m` Size of units in each set
- `r` Number of cycles
- `type` type of the modified RSS method. "r" for traditional RSS, "p" for Percentile RSS, "m" for Median RSS, "bg" for Balanced Groups RSS, "e" for Extreme RSS. Default value is "r"
- `p` Value of percentile for Percentile RSS method

**Details**

Concomitant variable `Y` must be a vector.

**References**


**rankedsets**

**See Also**

`con.Mrss, Mrss.rss`

**Examples**

```r
y=rexp(10000)
## Determining the observation numbers of the units which are chosen to sample

y=rexp(10000)
obsno.Mrss(y, m=3, r=5)
obsno.Mrss(y, m=5, r=6, type="m")
obsno.Mrss(y, m=7, r=3, type="e")
obsno.Mrss(y, m=4, r=5, type="p", p=0.3)
obsno.Mrss(y, m=6, r=2, type="bg")
```

**rankedsets**

*Selecting ranked sets*

**Description**

The `rankedsets` function selects ranked sets from a target population. The selection of units in a set is without replacement, but the sets are selecting with replacement.

**Usage**

```r
rankedsets(X, m, s=m)
```

**Arguments**

- `X` A vector of target population
- `m` Size of units in each set
- `s` Number of sets. (by default = m)

**Details**

Target population X must be a vector.

**Value**

It returns a matrix of ranked sets.

**References**

Examples

data=rnorm(10000,3)
## Creating m by m matrix (a regular cycle)
rankedsets(data,m=5)
## Creating m by s matrix
rankedsets(data,m=3,s=5)

regRSS

Regression estimator based on ranked set sampling

Description

It obtains the regression estimator for mean of interested population based on ranked set sampling.

Usage

regRSS(x,y,mu_y)

Arguments

x: An obtained ranked set sample for interested variable from target population
y: An obtained ranked set sample for concomitant variable from target population
mu_y: The known mean for population Y

Details

In this code, variable X and Y represents interested and concomitant variable, respectively, please note that notation is vice versa in the reference (Yu&Lam(1997)).

X and Y must be in same length.

Value

b: the B coefficient
x_reg: the regression estimate for mean of X based on ranked set sampling

References

Examples

library("LearnBayes")
mu=c(1,12,2)
Sigma <- matrix(c(1,2,0,0,2,5,0,5,0,0.5,0,3, 3, 3)
x <- rmmnorm(10000, mu, Sigma)
x=as.numeric(x[,1])
xy=as.numeric(x[,2])
samplerss=con.rss(xx,xy,m=4,r=8,sets=FALSE,concomitant=TRUE)
sample.x=samplerss$sample.x
sample.y=samplerss$sample.y

regRSS(sample.x,sample.y,mu_Y=mean(xy))

---

Rrss

Selecting a robust ranked set sample

Description

The Rrss function samples from a target population by using robust ranked set sampling methods.

Usage

Rrss(X,m,r=1,type="l",sets=FALSE,alpha)

Arguments

X      A vector of target population
m      Size of units in each set
r      Number of cycles. (By default = 1)
type   type of the modified RSS method. "l" for L-RSS, "tb" for truncation-based RSS, "re" for robust extreme RSS. (By default = "l")
sets   logical; if TRUE, ranked set samples are given with ranked sets (see rankedsets)
alpha  Coefficient of the method

Details

Target population X must be a vector. Coefficient of the method must be between 0 and 0.5.

Value

sets      the ranked sets where ranked set sample is chosen from
sample    the obtained ranked set sample of X
References


See Also

con.Mrss, Rrss, Drss

Examples

```r
data=rexp(10000)
## Selecting L-ranked set sample
Rrss(data, m=8, r=3, sets=TRUE, alpha=0.2)
## Selecting Truncation-based ranked set sample
Rrss(data, m=8, r=3, type="tb", sets=TRUE, alpha=0.1)
## Selecting Robust extreme ranked set sample
Rrss(data, m=8, r=3, type="re", sets=TRUE, alpha=0.4)
```

## Selecting classical ranked set sample

description

The `rss` function samples from a target population by using ranked set sampling method.

Usage

```r
rss(X, m, r=1, sets=FALSE)
```

Arguments

- `X`: A vector of target population
- `m`: Size of units in each set
- `r`: Number of cycles. (By default=1)
- `sets`: logical; if TRUE, ranked set samples are given with ranked sets (see `rankedsets`)

Details

Target population X must be a vector.
sign1testrss

Value

sets randomly chosen ranked sets
sample the obtained ranked set sample of X

References


See Also

con.rss

Examples

data=rnorm(10000,1,3)
## Selecting classical ranked set sample with set size \textit{m} and cycle size \textit{r}
rss(data,m=5,r=3,sets=TRUE)

sign1testrss \hspace{1cm} Sign Test with RSS

Description

It performs the RSS version of the sign test given by Chen et. al.(2003).

Usage

sign1testrss(sampledata,m,r,median0,alpha=0.05,alternative="two.sided")

Arguments

sampledata An obtained ranked set sample
m Number of units in each set (set size)
r Number of cycles
median0 The median value in the null hypothesis
alpha The significance level (by default = 0.05).
alternative Character string defining the alternative hypothesis, one of "two.sided", "less" or "greater" (by default = "two.sided")

Details

The test statistics and an approximate confidence intervals are constructed by using the normal approximation. Also note that, we assume that the ranking mechanism in the RSS is consistent. For more details please refer to Chen et. al.(2003, pg. 103-115).
Value

- median: The median value of the given set
- sign.test.stat: The value of the RSS sign test statistic
- C.I.: the confidence interval for median
- z.test: the z statistic for test
- p.value: the p value for the test

References


Examples

```r
data=rnorm(10000,0,1)
samplerss=as.numeric(rss(data,m=3,r=12))
sign1testrss(samplerss,m=3,r=12,median0=0.5)
```

<table>
<thead>
<tr>
<th>varRSS</th>
<th>Variance estimation based on ranked set sampling</th>
</tr>
</thead>
</table>

Description

The `varRSS` function estimates the variance based on ranked set sampling as types of Stokes or Montip&Sukuman.

Usage

`varRSS(X,m,r,type)`

Arguments

- `X`: An obtained ranked set sample
- `m`: Size of units in each set
- `r`: Number of cycles
- `type`: character string, one of "Stokes" or "Montip".

Details

An obtained ranked set sample X must be m by r matrix. Stokes (1980) showed that estimator for variance is biased. Montip and Sukuman(2003) showed that for one cycle there is no unbiased estimator for variance but for more than one cycle they proposed unbiased estimator for variance.
Value

\texttt{var} \quad \text{the estimated population variance based on ranked set sampling}

References


Examples

```r
data=rnorm(10000,2,1)
samplerss=rss(data,m=4,r=3,sets=FALSE)
## Estimation of variance based on ranked set sample by Stokes
varRSS(samplerss,m=4,r=3,type="Stokes")
## Estimation of variance based on ranked set sample by Montip&Sukuman
varRSS(samplerss,m=4,r=3,type="Montip")
```

\begin{tabular}{ll}
\textbf{wsrtestrss} & \textit{Wilcoxon signed rank test with RSS} \\
\end{tabular}

Description

It performs the RSS version of the Wilcoxon signed rank test given by Chen et. al.(2003).

Usage

```
wsrtestrss(sampledata,m,r,delta0=0,alpha=0.05,alternative="two.sided")
```

Arguments

\begin{itemize}
\item \texttt{sampledata} \quad \text{An obtained ranked set sample}
\item \texttt{m} \quad \text{Number of units in each set (set size)}
\item \texttt{r} \quad \text{Number of cycles}
\item \texttt{delta0} \quad \text{The median value of difference in the null hypothesis}
\item \texttt{alpha} \quad \text{The significance level (by default = 0.05).}
\item \texttt{alternative} \quad \text{Character string defining the alternative hypothesis, one of "two.sided", "less" or "greater" (by default = "two.sided")}
\end{itemize}
Details

The test statistics and an approximate confidence intervals are constructed by using the normal approximation. Also note that, we assume that the ranking mechanism in the RSS is consistent. For more details please refer to Chen et. al.(2003, pg. 124-133).

Value

- median: median value of the sample
- sign.rank.test.stat: The value of the Wilcoxon signed rank test statistic
- z.test: the z statistic for test
- p.value: the p value for the test

References


Examples

```r
library("LearnBayes")
mu=c(1,2,1.2,2)
Sigma <- matrix(c(1,2,0.2,0.5,0.5,0.2,0.5,3,3,3), 3, 3)
x <- rmnorm(10000, mu, Sigma)
xx=as.numeric(x[,1])
xy=as.numeric(x[,2])
samplerss=con.rss(xx,xy,m=3,r=12,concomitant=TRUE)
sample.x=as.numeric(samplerss$sample.x)
sample.y=as.numeric(samplerss$sample.y)
difference=sample.x-sample.y
wsrtestrss(difference,m=3,r=12,delta0=0)
```
Index

con.Mrss, 2, 5, 7, 10, 13, 16
con.Rrss, 3, 4, 4, 7
con.rss, 5, 5, 17

Drss, 3, 5, 6, 6, 10, 16

meanRSS, 8, 8
Mrss, 2, 3, 5, 7, 9, 9, 13
mwwutestrss, 10

obsno.Mrss, 12, 12

rankedsets, 2, 4, 5, 7, 9, 13, 13, 15, 16
regRSS, 14
Rrss, 3, 5, 7, 10, 15, 15, 16
rss, 6, 13, 16, 16

sign1testrss, 17

varRSS, 18, 18

wsrtestrss, 19