Package 'ProbSamplingI'

January 20, 2025

Type Package

Title Probabilistic Sampling Design and Strategies

Version 2.0

Date 2024-01-31

Description It allows the user to determine sample sizes, select probabilistic samples, make estimates of different parameters for the total finite population and in studio domains, using the main design drawings.

License GPL (>= 2)

Encoding UTF-8

RoxygenNote 7.2.1

NeedsCompilation no

Author Jorge Barón [aut, cre, cph], Guillermo Martínez [aut]

Maintainer Jorge Barón <jorgeabaron@correo.unicordoba.edu.co>

Repository CRAN

Date/Publication 2024-01-31 22:10:02 UTC

Contents

`ProbSamplingI-package` *Design and Sampling Strategies for Parameter Estimation and Sample Size Determination.*

Description

This package provides functions for selecting a sample and estimating parameters such as total, mean, proportion and ratio; through the main sampling designs.

Details

Index of help topics:

BER	Bernoulli Sampling Design
CONGL	Conglomerate Sampling
ESTRAT	Stratified Sampling
M.MET	Multi-Stage Sampling
MAS	Simple Random Sampling Design without Replacement
MCR	Simple Random Sampling Design with Replacement
PPT	Sampling Design with Replacement and Size Proportional Selection Probabilities
PiPT	Sampling Design without Replacement with Proportional Inclusion Probabilities for Sizes
ProbSamplingI-package	Design and Sampling Strategies for Parameter Estimation and Sample Size Determination.
R.SIS	R-Systematic Sampling Design
WHICH1	Positions of the components of a vector with respect to another vector
n.ESTMAS	Sample Size Through Stratified Sampling
n.MAS	Sample Size Using Simple Random Sampling Design Without Replacement
n.MASC	Sample size using simple random sampling design without conglomerate replacement.

Application of probabilistic sampling

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

Maintainer: Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

BER *Bernoulli Sampling Design*

Description

The BER function selects a random sample or estimates an interest parameter under a Bernoulli design.

Usage

```
BER(N,Pi,yk=NULL,zk=NULL,dk=NULL,type="selec",parameter="total",
    Nc=0.95,Ek=NULL)
# To selectionar: BER(N,Pi)
# To estimate: BER(yk,Pi,type="estm",parameter="total")
# To estimate in domains: BER(yk,Pi,type="estm.Ud",parameter="total")
```

Arguments

N	Size of the population.
Pi	Probability of inclusion.
yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.
dk	Factor that indicates the individuals that belong to each domain of interest, Only needed if "type" is equal to "estm.Ud".
type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If type is equal to "select" the function will make a selection, if it is equal to "estm" the function will perform the estimation of the indicated parameter and if it is equal to "estm.Ud" it will make an estimate in domain.
parameter	This argument indicates the parameter to be estimated ("total", "mean", "prop", or "ratio").
Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case the type is equal to "estm" or "estm.Ud".
Ek	Vector of random numbers of length equal to the size of the population. This argument is optional and by default the function generates them from a uniform distribution (0,1).

Value

This function returns two types of results using the Bernoulli sampling design, depending on the "type" argument, which indicates whether you want to select a sample ("select") or estimate a parameter ("estm" or "estm.Ud").

If type="select", the function returns a list with two elements:

`Ksel`	Vector with the positions of the selected individuals.
`nk`	Selected sample size.

If type="estm" or type="estm.Ud", the function returns a data frame with the estimation of the parameter of interest, the estimated variance of the estimator, the standard error, the coefficient of variation (in percentage), a confidence interval and the design effect.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
yk<-rnorm(100,10,2)
zk<-rnorm(100,10,2)
yk.p<-as.factor(ifelse(yk>10,1,0))


selection<-BER(N=100,Pi=0.3,type="selec")
BER(yk=yk[selection$Ksel],Pi=0.3,type="estm",parameter="total")
BER(Pi=0.3,yk=yk[selection$Ksel],type="estm",parameter="mean")
BER(yk=yk.p[selection$Ksel],Pi=0.3,type="estm",parameter="prop")
BER(yk=yk[selection$Ksel],zk=zk[selection$Ksel],Pi=0.3,
 type="estm",parameter="ratio")

# Domain Estimates

#Sex<-sample(2,100,replace=T)
Sex<-rep(1:2,each=50)
dk<-factor(Sex,labels=c("Man","Woman"))

BER(yk=yk[selection$Ksel],dk=dk[selection$Ksel],Pi=0.3,
 type="estm.Ud",parameter="total")
BER(yk=yk[selection$Ksel],dk=dk[selection$Ksel],Pi=0.3,
 type="estm.Ud",parameter="mean")
```

```
BER(yk=yk.p[selection$Ksel],dk=dk[selection$Ksel],Pi=0.3,
 type="estm.Ud",parameter="prop")
BER(yk=yk[selection$Ksel],zk=zk[selection$Ksel],
 dk=dk[selection$Ksel],Pi=0.3,type="estm.Ud",parameter="ratio")
```

CONGL *Conglomerate Sampling*

Description

The CONGL function selects a random sample or estimates a parameter of interest under a cluster sampling design

Usage

```
CONGL(Argt,cong,design="MAS",type="selec",parameter="total",yk=NULL,
 zk=NULL,dk=NULL,Ek=NULL,Nc=0.95)

# To select: CONGL(Argt=Argt,design)
# To estimate: CONGL(yk,cong,Argt,design,type="estm")
# To estimate in domains: CONGL(yk,dk,cong,Argt,design,type="estm.Ud")

# If the objective is to select a sample, the Argt argument is constructed as follows:

# "MAS": Argt<-list(NI,nI)
# "MCR": Argt<-list(NI,mI)
# "BER": Argt<-list(NI,PiI)
# "PPT": Argt<-list(txkI,mI)
# "PiPT": Argt<-list(txkI,nI)

# If the objective is to estimate a parameter of interest, the Argt argument is
# constructed as follows:

# "MAS": Argt<-list(NI,nI)
# "MCR": Argt<-list(NI,mI)
# "BER": Argt<-list(NI,PiI)
# "PPT": Argt<-list(pkI)
# "PiPT": Argt<-list(pikI,mpiklI)
```

Arguments

yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate.
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.

dk	Factor that indicates the individuals that belong to each domain of interest, Only needed if "type"" is equal to "estm.Ud".
type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If type is equal to "select" the function will make a selection, if it is equal to "estm" the function will make the estimation of the indicated parameter and if it is equal to "estm.Ud" it will make an estimate in domain.
Argt	List with the necessary arguments to select or estimate by the design that you want to use.
cong	Vector indicating which cluster each individual belongs to.
parameter	This argument indicates the parameter to be estimated ("total", "mean", "prop", or "ratio").
design	Sampling sampling design to be implemented ("BER", "MAS", "MCR", "PPT", "SIS" or "PiPT").
Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case "type" is equal to "estm" or "estm.Ud".
Ek	Vector of random numbers of length equal to the size of the population. This argument is optional and by default the function generates them from a uniform distribution (0,1).

Value

This function returns two types of results under the cluster sampling design, depending on the "type" argument, which indicates whether to select a sample ("select") or to estimate an interest parameter ("estm", "estm.Ud"). The results obtained in each case depend on the design implemented, in this way, such results are the same ones obtained for the case of element sampling, but nevertheless in the estimation of the total the intra-sample rate of variance is appended (IVI).

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
yk<-rnorm(120,10,2)
zk<-rnorm(120,12,2)
yk.p<-as.factor(ifelse(yk>10,1,0))
cong<-rep(1:12,each=10);cong
```

```
Sex<-rep(1:2,each=60)
dk<-factor(Sex,labels=c("Man","Woman"))
tyi<-tapply(yk,cong,sum)
txkI<-runif(12,0.95,1.1)*tyi
cor(tyi,txkI)
D1<-data.frame(cong,yk,yk.p,zk,dk)


# MAS-CONGLOMERATE

Argt<-list(NI=12,nI=3)
selection<-CONGL(Argt=Argt,design="MAS")
D.sel<-D1[WHICH1(selection$Ksel,cong),]
CONGL(yk=D.sel$yk,cong=D.sel$cong,Argt=Argt,design="MAS",type="estm")
CONGL(yk=D.sel$yk,cong=D.sel$cong,Argt=Argt,design="MAS",type="estm",
      parameter="mean")
CONGL(yk=D.sel$yk,zk=D.sel$zk,cong=D.sel$cong,Argt=Argt,design="MAS",
      type="estm",parameter="ratio")
CONGL(yk=D.sel$yk.p,cong=D.sel$cong,Argt=Argt,design="MAS",type="estm",
      parameter="prop")

#MCR-CONGLOMERATE

Argt<-list(NI=10,mI=3)
selection<-CONGL(Argt=Argt,design="MCR")
D.sel<-D1[WHICH1(selection$Ksel,cong),]
Ni<-table(cong)[selection$Ksel]
cong.s<-rep(1:3,Ni)
CONGL(yk=D.sel$yk,cong=cong.s,Argt=Argt,design="MCR",type="estm")
CONGL(yk=D.sel$yk,cong=cong.s,Argt=Argt,design="MCR",type="estm",parameter="mean")
CONGL(yk=D.sel$yk,zk=D.sel$zk,cong=cong.s,Argt=Argt,design="MCR",type="estm",
       parameter="ratio")
CONGL(yk=D.sel$yk.p,cong=cong.s,Argt=Argt,design="MCR",type="estm",parameter="prop")

#BER-CONGLOMERATE

Argt<-list(NI=10,PiI=0.4)
selection<-CONGL(Argt=Argt,design="BER")
D.sel<-D1[WHICH1(selection$Ksel,cong),]
CONGL(yk=D.sel$yk,cong=D.sel$cong,Argt=Argt,design="BER",type="estm")
CONGL(yk=D.sel$yk,cong=D.sel$cong,Argt=Argt,design="BER",type="estm",
      parameter="mean")
CONGL(yk=D.sel$yk,zk=D.sel$zk,cong=D.sel$cong,Argt=Argt,design="BER",
      type="estm",parameter="ratio")
CONGL(yk=D.sel$yk.p,cong=D.sel$cong,Argt=Argt,design="BER",type="estm",
      parameter="prop")

#PPT-CONGLOMERATE

Argt<-list(txkI=txkI,mI=4)
selection<-CONGL(Argt=Argt,design="PPT") ;selection
Argt<-list(pkI=selection$pksel)
D.sel<-D1[WHICH1(selection$Ksel,cong),]
```

```
Ni<-table(cong)[selection$Ksel]
cong.s<-rep(1:4,Ni)
CONGL(yk=D.sel$yk,cong=cong.s,Argt=Argt,design="PPT",type="estm")
CONGL(yk=D.sel$yk,cong=cong.s,Argt=Argt,design="PPT",type="estm",parameter="mean")
CONGL(yk=D.sel$yk,zk=D.sel$zk,cong=cong.s,Argt=Argt,design="PPT",type="estm",
      parameter="ratio")
CONGL(yk=D.sel$yk.p,cong=cong.s,Argt=Argt,design="PPT",type="estm",parameter="prop")


#PiPT-CONGLOMERATE

Argt<-list(txkI=txkI,nI=4)
selection<-CONGL(Argt=Argt,design="PiPT")
Argt<-list(pikI=selection$piksel,mpiklI=selection$mpikl.s)
D.sel<-D1[WHICH1(selection$Ksel,cong),]
CONGL(yk=D.sel$yk,cong=D.sel$cong,Argt=Argt,design="PiPT",type="estm")
CONGL(yk=D.sel$yk,cong=D.sel$cong,Argt=Argt,design="PiPT",type="estm",
       parameter="mean")
CONGL(yk=D.sel$yk,zk=D.sel$zk,cong=D.sel$cong,Argt=Argt,design="PiPT",
       type="estm",parameter="ratio")
CONGL(yk=D.sel$yk.p,cong=D.sel$cong,Argt=Argt,design="PiPT",type="estm",
      parameter="prop")


# Domain Estimate
# MAS-CONGLOMERATE

Argt<-list(NI=12,nI=3)
selection<-CONGL(Argt=Argt,design="MAS")
D.sel<-D1[WHICH1(selection$Ksel,cong),]
CONGL(yk=D.sel$yk,dk=D.sel$dk,cong=D.sel$cong,Argt=Argt,
      design="MAS",type="estm.Ud")
CONGL(yk=D.sel$yk,dk=D.sel$dk,cong=D.sel$cong,Argt=Argt,
      design="MAS",type="estm.Ud",parameter="mean")
CONGL(yk=D.sel$yk,zk=D.sel$zk,dk=D.sel$dk,cong=D.sel$cong,Argt=Argt,
       design="MAS",type="estm.Ud",parameter="ratio")
CONGL(yk=D.sel$yk.p,dk=D.sel$dk,cong=D.sel$cong,Argt=Argt,
      design="MAS",type="estm.Ud",parameter="prop")

# Domain Estimate
# MCR-CONGLOMERATE

Argt<-list(NI=10,mI=3)
selection<-CONGL(Argt=Argt,design="MCR")
D.sel<-D1[WHICH1(selection$Ksel,cong),]
Ni<-table(cong)[selection$Ksel]
cong.s<-rep(1:3,Ni)
CONGL(yk=D.sel$yk,dk=D.sel$dk,cong=cong.s,Argt=Argt,
      design="MCR",type="estm.Ud")
CONGL(yk=D.sel$yk,dk=D.sel$dk,cong=cong.s,Argt=Argt,design="MCR",
      type="estm.Ud",parameter="mean")
CONGL(yk=D.sel$yk,zk=D.sel$zk,dk=D.sel$dk,cong=cong.s,Argt=Argt,
      design="MCR",type="estm.Ud",parameter="ratio")
```

```
CONGL(yk=D.sel$yk.p,dk=D.sel$dk,cong=cong.s,Argt=Argt,design="MCR",
       type="estm.Ud",parameter="prop")

# Domain Estimate
# BER-CONGLOMERATE

Argt<-list(NI=10,PiI=0.4)
selection<-CONGL(Argt=Argt,design="BER")
D.sel<-D1[WHICH1(selection$Ksel,cong),]
CONGL(yk=D.sel$yk,dk=D.sel$dk,cong=D.sel$cong,Argt=Argt,
      design="BER",type="estm.Ud")
CONGL(yk=D.sel$yk,dk=D.sel$dk,cong=D.sel$cong,Argt=Argt,
      design="BER",type="estm.Ud",parameter="mean")
CONGL(yk=D.sel$yk,dk=D.sel$dk,zk=D.sel$zk,cong=D.sel$cong,Argt=Argt,
      design="BER",type="estm.Ud",parameter="ratio")
CONGL(yk=D.sel$yk.p,dk=D.sel$dk,cong=D.sel$cong,Argt=Argt,
      design="BER",type="estm.Ud",parameter="prop")
```

ESTRAT *Stratified Sampling*

Description

The ESTRAT function selects a random sample or estimates an interest parameter under a stratified sampling.

Usage

```
ESTRAT(strata,designs,nh,xk=NULL,yk=NULL,zk=NULL,dk=NULL,type="selec",
       Argt,parameter="total",rh=NULL,Ek=NULL,Nc=0.95)
# To select: ESTRAT(strata,nh,designs,xk,rh)
# To estimate: ESTRAT(yk,zk,strata,designs,type="estm",Argt,parameter)
# To estimate in domains: ESTRAT(yk,zk,dk,strata,designs,type="estm",Argt,parameter)
```

Arguments

strata	Vector indicating which stratum each individual belongs to.
nh	Vector that indicates the number of individuals to select in each stratum. This argument is required if the type argument is equal to "select".
yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate.
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.

xk	Vector of observations of the auxiliary variable. This vector is only necessary if it is desired to select in any stratum by means of a probability selection or inclusion probability proportional to size design.
dk	Factor that indicates the individuals that belong to each domain of interest, is only necessary if type is equal to "estm.Ud".
type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If type is equal to "select" the function will make a selection, if it is equal to "estm" the function will perform the estimation of the indicated parameter and if it is equal to "estm.Ud" will make an estimate in domain.
designs	Vector indicating the design to be used in each stratum ("BER", "MAS", "MCR", "PPT", "SIS" or "PiPT").
parameter	This argument indicates the parameter to be estimated ("total", "average", "prop" or "reason").
Argt	It is a list with the necessary arguments for the estimates under the respective designs used in the strata.
rh	Vector of size equal to the number of strata, necessary if it is desired to select under an r-sistematic design, which will have the number of starts to be used in the corresponding strata and zeros in the rest of the positions where this design is not used.
Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case the type is equal to "estm" or "estm.Ud".
Ek	Vector of random numbers of length equal to population size. This argument is optional and by default the function generates them from a uniform distribution (0,1).

Value

This function returns two types of results under the stratified sampling design depending on the "type" argument, which indicates whether to select ("select") or estimate ("estm", "estm.Ud"). If type is equal to "select" the function returns a list with two elements, the first is a data frame (Sample) in which one of its columns indicates the position of the selected individuals in each stratum and the second (Rtdos.h) is a list with the results obtained in each stratum which are necessary when making a certain estimate. If type is equal to "est" or "estm.Ud", the function returns a list with two data frames with the estimation of the parameter of interest, the estimated variance of the estimator, the standard error, the coefficient of variation (in percentage) and a confidence interval assuming normality; by stratum and in general.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
yk<-rnorm(1000,10,2)
xk<-rnorm(1000,10,3)
zk<-rnorm(1000,12,3)
yk.p<-factor(ifelse(yk>10,"A","B"))
strata<-rep(1:5,each=200)
Sex<-rep(1:2,length=1000)
dk<-factor(Sex,labels=c("Man","Woman"))


nh<-c(60,40,40,60,80)
designs<-c("MAS","MAS","MAS","MAS","MAS")
select<-ESTRAT(strata=strata,designs=designs,nh=nh)
Argt<-select$Rtdos.h
Strata<-strata[select$Sample$IND]
yksel<-yk[select$Sample$IND]
yk.psel<-as.factor(yk.p[select$Sample$IND])
zksel<-zk[select$Sample$IND]
ESTRAT(yk=yksel,strata=Strata,designs=designs,Argt=Argt,
       type="estm",parameter="total")
ESTRAT(yk=yksel,strata=Strata,designs=designs,Argt=Argt,
       type="estm",parameter="mean")
ESTRAT(yk=yk.psel,strata=Strata,designs=designs,Argt=Argt,
      type="estm",parameter="prop")
ESTRAT(yk=yksel,zk=zksel,strata=Strata,designs=designs,Argt=Argt,
       type="estm",parameter="ratio")


designs<-c("PiPT","PPT","MAS","MCR","BER")
select<-ESTRAT(xk=xk,strata=strata,designs=designs,nh)
Argt<-select$Rtdos.h
Strata<-strata[select$Sample$IND]
yksel<-yk[select$Sample$IND]
yk.psel<-yk.p[select$Sample$IND]
zksel<-zk[select$Sample$IND]
ESTRAT(yk=yksel,strata=Strata,designs=designs,Argt=Argt,
       type="estm",parameter="total")
ESTRAT(yk=yk.psel,strata=Strata,designs=designs,Argt=Argt,
      type="estm",parameter="prop")
ESTRAT(yk=yksel,strata=Strata,designs=designs,Argt=Argt,
       type="estm",parameter="mean")
ESTRAT(yk=yksel,zk=zksel,strata=Strata,designs=designs,Argt=Argt,
       type="estm",parameter="ratio")

# Estimates in Domains

designs<-c("MAS","MAS","MAS","MAS","MAS")
```

```
select<-ESTRAT(strata=strata,designs=designs,nh=nh)
Argt<-select$Rtdos.h
Strata<-strata[select$Sample$IND]
yksel<-yk[select$Sample$IND]
yk.psel<-yk.p[select$Sample$IND]
zksel<-zk[select$Sample$IND]
dksel<-dk[select$Sample$IND]
ESTRAT(yk=yksel,strata=Strata,dk=dksel,designs=designs,Argt=Argt,
        type="estm.Ud",parameter="total")
ESTRAT(yk=yksel,strata=Strata,dk=dksel,designs=designs,Argt=Argt,
      type="estm.Ud",parameter="mean")
ESTRAT(yk=yk.psel,strata=Strata,dk=dksel,designs=designs,Argt=Argt,
       type="estm.Ud",parameter="prop")
ESTRAT(yk=yksel,zk=zksel,strata=Strata,dk=dksel,designs=designs,
       Argt=Argt,type="estm.Ud",parameter="ratio")
```

M.MET *Multi-Stage Sampling*

Description

The M.MET function selects a random sample or estimates an interest parameter under multi-stage sampling (up to four stages).

Usage

```
M.MET(F.UM,designs,list.arg,p,type="selec",parameter="total",yk=NULL,
      zk=NULL,xk=NULL,dk=NULL,r=NULL,Nc=0.95)
# To select: M.MET(F.UM=F.UM,p=p,designs)
# To estimate: M.MET(yk,F.UM,p,designs,list.arg,type="estm",parameter)
```

Arguments

yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate.
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.
xk	Vector of observations of the auxiliary variable. This vector is only necessary if you want to select using a layout that uses an auxiliary variable.
dk	Factor that indicates the individuals that belong to each domain of interest, Only needed if type is equal to "estm.Ud".
F.UM	Data.frame that contains columns indicating which sampling unit each individual belongs to within each stage.
p	Vector indicating the proportion of individuals to be selected at each sampling stage. This argument is necessary if the type is equal to "select".

designs	Vector indicating the design to be used in each stage ("BER", "MAS", "MCR", "R.SIS", "PPT", "PiPT").
list.arg	List of arguments required for the estimate
r	Number of starts, this argument is only necessary if a r-systematic design is used in the last step.
type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If type is equal to "select" the function will make a selection, if it is equal to "estm" the function will perform the estimation of the indicated parameter and if it is equal to "estm.Ud" will make an estimate in domain.
parameter	This argument indicates the parameter to be estimated ("total", "mean", "prop" or "ratio").
Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case the type is equal to "estm" or "estm.Ud".

Value

This function returns two types of results through the multi-stage sampling strategy that needs to be implemented, depending on the "type" argument, which indicates whether you want to select a sample ("select") or estimate a parameter ("estm" or "estm.Ud").

-If type="select", the function will return a list with two elements:

Sample	Data frame with the location of the selected individuals
Results	List with the results obtained in each stage, which are necessary when making a certain estimate.

-If type = "estm" or type = "estm.Ud", the function returns a data frame with the estimation of the parameter of interest, the estimated variance of the estimator, the standard error, the coefficient of variation (in percent) and a confidence interval.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
#Selection and estimation using a 4-stage sampling

F.UPM<-rep(1:5,each=1000)
F.USM<-rep(1:5,each=200,length=5000)
F.UTM<-rep(1:10,each=20,length=5000)
F.UCM<-rep(1:20,length=5000)
F.UM<-data.frame(F.UPM,F.USM,F.UTM,F.UCM)
p<-c(0.3,0.3,0.3,0.2)
y<-rnorm(5000,10,2)
z<-rnorm(5000,12,2)
y.p<-as.factor(ifelse(y>10,"A","B"))
Sex<-rep(1:2,length=5000)
d<-factor(Sex,labels=c("Man","Woman"))

designs<-c("MAS","MAS","MAS","MAS")
select<-M.MET(F.UM=F.UM,p=p,designs=designs)
F.UM.s<-select$Sample[6:8]
yk<-y[select$Sample$IND]
yk.p<-y.p[select$Sample$IND]
zk<-z[select$Sample$IND]
dk<-d[select$Sample$IND]
list<-select$Results
M.MET(yk=yk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm",parameter="total")
M.MET(yk=yk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm",parameter="mean")
M.MET(yk=yk.p,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm",parameter="prop")
M.MET(yk=yk,zk=zk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm",parameter="ratio")
M.MET(yk=yk,dk=dk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm.Ud",parameter="total")
M.MET(yk=yk,dk=dk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm.Ud",parameter="mean")
M.MET(yk=yk.p,dk=dk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm.Ud",parameter="prop")
M.MET(yk=yk,zk=zk,dk=dk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
       type="estm.Ud",parameter="ratio")


xk<-rnorm(5000,10,2)
designs<-c("PiPT","MAS","PiPT","MAS")
select2<-M.MET(xk=xk,F.UM=F.UM,p=p,designs=designs)
F.UM.s<-select2$Sample[6:8]
yk<-y[select2$Sample$IND]
yk.p<-y.p[select2$Sample$IND]
zk<-z[select2$Sample$IND]
dk<-d[select2$Sample$IND]
list<-select2$Results
M.MET(yk=yk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm",parameter="total")
```

```
M.MET(yk=yk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm",parameter="mean")
M.MET(yk=yk.p,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm",parameter="prop")
M.MET(yk=yk,zk=zk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
       type="estm",parameter="ratio")
M.MET(yk=yk,dk=dk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
       type="estm.Ud",parameter="total")
M.MET(yk=yk,dk=dk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
       type="estm.Ud",parameter="mean")
M.MET(yk=yk.p,dk=dk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm.Ud",parameter="prop")
M.MET(yk=yk,zk=zk,dk=dk,F.UM=F.UM.s,p=p,designs=designs,list.arg=list,
      type="estm.Ud",parameter="ratio")
```

MAS *Simple Random Sampling Design without Replacement*

Description

The MAS function selects a random sample or estimates a parameter of interest under a simple random sampling design without replacement.

Usage

```
MAS(N,n,yk=NULL,zk=NULL,dk=NULL,type="selec",method="fmuller",
              parameter="total",Nc=0.95,Ek=NULL)
# To select: MAS(N,n,method="fmuller")
# To estimate: MAS(yk,N,n,type="estm",parameter="total")
# To estimate in domains: MAS(yk,dk,N,n,type="estm.Ud",parameter="total")
```

Arguments

N	Size of the population
n	Sample size.
yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate.
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.
dk	Factor that indicates the individuals that belong to each domain of interest, Only needed if "type"" is equal to "estm.Ud".
type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If type is equal to "select" the function will make a selection, if it is equal to "estm" the function will make the estimation of the indicated parameter and if it is equal to "estm.Ud" it will make an estimate in domain.

method	Indicates the method or selection mechanism. If Method is equal to "fmuller" the function uses the Fan-Muller method or if it is equal to "cnegative" the function uses the negative coordinate method.
parameter	This argument indicates the parameter to be estimated ("total", "mean", "prop", or "ratio").
Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case "type" is equal to "estm" or "estm.Ud".
Ek	Vector of random numbers of length equal to the size of the population. This argument is optional and by default the function generates them from a uniform distribution (0,1).

Value

This function returns two types of results using the simple random sample design without replacement depending on the "type" argument, which indicates whether to select a sample ("select") or to estimate a parameter ("estm" or "estm.Ud").

If type="select", the function returns a list with a vector (Ksel) with the selected individuals' positions.

If type="estm" or type="estm.Ud", the function returns a data frame with the estimation of the parameter of interest, the estimated variance of the estimator, the standard error, the coefficient of variation (in percent) and an interval of trust.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martinez Florez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
zk<-rnorm(200,15,2)
yk<-rnorm(200,10,3)
yk.p<-as.factor(ifelse(yk>10,"A","B"))
Sex<-rep(1:2,length=200)
dk<-factor(Sex,labels=c("Man","Woman"))
selection<-MAS(N=200,n=40,type="selec",method="fmuller")
MAS(N=200,n=40,type="selec",method="cnegativo")

MAS(yk=yk[selection$K],N=200,n=40,type="estm",parameter="total")
MAS(yk=yk[selection$K],N=200,n=40,type="estm",parameter="mean")
MAS(yk=yk.p[selection$K],N=200,n=40,type="estm",parameter="prop")
```

```
MAS(yk=yk[selection$K],zk=zk[selection$K],N=200,n=40,type="estm",
    parameter="ratio")

# Domain Estimate

MAS(yk=yk[selection$K],dk=dk[selection$K],N=200,n=40,type="estm.Ud",
    parameter="total")
MAS(yk=yk[selection$K],dk=dk[selection$K],N=200,n=40,type="estm.Ud",
    parameter="mean")
MAS(yk=yk.p[selection$K],dk=dk[selection$K],N=200,n=40,type="estm.Ud",
    parameter="prop")
MAS(yk=yk[selection$K],zk=zk[selection$K],dk=dk[selection$K],N=200,n=40,
    type="estm.Ud",parameter="ratio")
```

MCR *Simple Random Sampling Design with Replacement*

Description

The MCR function selects a random sample or estimates an interest parameter under a simple random sampling design without replacement.

Usage

```
MCR(N,m,yk=NULL,zk=NULL,dk=NULL,type="selec",parameter="total",
    Ek=NULL,Nc=0.95)
# To select: MCR(N,m)
# To estimate: MCR(yk,N,m,type="estm",parameter)
# To domain estimate: MCR(yk,dk,N,m,type="est.Ud",parameter)
```

Arguments

N	Size of the population.
m	Sample size.
yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate.
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.
dk	Factor that indicates the individuals that belong to each domain of interest, Only needed if type is equal to "estm.Ud".
type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If type is equal to "select" the function will make a selection, if it is equal to "estm" the function will make the estimation of the indicated parameter and if it is equal to "estm.Ud" it will make an estimate in domain.
parameter	This argument indicates the parameter to be estimated ("total", "mean", "prop", or "ratio").

Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case the type is equal to "estm" or "estm.Ud".
Ek	Vector of random numbers of length equal to the size of the population. This argument is optional and by default the function generates them from a uniform distribution (0,1).

Value

This function returns two types of results using the simple random sample design with replacement, depending on the "type" argument with which it is indicated to select a sample ("select") or to estimate a parameter ("estm" or "estm.Ud").

If type="select", the function returns a list with two elements:

Ksel	Vector with the positions of the selected individuals.
pksel	Vector with the probabilities of selection of individuals.

If type="estm" or type="estm.Ud", the function returns a data frame with the estimation of the parameter of interest, the estimated variance of the estimator, the standard error, the coefficient of variation (in percentage), a confidence interval and the design effect.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
yk<-rnorm(200,10,2)
zk<-rnorm(200,15,3)
yk.p<-as.factor(ifelse(yk>10,1,0))
selection<-MCR(N=200,m=40)
MCR(yk=yk[selection$Ksel],N=200,m=40,type="estm",parameter="total")
MCR(yk=yk[selection$Ksel],N=200,m=40,type="estm",parameter="mean")
MCR(yk=yk.p[selection$Ksel],N=200,m=40,type="estm",parameter="prop")
MCR(yk=yk[selection$Ksel],zk=zk[selection$Ksel],N=200,m=40,
     type="estm",parameter="ratio")

# Domain Estimate

Sex<-rep(1:2,length=200)
dk<-factor(Sex,labels=c("Man","Woman"))
```

```
MCR(yk=yk[selection$K],dk=dk[selection$K],N=200,m=40,type="estm.Ud")
MCR(yk=yk[selection$K],dk=dk[selection$K],N=200,m=40,type="estm.Ud",
    parameter="mean")
MCR(yk=yk.p[selection$Ksel],dk=dk[selection$K],N=200,m=40,
    type="estm.Ud",parameter="prop")
MCR(yk=yk[selection$Ksel],zk=zk[selection$Ksel],dk=dk[selection$K],
    N=100,m=40,type="estm.Ud",parameter="ratio")
```

---

n.ESTMAS *Sample Size Through Stratified Sampling*

Description

The n.ESTMAS function determines the sample size with its corresponding allocation by stratum, using a stratified sampling strategy, where a simple random sampling design with no replacement (ESTMAS) is applied in each stratum; taking into account whether the parameter of interest is the average (or total) or a proportion.

Usage

```
n.ESTMAS(Nh,Sh,Ch,Ph,Emax.a,Nc=0.95,parameter="mean",Asig="Optima")

# n.ESTMAS(Nh,Sh,Ch,Emax.a,Nc=0.95,parameter="mean",Asig="Optima")
# n.ESTMAS(Nh,Ph,Ch,Emax.a,Nc=0.95,parameter="prop",Asig="Optima")

# n.ESTMAS(Nh,Sh,Emax.a,Nc=0.95,parameter="mean",Asig="Neyman")
# n.ESTMAS(Nh,Ph,Emax.a,Nc=0.95,parameter="prop",Asig="Neyman")

# n.ESTMAS(Nh,Sh,Emax.a,Nc=0.95,parameter="mean",Asig="Proportional")
# n.ESTMAS(Nh,Ph,Emax.a,Nc=0.95,parameter="prop",Asig="Proportional")
```

Arguments

Nh	Numerical vector with the respective sizes of strata.
Sh	Numerical vector with the respective standard deviations of the variable of interest of each stratum. This argument is necessary only if the parameter of interest is the mean.
Ch	Numerical vector with the costs of sampling an element within each stratum. This argument is only necessary if the allocation by stratum is the optimal allocation.
Ph	Numerical vector with estimated proportions within each stratum.
Emax.a	Absolute maximum error.
parameter	Type of parameter to be estimated, either the mean or a proportion ("mean", "prop").
Nc	Confidence level (between 0 and 1) that you want to set.
Asig	Assignment by stratum ("Optima", "Neyman" or "Proportional")

Value

This function returns the sample size and the allocation by stratum, through the conditions established in the arguments.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
Nc<-0.95
E<-0.3
Nh<-c(400,220,380)
Sh<-sqrt(c(0.7521,1.4366,1.1361))
Ph<-c(0.4,0.2,0.6)
Ch<-c(1000,1200,1500)

# Optimal Assignment
n.ESTMAS(Nh=Nh,Sh=Sh,Ch=Ch,E=E,Nc=0.95,parameter="mean",Asig="Optima")
n.ESTMAS(Nh=Nh,Ph=Ph,Ch=Ch,E=E,Nc=0.95,parameter="prop",Asig="Optima")

# Neyman Assignment
n.ESTMAS(Nh=Nh,Sh=Sh,E=E,Nc=0.95,parameter="mean",Asig="Neyman")
n.ESTMAS(Nh=Nh,Ph=Ph,E=E,Nc=0.95,parameter="prop",Asig="Neyman")

# Proportional Assignment
n.ESTMAS(Nh=Nh,Sh=Sh,E=E,Nc=0.95,parameter="mean",Asig="Proportional")
n.ESTMAS(Nh=Nh,Ph=Ph,E=E,Nc=0.95,parameter="prop",Asig="Proportional")
```

---

n.MAS *Sample Size Using Simple Random Sampling Design Without Replacement*

Description

The n.MAS function determines the sample size by a simple random sample design without replacement, taking into account whether the parameter of interest is the mean (or total) or a proportion.

Usage

```
n.MAS(N,Argt,Nc=0.95,opc=2)

# n.MAS(N,Argt=c(S,Emax.a),opc=1,Nc=0.95)
# n.MAS(N,Argt=c(Cve,Emax.r),opc=2,Nc=0.95)
# n.MAS(N,Argt=c(p,Emax.a),opc=3,Nc=0.95)
# n.MAS(N,Argt=c(p,Emax.r),opc=4,Nc=0.95)
```

Arguments

N	Population size.
opc	Numeric value from 1 to 4, which indicates the option to choose.
Argt	Vector of length two, in which its components depends on the chosen option ("opc"). If option 1, (opc = 1) is chosen, the components of the Argt vector are in their order, the standard deviation of the variable of interest and the respective absolute maximum error that can be admitted; If option 2 (opc = 2) is chosen, the components of the Argt vector are respectively the estimated coefficient of variation and the relative maximum error to be controlled; If option 3 (opc = 3) is chosen, the components are the estimated proportion and absolute maximum error that can be admitted; And if option 4 (opc = 4) is chosen, the components are the estimated ratio and the relative maximum error respectively.
Nc	Confidence level (between 0 and 1) that you want to set.

Value

This function returns the sample size through the conditions set in the arguments.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
# Sample size for the mean (or total) when you want to control the absolute maximum error.

Nc<-0.95
S<-sqrt(6.0590)
Emax.a<-0.2
N<-10000
```

```
n.MAS(N=N,Argt=c(S,Emax.a),opc=1)


# Sample size for the mean (or total) when you want to control the relative maximum error.

Cve<-0.4346
Emax.r<-0.05
N<-10000
n.MAS(N=N,Argt=c(Cve,Emax.r))

# Sample size for proportions when you want to control the absolute maximum error.

N<-10000
p<-14/30
Emax.a<-0.04
Nc<-0.9
n.MAS(N=N,Argt=c(p,Emax.a),opc=3,Nc=Nc)

# Sample size for proportions when you want to control the relative maximum error.

N<-10000
p<- 14/30
Emax.r<-0.1
Nc<-0.9
n.MAS(N=N,Argt=c(p,Emax.r),opc=4,Nc=Nc)
```

---

n.MASC *Sample size using simple random sampling design without conglomerate replacement.*

Description

The n.MASC function determines sample size using a simple random sampling design without replacement of Conglomerates.

Usage

```
n.MASC(N,NI,Ni,St,Emax.a,Nc=0.95,n.equal=TRUE)

# For clusters with equal sizes.
# n.MASC(NI,Ni,St,Emax.a,Nc)

# For clusters with different sizes.
# n.MASC(N,NI,St,Emax.a,Nc,n.equal=FALSE)
```

Arguments

N	Size of the population, this argument is only necessary if the size of the conglomerates is different.

NI	Number of clusters in the population.
Ni	Size of the clusters, this argument is only necessary if the conglomerates have equal size (constant size).
St	Standard deviation of conglomerate totals.
Emax.a	Absolute maximum error.
Nc	Confidence level (between 0 and 1) to be set.
n.equal	Logical value indicating whether clusters have the same size

Value

This function returns the sample size under the conditions set in the arguments, that is, the number of clusters to select.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
# Sample size for populations with clusters of equal size.

st<-sqrt(1417.8668)
NI<-2000
Ni<-6
e<-2
nc=0.9
n.MASC(St=st,NI=NI,Ni=Ni,Emax.a=e,Nc=nc)

# Sample size for populations with clusters of different sizes.

st=sqrt(2019760.760)
N<-11000
NI<-400
e=10
nc=0.95
n.MASC(St=st,N=N,NI=NI,Emax.a=e,Nc=nc,n.equal=FALSE)
```

PiPT *Sampling Design without Replacement with Proportional Inclusion Probabilities for Sizes*

Description

The PiPT function selects a random sample or estimates an interest parameter under a sampling design with proportional inclusion probabilities proportional to size.

Usage

```
PiPT(xk,n,yk=NULL,zk=NULL,pik=NULL,mpikl=NULL,dk=NULL,type="selec",
     parameter="total",Nc=0.95,Ek=NULL)

# To select: PiPT(xk,n)

# To estimate: PiPT(yk,pik,mpikl,type="estm",parameter="total")

# To estimate in domains
# PiPT(yk,pik,mpikl,dk,type="estm",parameter="total")
```

Arguments

xk	Vector of observations of the auxiliary variable. This vector is only necessary if you wish to select.
n	Sample size.
yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate.
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.
pik	Vector of the first-order inclusion probabilities.
mpikl	Matrix of second-order inclusion probabilities.
type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If type is equal to "select" the function will make a selection, if it is equal to "estm" the function will make the estimation of the indicated parameter and if it is equal to "estm.Ud" it will make an estimate in domain.
dk	Factor that indicates the individuals that belong to each domain of interest, Only needed if "type"" is equal to "estm.Ud".
parameter	This argument indicates the parameter to be estimated ("total", "mean", "prop", or "ratio").
Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case "type" is equal to "estm" or "estm.Ud".
Ek	Vector of random numbers of length equal to the size of the population. This argument is optional and by default the function generates them from a uniform distribution (0,1).

Value

The PiPT function returns two types of results using a sampling design with inclusion probabilities proportional to size, depending on the argument "type", which indicates whether to select ("select") or estimate ("estm" or "estm.Ud").

If type="select" the function will return a list with three elements:

`Ksel`	Vector with the positions of the selected individuals
`piksel`	First order inclusion probability vector of selected individuals
`mpikl.s`	Matrix of the second-order inclusion probabilities

If type="estm" or type="estm.Ud", the function returns a data frame with the estimation of the parameter of interest, the estimated variance of the estimator, the standard error, the coefficient of variation (in percentage) and a confidence interval.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
set.seed(12265)
yk<-rnorm(100,mean=50,sd=5)
zk<-rnorm(100,mean=51,sd=5)
yk.p<-as.factor(ifelse(yk>50,"A","B"))
set.seed(12245)
# Información Auxiliar
xk<-yk*runif(100,min=0.9,max=1.1)
r<-cor(yk,xk)

selection<-PiPT(xk=xk,n=10,type="selec")
PiPT(yk=yk[selection$Ksel],pik=selection$pik,mpikl=selection$mpikl.s,
     type="estm",parameter="total")
PiPT(yk=yk[selection$Ksel],pik=selection$pik,mpikl=selection$mpikl.s,
     type="estm",parameter="mean")
PiPT(yk=yk.p[selection$Ksel],pik=selection$pik,mpikl=selection$mpikl.s,
     type="estm",parameter="prop")
PiPT(yk=yk[selection$Ksel],zk=zk[selection$Ksel],pik=selection$pik,
     mpikl=selection$mpikl.s,type="estm",parameter="ratio")

# Domain Estimate
```

```
Sex<-rep(1:2,length=100)
dk<-factor(Sex,labels=c("Man","Woman"))
PiPT(yk=yk[selection$Ksel],pik=selection$pik,mpikl=selection$mpikl.s,
    dk=dk[selection$Ksel],type="estm.Ud",parameter="total")
PiPT(yk=yk[selection$Ksel],pik=selection$pik,mpikl=selection$mpikl.s,
     dk=dk[selection$Ksel],type="estm.Ud",parameter="mean")
PiPT(yk=yk.p[selection$Ksel],pik=selection$pik,mpikl=selection$mpikl.s,
     dk=dk[selection$Ksel],type="estm.Ud",parameter="prop")
PiPT(yk=yk[selection$Ksel],zk=zk[selection$Ksel],pik=selection$pik,
     mpikl=selection$mpikl.s,dk=dk[selection$Ksel],type="estm.Ud",
     parameter="ratio")
```

PPT *Sampling Design with Replacement and Size Proportional Selection Probabilities*

Description

The PPT function selects a random sample or estimates a parameter of interest under a sampling design with proportional proportional selection probabilities (PPT).

Usage

```
PPT(xk,m,yk=NULL,zk=NULL,pk=NULL,dk=NULL,type="selec",parameter="total",
    method ="acum.total",Nc=0.95,Ek=NULL)

# To select: PPT(xk,m,method="acum.total")
# To estimate: PPT(yk,pk,type="estm",parameter)
# To estimate in domains: PPT(yk,pk,dk,type="estm.Ud",parameter)
```

Arguments

xk	Vector of observations of the auxiliary variable. This vector is only necessary if you wish to select.
m	Sample size.
yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate.
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.
pk	Vector of the probabilities of selection of individuals.
dk	Factor that indicates the individuals that belong to each domain of interest, is only necessary if "type" is equal to "estm.Ud".

type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If "type" is equal to "select" the function will make a selection, if it is equal to "estm" the function will perform the estimation of the indicated parameter and if it is equal to "estm.Ud" it will make an estimate in domain.
method	Indicates the method or selection mechanism. If method is equal to "total cum." The function uses the total cumulative method or if it is equal to "lahiri" the function uses the method of Lahiri.
parameter	This argument indicates the parameter to be estimated ("total", "mean", "prop", or "ratio").
Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case the type is equal to "estm" or "estm.Ud".
Ek	Vector of random numbers of length equal to the size of the population. This argument is optional and by default the function generates them from a uniform distribution (0,1).

Value

This function returns two types of results using the PPT sampling design, depending on the "type" argument with which to select ("select") or estimate ("estm" or "estm.Ud").

If type is equal to "select" the function will return a list with two elements:

Ksel	Vector with the positions of the selected individuals.
pksel	Selection probabilities vector of selected individuals.

-If type="estm" or type="estm.Ud", the function returns a data frame with the estimation of the parameter of interest, the estimated variance of the estimator, the standard error, the coefficient of variation (in percentage) and a confidence interval.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
set.seed(12265)
yk<-rnorm(100,50,5)
zk<-rnorm(100,12,4)
set.seed(12245)
```

```
xk<-yk*runif(100,min=0.9,max=1.1)
r<-cor(yk,xk)
yk.p<-as.factor(ifelse(yk>50,"A","B"))

selection<-PPT(xk=xk,m=10,type="selec",method="acum.total")
PPT(yk=yk[selection$Ksel],pk=selection$pksel,type="estm",parameter="total")
PPT(yk=yk[selection$Ksel],pk=selection$pksel,type="estm",parameter="mean")
PPT(yk=yk.p[selection$Ksel],pk=selection$pksel,type="estm",parameter="prop")
PPT(yk=yk[selection$Ksel],zk=zk[selection$Ksel],pk=selection$pksel,
    type="estm",parameter="ratio")

# Domain Estimate

Sex<-rep(1:2,length=100)
dk<-factor(Sex,labels=c("Man","Woman"))
PPT(yk=yk[selection$Ksel],dk=dk[selection$Ksel],pk=selection$pksel,
    type="estm.Ud",parameter="total")
PPT(yk=yk[selection$Ksel],dk=dk[selection$Ksel],pk=selection$pksel,
    type="estm.Ud",parameter="mean")
PPT(yk=yk.p[selection$Ksel],dk=dk[selection$Ksel],pk=selection$pksel,
    type="estm.Ud",parameter="prop")
PPT(yk=yk[selection$Ksel],zk=zk[selection$Ksel],dk=dk[selection$Ksel],
    pk=selection$pksel,type="estm.Ud",parameter="ratio")
```

R.SIS *R-Systematic Sampling Design*

Description

The R.SIS function selects a random sample or estimates a parameter of interest under a r-systematic sampling design.

Usage

```
R.SIS(N,n,r,yk=NULL,zk=NULL,fact=NULL,dk=NULL,type="selec",
      parameter="total",Nc=0.95,Ek=NULL)

# To select: R.SIS(N,n,r)

#To estimate: R.SIS(N,n,r,fact,yk,type="estm",parameter)

# To estimate in domains
# R.SIS(yk,fact,N,n,r,type="estm.Ud",parameter)
```

Arguments

N	Size of the population.
n	Sample size.

r	Number of starts.
yk	Vector of observations of the characteristic of interest. This vector is only necessary if you want to estimate.
zk	Vector of observations of the characteristic of interest of equal length that yk. This vector is necessary if the parameter of interest is the ratio and refers to the variable involved in the denominator of the ratio.
fact	Factor indicating that Ur belongs to the observations of the variable of interest and yk. This factor is only necessary if type is equal to "estm" or "estm.Ud".
dk	Factor that indicates the individuals that belong to each domain of interest, is only necessary if type is equal to "estm.Ud".
type	This argument indicates the procedure that will have the function ("select", "estm" or "estm.Ud"). If "type" is equal to "select" the function will make a selection, if it is equal to "estm" the function will perform the estimation of the indicated parameter and if it is equal to "estm.Ud" it will make an estimate in domain.
parameter	This argument indicates the parameter to be estimated ("total", "mean", "prop", or "ratio").
Nc	Confidence level (between 0 and 1), for the confidence interval of the estimator in case "type" is equal to "estm" or "estm.Ud".
Ek	Vector of random numbers of length equal to the size of the population. This argument is optional and by default the function generates them from a uniform distribution (0,1).

Value

This function returns two types of results using the r-systematic sampling design, depending on the "type" argument with which to select ("select") or estimate "estm.Ud").

If type="select", the function returns a list with four elements:

Sel	Array with r columns that refers to the clusters selected by each boot
Ksel	Vector with selected individuals
fact	factor indicating which start each selected individual belongs to
n.s	Sample size

If type="estm" or type="estm.Ud" the function returns a data frame with the estimation of the parameter of interest, the estimated variance of the estimator, the standard error, the coefficient of variation (in percent), an interval of confidence, the intraclass correlation coefficient and the intra-sample rate of variance.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
yk<-rnorm(100,40,2)
zk<-rnorm(100,12,2)
yk.p<-as.factor(ifelse(yk>40,"A","B"))
selection<-R.SIS(N=100,n=20,r=3,type="selec")


R.SIS(yk=yk[selection$Ksel],fact=selection$fact,N=100,n=20,r=3,
      type="estm",parameter="total")
R.SIS(yk=yk[selection$Ksel],fact=selection$fact,N=100,n=20,r=3,
       type="estm",parameter="mean")
R.SIS(yk=yk.p[selection$Ksel],fact=selection$fact,N=100,n=20,r=3,
       type="estm",parameter="prop")
R.SIS(yk=yk[selection$Ksel],zk=zk[selection$Ksel],fact=selection$fact,
       N=100,n=20,r=3,type="estm",parameter="ratio")


#Domain Estimate

Sex<-rep(1:2,length=100)
dk<-factor(Sex,labels=c("Man","Woman"))
R.SIS(yk=yk[selection$Ksel],fact=selection$fact,dk=dk[selection$Ksel],
      N=100,n=20,r=3,type="estm.Ud",parameter="total")
R.SIS(yk=yk[selection$Ksel],fact=selection$fact,dk=dk[selection$Ksel],
      N=100,n=20,r=3,type="estm.Ud",parameter="mean")
R.SIS(yk=yk.p[selection$Ksel],fact=selection$fact,dk=dk[selection$Ksel],
      N=100,n=20,r=3,type="estm.Ud",parameter="prop")
R.SIS(yk=yk[selection$Ksel],zk=zk[selection$Ksel],fact=selection$fact,
      dk=dk[selection$Ksel],N=100,n=20,r=3,type="estm.Ud",parameter="ratio")
```

WHICH1 *Positions of the components of a vector with respect to another vector*

Description

The WHICH1 function returns the positions in which the vector components (V1) are located in another vector (V2).

Usage

```
WHICH1(V1,V2)
```

Arguments

V1	Vector initial.
V2	Vector containing replicates of the components of the initial vector.

Value

This function is used to extract the positions of all the individuals that are part of the selected clusters, in a cluster sampling.

Author(s)

Jorge Alberto Barón Cárdenas <jorgeabaron@correo.unicordoba.edu.co>

Guillermo Martínez Flórez <guillermomartinez@correo.unicordoba.edu.co>

References

Särndal, C. E., J. H. Wretman, and C. M. Cassel (1992). Foundations of Inference in Survey Sampling. Wiley New York.

Cochran, W. G. (1977). Sampling Techniques, 3ra ed. New York: Wiley.

Thompson, S. K. (1945). Wiley Series in Probability and Statistics, Sampling, 1ra ed. United States of America.

Examples

```
cong<-rep(1:12,each=10)
Argt<-list(NI=12,nI=3)
selection<-CONGL(Argt=Argt,design="MAS")
WHICH1(selection$Ksel,cong)
```

Index