Package 'stratbr'

May 9, 2026

Type Package

Title Optimal Stratification in Stratified Sampling

Version 1.2

Date 2017-05-19

Author Jose Brito, Pedro Silva and Tomas Veiga

Maintainer Jose Brito <jambrito@gmail.com>

Description An Optimization Algorithm Applied to
Stratification Problem.This function aims
at constructing optimal strata with an optimization algorithm
based on a global optimisation technique called Biased
Random Key Genetic Algorithms.

Depends Rglpk,snowfall,stratification

Encoding UTF-8

License GPL-2

LazyData true

RoxygenNote 5.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2017-05-26 21:44:59 UTC

Contents

BSSM_FD *Optimal Allocation - Minimum Coefficient of Variation*

Description

Function that uses an integer programming formulation for allocation of the overall sample size n to the strata, for the following purpose: Coefficient of Variation of the estimate of total for the survey variable is minimized.

Usage

```
BSSM_FD(Nh, Sh2x, n, H, nmin = 2, X, takeall = FALSE)
```

Arguments

Nh	Vector with number of population elements, or population size, in stratum h
Sh2x	Vector with population variance of the variable X in stratum h.
n	Sample size.
H	Number of strata.
nmin	Minimum sample size (smallest possible sample size in any stratum). The default is 2.
X	Population Total
takeall	Take-all stratum (takeall=TRUE) => nH=NH.

Value

solution	Vector with sample of size by stratum and coefficient of variation for the estimator of total of the stratification variable considered.

Author(s)

Jose Brito (jambrito@gmail.com), Pedro Silva, Gustavo Semaan and Nelson Maculan.

References

Brito, J.A.M, Silva, P.L.N.,Semaan, G.S. and Maculan, N. (2015). Integer Programming Formulations Applied to Optimal Allocation in Stratified Sampling. Survey Methodology, 41: 427-442.

Examples

```
X<-round(100*runif(50))
Nh<-c(10,20,20)
Sh2x<-c(var(X[1:10]),var(X[11:30]),var(X[31:50]))
aloc1<-BSSM_FD(Nh,Sh2x,n=40,H=3,nmin=2,sum(X),takeall=TRUE)
Nh<-c(49,78,20,39,73,82,89)
X<-542350
Sh2x<-c(4436978,5581445,33454902,5763294,8689167,3716130,13938505)
aloc2<-BSSM_FD(Nh,Sh2x,n=100,H=7,nmin=2,X)
```

`stratbr` *Optimization Algorithm to solve stratification problem*

Description

This function aims at constructing optimal strata with an optimization algorithm based on a global optimisation technique called Biased Random Key Genetic Algorithms(BRKGA). The optimization algorithm is applied to solve the one dimensional case, which reduces the stratification problem to just determining strata boundaries. Assuming that the number H of strata and the total sample size n are fixed, it is possible to produce the strata boundaries by taking into consideration an objective function associated with the variance. This function determines strata boundaries so that the elements in each stratum are more homogeneous among themselves.

Usage

```
stratbr(X, H = 3, n = 30, nmin = 2, takeall = FALSE, tampop = 100,
  totgen = 1500, pelite = 0.2, pmutant = 0.3, rc = 0.6, cores = 2)
```

Arguments

`X`	Stratification variable.
`H`	Number of strata.
`n`	Sample size.
`nmin`	Minimum sample size (smallest possible sample size in any stratum).
`takeall`	Take-all stratum (takeall=TRUE) => nH=NH.
`tampop`	Number of chromosomes BRKGA.The default is 100.
`totgen`	Maximum number of generations BRKGA.The default is 1500.
`pelite`	Percentage elite solutions BRKGA.The default is 0.2.
`pmutant`	Percentage mutant solutions BRKGA.The default is 0.3.
`rc`	Crossover probability BRKGA. The default is 0.6.
`cores`	Numerical amount of CPUs requested for the cluster.

Value

`cvtot`	Coefficient of variation for the estimator of total of the stratification variable considered.
`nh`	Number of sample elements, or sample size, in stratum h.
`Nh`	Number of population elements, or population size, in stratum h.
`Sh2`	Population variance of the stratification variable x in stratum h.
`bk`	Strata boundaries
`cputime`	Time consumed by the algorithm in seconds.

Author(s)

Jose Brito (jambrito@gmail.com), Pedro Luis and Tomas Veiga.

References

Brito, J.A.M, Silva, P.L.N.,Semaan, G.S. and Maculan, N. (2015). Integer Programming Formulations Applied to Optimal Allocation in Stratified Sampling. Survey Methodology, 41: 427-442.

Brito, J.A.M, Semaan, G.S., Fadel, A.C. and Brito, L.R.(2017). An optimization approach applied to the optimal stratification problem, Communications in Statistics - Simulation and Computation.

Gonçalves, J.R. and Resende, M.G.C. (2011). Biased random-key genetic algorithms for combinatorial optimization, Journal of Heuristics, 17: 487-525.

Examples

```
data(Sweden)
REV84<-Sweden[,9]
solution1<-stratbr(REV84,H=3,n=50,nmin=10,totgen=2,cores=4)
data(USbanks)
solution2<-stratbr(USbanks,H=3,n=50,totgen=2,cores=4,takeall=TRUE)
```

Index