Package 'stratvns'

October 14, 2022

Type Package

Title Optimal Stratification in Stratified Sampling

Version 1.1

Author Leonardo de Lima, Jose Brito, Pedro Gonzalez and Breno Oliveira

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 vns.

Depends R (>= 3.5.0), parallel, purrr, partitions, MultAlloc

Encoding UTF-8

License GPL-2

LazyData true

RoxygenNote 7.1.0

NeedsCompilation no

Repository CRAN

Date/Publication 2020-07-09 12:30:06 UTC

R topics documented:

STRATENUM *Enumeration Algorithm*

Description

This function enumerates all feasible solutions to the stratification problem and produces the global optimum, applying an integer formulation proposed by Brito et al (2015).

Usage

```
STRATENUM(X, L, cvt = 0.1, nhmin = 2)
```

Arguments

X	Stratification Variable
L	Number of strata
cvt	Target cv
nhmin	Mininum sample size by stratum

Details

STRATENUM

Value

n	Sample size
nh	Sample size by strata
cv	coefficient of variation
Nh	Strata sizes
Vh	Strata variances
totoptg	Total global optimal solutions
tfeasible	Total feasible solutions
cputime	Runtime in seconds

Author(s)

Leonardo de Lima, Jose Brito, Pedro Gonzalez and Breno Oliveira

References

1. Brito, J.A.M., Silva, P.L.N., Semaan, G.S., Maculan, N., 2015. Integer programming formulations applied to optimal allocation in stratified sampling. Survey Methodology 41, 2, 427–442.

Examples

```
## Not run:
Example1:
s<-STRATENUM(U21,L=3,cvt=0.05)
Example2:
s<-STRATENUM(U15,L=4)
Example3:
s<-STRATENUM(U1,L=3,nhmin=4)

## End(Not run)
```

STRATVNS *Vns Algorithm*

Description

This function aims at constructing optimal strata with an optimization algorithm based on a global optimisation technique called Variable neighborhood search (VNS). 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 L of strata and the coefficient of variation are fixed, it is possible to produce the strata boundaries by taking into consideration an objective function associated with the sample size. This function determines strata boundaries so that the elements in each stratum are more homogeneous among themselves and produce minimum sample size applying an integer formulation proposed by Brito et al (2015).

Usage

```
STRATVNS(
  X,
  L = 3,
  cvt = 0.1,
  nhmin = 2,
  maxstart = 3,
  imax = 3,
  kmax = 3,
  s = 30,
  sl = 50,
  tmax = 15,
  nsols = 20,
  cputime = 3600,
  nIterWithNoImpMax = 5,
  parallelize = TRUE
)
```

Arguments

X	Stratification Variable

L	Number of strata
cvt	Target cv
nhmin	Mininum sample size by stratum
maxstart	Number of iterations in multstart
imax	Maximum Number Iterations - VNS
kmax	Maximum Neighborhoods = number of cut points selected to apply shaking and local search
s	Range of shaking procedure
sl	Range of RVNS procedure
tmax	Maximum number cut points in neighborhoods
nsols	Number of initial solutions generated
cputime	Maximum cpu time in seconds
nIterWithNoImpMax	Maximum number of iterations without improvement in VNS
parallelize	TRUE = Performs multiple vns calls in parallel

Details

STRATVNS

Value

bk	Cut points
n	Minimum sample size
nh	Sample size by strata
cv	coefficient of variation
Nh	Strata sizes
Vh	Strata variances
cputime	Runtime in seconds

Author(s)

Leonardo de Lima, Jose Brito, Pedro Gonzalez and Breno Oliveira

References

1. Hansen, P., Mladenovi´c, N., 2001. Variable neighborhood search: Principles and applications. European Journal of Operational Research 130, 3, 449 – 467.

2. Brito, J.A.M., Silva, P.L.N., Semaan, G.S., Maculan, N., 2015. Integer programming formulations applied to optimal allocation in stratified sampling. Survey Methodology 41, 2, 427–442.

Examples

```
## Not run:
Example1:
s<-STRATVNS(U1,L=4,cvt=0.05,nhmin=3)
Example2:
s<-STRATVNS(U15,L=3)
#'Example3:
s<-STRATVNS(U21,L=5)
Example4:
s<-STRATVNS(U1,L=3,nhmin=4)

## End(Not run)
```

U1 *Population U1*

Description

Australian cattle farms stratified by industrial regions

Usage

```
U1
```

Format

A vector 430x1:

Details

U1

References

1. Chambers, R., Dunstan, R., 1986. Estimating distribution functions from survey data. Biometrika 73, 3, 597–604.

U15 *Population U15*

Description

Population in thousands of 284 municipalities in Sweden in 1975

Usage

```
U15
```

Format

A vector 284x1:

Details

U15

References

Särndal, C.E., Swensson,B., Wretman, J. (2003). Model Assisted Survey Sampling, 1st edition, Springer.

U21 *Population U21*

Description

Million dollar funds from major US commercial banks

Usage

```
U21
```

Format

A vector 357x1:

Details

U21

References

Särndal, C.E., Swensson,B., Wretman, J. (2003). Model Assisted Survey Sampling, 1st edition, Springer.

Index