Package: MultAlloc (via r-universe)

January 26, 2025

Type Package

Title Optimal Allocation in Stratified Sampling

Version 1.2

Date 2015-12-26

Author Jose Brito, Pedro Silva, Nelson Maculan and Gustavo Semaan

Maintainer Jose Brito <jambrito@gmail.com>

Description Integer Programming Formulations Applied to Univariate and Multivariate Allocation Problems.

Imports Rglpk

License GPL-2

LazyData TRUE

RoxygenNote 5.0.1

NeedsCompilation no

Date/Publication 2015-12-27 10:25:54

Additional_repositories https://cranhaven.r-universe.dev

Config/pak/sysreqs libglpk-dev

Repository https://cranhaven.r-universe.dev

RemoteUrl https://github.com/cranhaven/cranhaven.r-universe.dev

RemoteRef package/MultAlloc

RemoteSha c17ad65ab01e96cf8bae266c937036617530e320

RemoteSubdir MultAlloc

Contents

BSSM_FC *Optimal Allocation - Minimum Sample Size*

Description

Allocation of the overall sample size n to the strata for the following purpose:

The total variable survey cost C (c1.n1+c2.n2+...+cH.nH)is minimized, subject to having Coefficients of Variation (CVs) for the estimates of totals of the m survey variables below specified thresholds.

If the unit level survey costs for sampling from the various strata are unknown or are assumed to be the same, then c1,c2,...,cH may all be set to one and the alternative objective function to minimize is n1+n2+...+nH.

Usage

```
BSSM_FC(Nh,Sh2j,Yj,cvt,nmin,ch,certain)
```

Arguments

Nh	Vector with total number of population units in each stratum (h=1,...,H)
Sh2j	Matrix (or vector) mxH (m = number of variables and H =number of strata) with Population variance for each variable of the hth stratum
Yj	Vector with population total Yj for the jth survey variable
cvt	Vector with target cvs asociated with survey variables
nmin	Smallest possible sample size in any stratum
ch	Vector with the unit level survey costs for sampling from stratum h
certain	if (nH=NH) => certain=TRUE else certain=FALSE

Details

Function that uses an integer programming formulation

Value

n	Sample size
nh	Sample of size by stratum
cvs	Coefficients of variation for the estimators of totals of the survey variables considered
time_cpu	Time consumed by the algorithm (seconds)

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, No.2, pp.427-442.

See Also

BSSM_FD

Examples

```
#Example1 - Number of survey variables (m=2) and seven strata (H=7)
Nh<-c(49,78,20,39,73,82,89)
Yj<-c(542350,56089251)
Sh2j<-rbind(c(4436978,5581445,33454902,5763294,8689167,3716130,13938505),
        c(11034299660,40919330279,33519355946,18228286901,74247764986,49062224184,5783096806))
cvt<-c(0.02,0.02)
result<-BSSM_FC(Nh,Sh2j,Yj,cvt)

#Example2
#nmin>2
Nh<-c(49,78,20,39,73,82,89)
Yj<-c(542350,56089251)
Sh2j<-rbind(c(4436978,5581445,33454902,5763294,8689167,3716130,13938505),
        c(11034299660,40919330279,33519355946,18228286901,74247764986,49062224184,5783096806))
cvt<-c(0.1,0.1)
nmin<-20
result<-BSSM_FC(Nh,Sh2j,Yj,cvt,nmin)

#Example3
#certain=TRUE
Nh<-c(49,78,20,39,73,82,89)
Yj<-c(542350,56089251)
Sh2j<-rbind(c(4436978,5581445,33454902,5763294,8689167,3716130,13938505),
        c(11034299660,40919330279,33519355946,18228286901,74247764986,49062224184,5783096806))
cvt<-c(0.1,0.1)
result<-BSSM_FC(Nh,Sh2j,Yj,cvt,certain=TRUE)


#Example4
#Number of survey variables m=1
Nh<-c(49,78,20,39,73,82,89)
Yj<-542350
Sh2j<-c(4436978,5581445,33454902,5763294,8689167,3716130,13938505)
cvt<-0.1
result<-BSSM_FC(Nh,Sh2j,Yj,cvt)
```

BSSM_FD *Optimal Allocation - Minimum Sum of Relative Variances*

Description

Allocation of the overall sample size n to the strata for the following purpose:

A weighted sum of variances (or relative variances) of the estimates of totals for the m survey variables is minimized.

Usage

```
BSSM_FD(Nh,Sh2j,Yj,Cust,nmin,ch,w,certain)
```

Arguments

Nh	Vector with total number of population units in each stratum (h=1,...,H)
Sh2j	Matrix (or vector) mxH (m = number of variables and H =number of strata) with Population variance for each variable of the hth stratum
Yj	Vector with population total Yj for the jth survey variable
Cust	Corresponds to the overall variable cost budget for the survey C
nmin	Smallest possible sample size in any stratum
ch	Vector with the unit level survey costs for sampling from stratum h
w	Vector with Variable-specific weights, set a priori to represent the relative importance of the survey variables
certain	if (nH=NH) => certain=TRUE else certain=FALSE

Details

Function that uses an integer programming formulation

Value

n	Sample size
nh	Sample of size by stratum
cvs	Coefficients of variation for the estimators of totals of the survey variables considered
time_cpu	Time consumed by the algorithm (seconds)

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.Srogramming Formulations Applied to Optimal Allocation in Stratified Sampling. Survey Methodology, 41, No.2, pp.427-442.

See Also

BSSM_FC

Examples

```
#Example1
#Unit level survey costs for sampling from the strata are assumed to be the same.
#Number of survey variables (m=2) and seven strata (H=7)
#ch=1 ==> Cust = n
Nh<-c(49,78,20,39,73,82,89)
Yj<-c(542350,56089251)
Sh2j<-rbind(c(4436978,5581445,33454902,5763294,8689167,3716130,13938505),
        c(11034299660,40919330279,33519355946,18228286901,74247764986,49062224184,5783096806))
n<-200 #sample size
result<-BSSM_FD(Nh,Sh2j,Yj,Cust=n)

#Example2
#Unit level survey costs for sampling from the strata are assumed to be the same.
#ch=1 ==> Cust = n
#nmin>2
Nh<-c(49,78,20,39,73,82,89)
Yj<-c(542350,56089251)
Sh2j<-rbind(c(4436978,5581445,33454902,5763294,8689167,3716130,13938505),
        c(11034299660,40919330279,33519355946,18228286901,74247764986,49062224184,5783096806))
nmin<-20
n<-200
result<-BSSM_FD(Nh,Sh2j,Yj,nmin,Cust=n)

#Example3
#certain=TRUE
Nh<-c(49,78,20,39,73,82,89)
Yj<-c(542350,56089251)
Sh2j<-rbind(c(4436978,5581445,33454902,5763294,8689167,3716130,13938505),
        c(11034299660,40919330279,33519355946,18228286901,74247764986,49062224184,5783096806))
n<-200
result<-BSSM_FD(Nh,Sh2j,Yj,Cust=n,certain=TRUE)


#Example4
#Relative importance of the survey variables is different
w<-c(0.3,0.7)
Nh<-c(49,78,20,39,73,82,89)
Yj<-c(542350,56089251)
Sh2j<-rbind(c(4436978,5581445,33454902,5763294,8689167,3716130,13938505),
        c(11034299660,40919330279,33519355946,18228286901,74247764986,49062224184,5783096806))
n<-200
result<-BSSM_FD(Nh,Sh2j,Yj,Cust=n,w=w)
```

```
#Example5
#Number of survey variables m=1
Nh<-c(49,78,20,39,73,82,89)
Yj<-542350
Sh2j<-c(4436978,5581445,33454902,5763294,8689167,3716130,13938505)
n<-100
result<-BSSM_FD(Nh,Sh2j,Yj,Cust=n)
```

Index