auditsampling

May 21, 2019

Title A auxiliary package to calculate some basic things in sampling. Most used para CGU auditors.

Version 0.0.0.9000

Description The package contains some functions the help auditors in sampling process.

Depends R (>= 3.4.3)

License GPL-2 | GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1.9000

R topics documented:

---

nClusterProp   *Calculates the number os cluster to a cluster sample design*

Description

This function returns the number of clusters to be sampled in a proportion based survey. By default, the argument 'pq' is 0.25 (maximum variance).

Usage

```
nClusterProp(data, clustername, alpha, moe, pq = 0.25)
```

Arguments

data	The data.frame containing the variable denoted as the cluster.
clustername	The variable name of the considered cluster.
alpha	Confidence level.
moe	Margin of error.
pq	The variance, as $P(1-P)$. By default, it is 0.25.

Value

The value (integer) of the number of clusters to sample in cluster sampling.

Examples

```
library(survey)
data("api")
nClusterProp(data=apipop, clustername = "dnum", alpha = 0.95, moe = 0.1)
nClusterProp(data=apipop, clustername = "dnum", alpha = 0.95, moe = 0.05)
```

---

nPilot	*Calculates the remaing sample size when using pilot sample*

Description

Based on a pilot sample, this function calculates the remain elements to the final sample given the variance considered.

Usage

```
nPilot(s2 = NULL, pq = NULL, n1, V)
```

Arguments

pq	The variance, as $P(1 - P)$. Not necessary if 'S2' is not 'NULL'.
n1	The size of the pilot sample.
V	The target variance.
S2	The variance, as σ^2. Not necessary if 'S2' is not 'NULL'.

Value

The final sample is

$$n = n_1 + n_2$$

Where n_2 is the value returned by this function.

References

COCHRAN, William Gemmell. Sampling techniques-3. 1977.

Examples

```
nPilot(s2=13.5,n1=30,V=1)
nPilot(pq=0.15,n1=30,V=0.025)
```

---

nStrata   *Calculates the sample size in a stratified desgin*

Description

This function returns the number of elements to be sampled in a stratified based survey. If the method choosen is 'prop' or 'optimum', the function return the allocation too.

Usage

```
nStrata(data, stratanames, alpha, moe, S2 = NULL, pq = NULL,
  V = NULL, N = Inf, method = c("none", "prop", "optimum"))
```

Arguments

data	The data.frame containing the variable denoted as the cluster.
stratanames	The variable name of the considered strata. Can be a vector of variable names.
alpha	Confidence level.
moe	Margin of error.
S2	The variance, as σ^2. Not necessary if 'pq' or 'V' is not 'NULL'.
pq	The variance, as $P(1 - P)$. Not necessary if 'S2' or 'V' is not 'NULL'.
V	The variance, as $(d/t)^2$. Not necessary if 'S2' or 'pq' is not 'NULL'.
N	Population size.
method	A string with 'none', 'prop' or 'optimum'. With 'optimum', it uses Neyman allocation.

Value

The value (integer) of the number of elements to sample in stratified sampling. If 'prop' or 'optimum' where used, the function returns a list with the components

n Total number of elements in the sample.

nh In case of method 'prop' or 'optimum', the allocation in each strata.

Examples

```
library(survey)
data("api")
nStrata(apipop,stratanames = "stype", alpha = 0.95, moe = 0.05,
        pq = 0.25, N = Inf, method = "none")
nStrata(apipop,stratanames = "stype", alpha = 0.95, moe = 0.05,
        pq = 0.25, N = nrow(apipop), method = "prop")
nStrata(apipop,stratanames = "stype", alpha = 0.95, moe = 0.05,
        pq = 0.25, N = nrow(apipop), method = "optimum")
### Cochran's example (Cochran, W. G. (2007). Sampling techniques. John Wiley & Sons. p. 106-107)
data_cochran=data.frame(stratum=rep(c(1:6),c(13,18,26,42,73,24)))
nStrata(data_cochran,stratanames = "stratum", N=196, method = "optimum",
S2=c(325^2,190^2,189^2,82^2,86^2,190^2),V=7974976)
```

NSubpop   *Calculates the new subpopulation due elements thar not belong to the*
                                 populational universe

Description

This function returns the number of the new subpopulation.

Usage

```
NSubpop(N, n, p, z = 3)
```

Arguments

N	Population size.
n	Sample size.
p	Proportion or the number of elements that belong to the universe intended.
z	The z-value of α.
data	The data.frame containing the variable denoted as the cluster.

Value

The value (integer) of the new population (subpopulation) number.

Examples

```
NSubpop(45000,100,59)
NSubpop(45000,100,5)
```

Index