

Package ‘CARE1’

February 19, 2015

Type Package

Title Statistical package for population size estimation in capture-recapture models.

Version 1.1.0

Date 2012-10-23

Author T.C. Hsieh

Maintainer Anne Chao <chao@stat.nthu.edu.tw>

Description The R package CARE1, the first part of the program CARE (Capture-Recapture) in <http://chao.stat.nthu.edu.tw/softwareCE.html>, can be used to analyze epidemiological data via sample coverage approach (Chao et al. 2001a). Based on the input of records from several incomplete lists (or samples) of individuals, the R package CARE1 provides output of population size estimate and related statistics.

License GPL-3

LazyLoad yes

Repository CRAN

Date/Publication 2012-10-24 12:31:15

NeedsCompilation no

R topics documented:

CARE1-package	2
as.record	2
CARE1.print	3
Congenital	4
Diabetes	5
estN	6
estN.n	7
estN.pair	7
estN.para	8

estN.stat	9
HAV	10

Index	11
--------------	-----------

CARE1-package	<i>Statistical package for population size estimation in capture-recapture models.</i>
---------------	--

Description

The R package CARE1, the first part of the program CARE (Capture-Recapture) in <http://chao.stat.nthu.edu.tw/softwareCE.h> can be used to analyze epidemiological data via sample coverage approach (Chao et al. 2001a). Based on the input of records from several incomplete lists (or samples) of individuals, the R package CARE1 provides output of population size estimate and related statistics.

Details

Package:	CARE1
Type:	Package
Version:	1.1.0
Date:	2012-10-23
License:	GPL-3
LazyLoad:	yes

Author(s)

T.C. Hsieh
 Maintainer: Anne Chao <chao@stat.nthu.edu.tw>

References

Chao, A., and Tsay, P. K. (1998). A sample coverage approach to multiple-system estimation with application to census undercount. *Journal of the American Statistical Association*, 93: 283-293.
 Chao, A., Tsay, P. K., Lin, S. H., Shau, W. Y., and Chao, D. Y. (2001a). The applications of capture-recapture models to epidemiological data. *Statistics in Medicine*, 20: 3123-3157.

as.record	<i>Transform observed data to ascertainment records</i>
-----------	---

Description

Transform observed data to ascertainment records.

Usage

```
as.record(x)
```

Arguments

x the matrix of the observed capture histories

Details

x has one row per unit captured in the experiment. Each row is an observed capture history. It must contain only zeros and ones; the number one indicates a capture. In this case, the number of columns in the table represents the number of capture occasions in the experiment (noted t). Here is a fictive example of a data set of this type for t=3:

```
[1,] 0 1 0
[2,] 1 0 1
[3,] 0 1 0
[4,] 0 1 1
[5,] 1 0 1
[6,] 1 0 1
[7,] 0 0 1
[8,] 0 1 0
[9,] 0 1 0
[10,] 1 1 0
```

Value

the vector of capture histories or ascertainment records.

Examples

```
x=matrix(sample(0:1,300,TRUE),ncol=3)
as.record(x)
```

CARE1.print

A quick analysis function for capture-recapture data

Description

A quick analysis function for of capture-recapture data which integrating all output in the same time.

Usage

```
CARE1.print(z)
```

Arguments

`z` the vector of capture histories or ascertainment records.

Value

integrated output

References

- Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics*, 43: 783-791.
- Chao, A., Tsay, P. K., Lin, S. H., Shau, W. Y., and Chao, D. Y. (2001). The applications of capture-recapture models to epidemiological data. *Statistics in Medicine*, 20: 3123-3157.
- Seber, G. A. F. (1982). *The estimation of animal abundance* (2 ed.): Griffin, London.

Examples

```
data(HAV)
CARE1.print(HAV)
```

Congenital

Congenital anomaly data

Description

Infants born data with a specific congenital anomaly in Massachusetts

Usage

```
data(Congenital)
```

Format

A five lists data with 31 variables.

Details

An epidemiological data with five distinct types of sources, which contains: obstetric records (183 cases); other hospital records (215 cases); list maintained by state Department of Public Health (36 cases); list maintained by state Department of Mental Health (263cases), and records by special schools (252 cases). The total number of cases identified was 537.

References

- Fienberg, S. E. (1972). The multiple recapture census for closed populations and incomplete 2k contingency tables. *Biometrika*, 59: 591-603.
- Wittes, J. T., Colton, T., and Sidel, V. W. (1974). Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *Journal of Chronic Diseases*, 27: 25.

Examples

```
data(Congenital)
CARE1.print(Congenital)
```

Diabetes

Diabetes data

Description

Data set on diabetes in a community in Italy.

Usage

```
data(Diabetes)
```

Format

A four lists data with 15 variables.

Details

An epidemiological data with four distinct types of sources, which contains: diabetic clinic and/or family physician visits (1754 cases); hospital discharges (452 cases); prescriptions (1135 cases), and purchases of reagent strips and insulin syringes (173 cases). A total of 2069 cases were identified.

References

Bruno G, LaPorte R, Merletti F, Biggeri A, McCarty D, Pagano G. 1994. National diabetes programs. Application of capture-recapture to count diabetes? *Diabetes Care* 17(6): 548.

Examples

```
data(Diabetes)
CARE1.print(Diabetes)
```

estN

*Estimating population size via sample coverage***Description**

Population size estimation based on sample coverage approach for closed capture-recapture models.

Usage

```
estN(z, method="Indep", se=FALSE, nboot=200)
```

Arguments

z	the vector of capture histories or ascertainment records.
method	the method employed to estimate population size. method="Indep": population size estimate for independent samples; method="HSC": population size estimate for sufficiently high sample coverage cases; method="LSC": one-step population size estimate for low sample coverage cases.
se	should calculus bootstrap standard error?
nboot	the number of bootstrap resampling times.

Value

population size estimator

Note

- se estimated standard error of the population size estimation based on bootstrap replications, which might vary with trials.
- cil confidence interval lower limit (using a log-transformation).
- ciu 95% confidence interval upper limit (using a log-transformation).
- \hat{N}_{Indep} population size estimate for independent samples; see Equation (9) of Chao et al. (2001).
- \hat{N}_{HSC} Population size estimate for sufficiently high sample coverage cases; see Equation (12) of Chao et al. (2001).
- \hat{N}_{LSC} One-step population size estimate for low sample coverage cases; see Equation (13) of Chao et al. (2001). This estimator is suggested for use when the estimated s.e. of \hat{N}_{HSC} is relatively large.

References

Chao, A., and Tsay, P. K. (1998). A sample coverage approach to multiple-system estimation with application to census undercount. *Journal of the American Statistical Association*, 93: 283-293.

Chao, A., Tsay, P. K., Lin, S. H., Shau, W. Y., and Chao, D. Y. (2001). The applications of capture-recapture models to epidemiological data. *Statistics in Medicine*, 20: 3123-3157.

Examples

```
data(HAV)
estN(HAV,method="LSC",se=TRUE,nboot=200)
```

estN.n	<i>Number of identified cases in each records</i>
--------	---

Description

Number of identified cases in each records

Usage

```
estN.n(z)
```

Arguments

`z` the vector of capture histories or ascertainment records.

Value

Number of identified cases.

Examples

```
data(HAV)
estN.n(HAV)
```

estN.pair	<i>Pairwise estimation</i>
-----------	----------------------------

Description

Estimating population size based on any pair of samples

Usage

```
estN.pair(z)
```

Arguments

`z` the vector of capture histories or ascertainment records.

Value

estimates based on any pair of samples

Note

1. Refer to Seber (1982, pages 59 and 60) for the Petersen estimator and the Chapman estimator as well as s.e. formula.
2. A log-transformation is used to obtain the confidence interval so that the lower limit is always greater than the number of ascertained. Refer to Chao (1987, Biometrics,43, 783-791) for the construction of the confidence interval.

References

Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics*, 43: 783-791.

Chao, A., Tsay, P. K., Lin, S. H., Shau, W. Y., and Chao, D. Y. (2001). The applications of capture-recapture models to epidemiological data. *Statistics in Medicine*, 20: 3123-3157.

Seber, G. A. F. (1982). *The estimation of animal abundance* (2 ed.): Griffin, London.

Examples

```
data(HAV)
estN.pair(HAV)
```

estN.para

Parameter estimates

Description

Report the estimated mean probabilities depending on the estimate of N, and estimated coefficient of covariation (CCV) depending on the estimate of N.

Usage

```
estN.para(z, nhat)
```

Arguments

z the vector of capture histories or ascertainment records.
nhat the estimation of population size.

Value

Parameter estimates

Note

- u estimated mean probabilities depending on the estimate of N.
- r estimated coefficient of covariation (CCV) depending on the estimate of N.

References

Chao, A., Tsay, P. K., Lin, S. H., Shau, W. Y., and Chao, D. Y. (2001). The applications of capture-recapture models to epidemiological data. *Statistics in Medicine*, 20: 3123-3157.

Examples

```
data(HAV)
nhat <- estN(HAV,method="LSC",se=FALSE)
estN.para(HAV, nhat)
```

estN.stat

Statistics of sample coverage approach

Description

Report some statistics of sample coverage approach.

Usage

```
estN.stat(z)
```

Arguments

z the vector of capture histories or ascertainment records.

Value

- M number of individuals ascertained in at least one list.
- D the average of the number of individuals listed in the combination of two lists omitting the third.
- C^{\wedge} sample coverage estimate, see Equation (14) of Chao and Tsay (1998).

References

Chao, A., and Tsay, P. K. (1998). A sample coverage approach to multiple-system estimation with application to census undercount. *Journal of the American Statistical Association*, 93: 283-293.

Examples

```
data(HAV)
estN.stat(HAV)
```

HAV

Hepatitis A virus data

Description

An outbreak of the hepatitis A virus (HAV) occurred in and around a college in northern Taiwan from April to July 1995.

Usage

`data(HAV)`

Format

A three lists data with 7 variables.

Details

Cases of students in that college were ascertained by three sources: (i) P-list, records based on a serum test taken by the Institute of Preventive Medicine, Department of Health of Taiwan (135 cases); (ii) Q-list, local hospital records reported by the National Quarantine Service (122 cases); (iii) E-list, records collected by epidemiologists (126 cases). Merging the three lists by eliminating duplicate records resulted in 271 ascertained cases.

References

Chao, D., Shau, W., Lu, C., Chen, K., Chu, C., Shu, H., and Horng, C. (1997). A large outbreak of hepatitis A in a college school in Taiwan: associated with contaminated food and water dissemination. *Epidemiology Bulletin*, Department of Health, Executive Yuan, Taiwan Government.

Examples

`data(HAV)`

Index

*Topic **datasets**

Congenital, [4](#)

Diabetes, [5](#)

HAV, [10](#)

as.record, [2](#)

CARE1 (CARE1-package), [2](#)

CARE1-package, [2](#)

CARE1.print, [3](#)

Congenital, [4](#)

Diabetes, [5](#)

estN, [6](#)

estN.n, [7](#)

estN.pair, [7](#)

estN.para, [8](#)

estN.stat, [9](#)

HAV, [10](#)