

# Package ‘Frames2’

May 7, 2026

**Type** Package

**Title** Estimation in Dual Frame Surveys

**Version** 0.2.1

**Date** 2015-12-11

**Author** Antonio Arcos <arcos@ugr.es>, Maria del Mar Rueda <mrueda@ugr.es>, Maria Giovanna Ranalli <giovanna.ranalli@stat.unipg.it> and David Molina <dmolinam@ugr.es>

**Maintainer** David Molina <dmolinam@ugr.es>

**Description** Point and interval estimation in dual frame surveys. In contrast to classic sampling theory, where only one sampling frame is considered, dual frame methodology assumes that there are two frames available for sampling and that, overall, they cover the entire target population. Then, two probability samples (one from each frame) are drawn and information collected is suitably combined to get estimators of the parameter of interest.

**License** GPL (>= 2)

**Imports** sampling, MASS, nnet

**RoxygenNote** 5.0.1

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2015-12-12 09:49:54

## Contents

BKA . . . . .	2
CalDF . . . . .	5
CalSF . . . . .	8
Compare . . . . .	11
CovHT . . . . .	13
Dat . . . . .	14
DatA . . . . .	16

DatB . . . . .	17
DatMA . . . . .	18
DatMB . . . . .	19
DatPopM . . . . .	20
Domains . . . . .	21
FB . . . . .	22
Hartley . . . . .	24
HT . . . . .	26
JackBKA . . . . .	27
JackCalDF . . . . .	30
JackCalSF . . . . .	33
JackFB . . . . .	36
JackHartley . . . . .	38
JackMLCDF . . . . .	40
JackMLCDW . . . . .	43
JackMLCSW . . . . .	46
JackMLDF . . . . .	49
JackMLDW . . . . .	52
JackMLSW . . . . .	54
JackPEL . . . . .	57
JackPML . . . . .	59
JackSFRR . . . . .	62
MLCDF . . . . .	64
MLCDW . . . . .	67
MLCSW . . . . .	69
MLDF . . . . .	72
MLDW . . . . .	74
MLSW . . . . .	76
PEL . . . . .	79
Pik1A . . . . .	82
Pik1B . . . . .	82
PML . . . . .	83
SFRR . . . . .	85
VarHT . . . . .	87
WeightsCalDF . . . . .	89
WeightsCalSF . . . . .	91

<b>Index</b>	<b>94</b>
--------------	-----------

---

BKA

*Bankier-Kalton-Anderson estimator*

---

### Description

Produces estimates for population total and mean using the Bankier-Kalton-Anderson estimator from survey data obtained from a dual frame sampling design. Confidence intervals are also computed, if required.

**Usage**

BKA(ysA, ysB, pi\_A, pi\_B, pik\_ab\_B, pik\_ba\_A, domains\_A, domains\_B,  
conf\_level = NULL)

**Arguments**

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable(s) of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable(s) of interest from $s_B$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

**Details**

BKA estimator of population total is given by

$$\hat{Y}_{BKA} = \sum_{i \in s_A} \tilde{d}_i^A y_i + \sum_{i \in s_B} \tilde{d}_i^B y_i$$

where  $\tilde{d}_i^A = \begin{cases} d_i^A & \text{if } i \in a \\ (1/d_i^A + 1/d_i^B)^{-1} & \text{if } i \in ab \end{cases}$  and  $\tilde{d}_i^B = \begin{cases} d_i^B & \text{if } i \in b \\ (1/d_i^A + 1/d_i^B)^{-1} & \text{if } i \in ba \end{cases}$  being  $d_i^A$  and  $d_i^B$  the design weights, obtained as the inverse of the first order inclusion probabilities, that is,  $d_i^A = 1/\pi_i^A$  and  $d_i^B = 1/\pi_i^B$ .

To estimate variance of this estimator, one uses following approach proposed by Rao and Skinner (1996)

$$\hat{V}(\hat{Y}_{BKA}) = \hat{V}\left(\sum_{i \in s_A} \tilde{z}_i^A\right) + \hat{V}\left(\sum_{i \in s_B} \tilde{z}_i^B\right)$$

with  $\tilde{z}_i^A = \delta_i(a)y_i + (1 - \delta_i(a))y_i\pi_i^A/(\pi_i^A + \pi_i^B)$  and  $\tilde{z}_i^B = \delta_i(b)y_i + (1 - \delta_i(b))y_i\pi_i^B/(\pi_i^A + \pi_i^B)$ , being  $\delta_i(a)$  and  $\delta_i(b)$  the indicator variables for domain  $a$  and domain  $b$ , respectively. If both first

and second order probabilities are known, variances and covariances involved in calculation of  $\hat{\beta}$  and  $\hat{V}(\hat{Y}_{FB})$  are estimated using functions VarHT and CovHT, respectively. If only first order probabilities are known, variances are estimated using Deville's method and covariances are estimated using following expression

$$\widehat{Cov}(\hat{X}, \hat{Y}) = \frac{\hat{V}(X + Y) - \hat{V}(X) - \hat{V}(Y)}{2}$$

### Value

BKA returns an object of class "EstimatorDF" which is a list with, at least, the following components:

Call                    the matched call.  
 Est                    total and mean estimation for main variable(s).  
 VarEst                variance estimation for main variable(s).

If parameter `conf_level` is different from NULL, object includes component

ConfInt                total and mean estimation and confidence intervals for main variables(s).

In addition, components TotDomEst and MeanDomEst are available when estimator is based on estimators of the domains. Component Param shows value of parameters involved in calculation of the estimator (if any). By default, only Est component (or ConfInt component, if parameter `conf_level` is different from NULL) is shown. It is possible to access to all the components of the objects by using function `summary`.

### References

- Bankier, M. D. (1986) *Estimators Based on Several Stratified Samples With Applications to Multiple Frame Surveys*. Journal of the American Statistical Association, Vol. 81, 1074 - 1079.
- Kalton, G. and Anderson, D. W. (1986) *Sampling Rare Populations*. Journal of the Royal Statistical Society, Ser. A, Vol. 149, 65 - 82.
- Rao, J. N. K. and Skinner, C. J. (1996) *Estimation in Dual Frame Surveys with Complex Designs*. Proceedings of the Survey Method Section, Statistical Society of Canada, 63 - 68.
- Skinner, C. J. and Rao, J. N. K. (1996) *Estimation in Dual Frame Surveys with Complex Designs*. Journal of the American Statistical Association, Vol. 91, 433, 349 - 356.

### See Also

[JackBKA](#)

### Examples

```
data(Data)
data(DatB)
data(Pik1A)
data(Pik1B)
```

```
#Let calculate BKA estimator for population total for variable Leisure
BKA(Data$Lei, DatB$Lei, Pik1A, Pik1B, Data$ProbB, DatB$ProbA,
```

```

Data$Domain, DatB$Domain)

#Now, let calculate BKA estimator and a 90% confidence interval for population
#total for variable Feeding considering only first order inclusion probabilities
BKA(Data$Feed, DatB$Feed, DatA$ProbA, DatB$ProbB, DatA$ProbB,
DatB$ProbA, Data$Domain, DatB$Domain, 0.90)

```

CalDF

*DF calibration estimator***Description**

Produces estimates for population totals and means using the DF calibration estimator from survey data obtained from a dual frame sampling design. Confidence intervals are also computed, if required.

**Usage**

```

CalDF(ysA, ysB, pi_A, pi_B, domains_A, domains_B, N_A = NULL, N_B = NULL,
N_ab = NULL, xsAFrameA = NULL, xsBFrameA = NULL, xsAFrameB = NULL, xsBFrameB = NULL,
xsT = NULL, XA = NULL, XB = NULL, X = NULL, met = "linear", conf_level = NULL)

```

**Arguments**

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable(s) of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable(s) of interest from $s_B$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of length $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of length $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	(Optional) A numeric value indicating the size of frame A.
N_B	(Optional) A numeric value indicating the size of frame B.
N_ab	(Optional) A numeric value indicating the size of the overlap domain.
xsAFrameA	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .

xsBFrameA	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.
xsAFrameB	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.
xsBFrameB	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
xsT	(Optional) A numeric vector of length $n$ or a numeric matrix or data frame of dimensions $n \times m_T$ , with $m_T$ the number of auxiliary variables in both frames, containing auxiliary information for all units in the entire sample $s = s_A \cup s_B$ .
XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
X	(Optional) A numeric value or vector of length $m_T$ , with $m_T$ the number of auxiliary variables in both frames, indicating the population totals for the auxiliary variables considered in both frames.
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

## Details

DF calibration estimator of population total is given by

$$\hat{Y}_{CalDF} = \hat{Y}_a + \hat{\eta}\hat{Y}_{ab} + \hat{Y}_b + (1 - \hat{\eta})\hat{Y}_{ba}$$

where  $\hat{Y}_a = \sum_{i \in s_a} \tilde{d}_i y_i$ ,  $\hat{Y}_{ab} = \sum_{i \in s_{ab}} \tilde{d}_i y_i$ ,  $\hat{Y}_b = \sum_{i \in s_b} \tilde{d}_i y_i$  and  $\hat{Y}_{ba} = \sum_{i \in s_{ba}} \tilde{d}_i y_i$ , with  $\tilde{d}_i$  calibration weights which are calculated having into account a different set of constraints, depending on the case. For instance, if  $N_A, N_B$  and  $N_{ab}$  are all known and no other auxiliary information is available, calibration constraints are

$$\sum_{i \in s_a} \tilde{d}_i = N_a, \sum_{i \in s_{ab}} \tilde{d}_i = N_{ab}, \sum_{i \in s_{ba}} \tilde{d}_i = N_{ba}, \sum_{i \in s_b} \tilde{d}_i = N_b$$

Optimal value for  $\hat{\eta}$  to minimize variance of the estimator is given by  $\hat{V}(\hat{N}_{ba}) / (\hat{V}(\hat{N}_{ab}) + \hat{V}(\hat{N}_{ba}))$ . If both first and second order probabilities are known, variances are estimated using function VarHT. If only first order probabilities are known, variances are estimated using Deville's method.

Function covers following scenarios:

- There is not any additional auxiliary variable
  - $N_A, N_B$  and  $N_{ab}$  unknown
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_{ab}$  known and  $N_A$  and  $N_B$  unknown
  - $N_A, N_B$  and  $N_{ab}$  known
- At least, information about one additional auxiliary variable is available
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_{ab}$  known and  $N_A$  and  $N_B$  unknown
  - $N_A, N_B$  and  $N_{ab}$  known

To obtain an estimator of the variance for this estimator, one can use Deville's expression

$$\hat{V}(\hat{Y}_{CalDF}) = \frac{1}{1 - \sum_{k \in s} a_k^2} \sum_{k \in s} (1 - \pi_k) \left( \frac{e_k}{\pi_k} - \sum_{l \in s} a_l \frac{e_l}{\pi_l} \right)^2$$

where  $a_k = (1 - \pi_k) / \sum_{l \in s} (1 - \pi_l)$  and  $e_k$  are the residuals of the regression with auxiliary variables as regressors.

## Value

CalDF returns an object of class "EstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	total and mean estimation for main variable(s).
VarEst	variance estimation for main variable(s).

If parameter `conf_level` is different from NULL, object includes component

ConfInt	total and mean estimation and confidence intervals for main variables(s).
---------	---

In addition, components `TotDomEst` and `MeanDomEst` are available when estimator is based on estimators of the domains. Component `Param` shows value of parameters involved in calculation of the estimator (if any). By default, only `Est` component (or `ConfInt` component, if parameter `conf_level` is different from NULL) is shown. It is possible to access to all the components of the objects by using function `summary`.

## References

- Ranalli, M. G., Arcos, A., Rueda, M. and Teodoro, A. (2013) *Calibration estimation in dual frame surveys*. arXiv:1312.0761 [stat.ME]
- Deville, J. C., Sarndal, C. E. (1992) *Calibration estimators in survey sampling*. Journal of the American Statistical Association, 87, 376 - 382

## See Also

[JackCalDF](#)

## Examples

```

data(DatA)
data(DatB)
data(PiklA)
data(PiklB)

#Let calculate DF calibration estimator for variable Feeding, without
#considering any auxiliary information
CalDF(DatA$Feed, DatB$Feed, PiklA, PiklB, DatA$Domain, DatB$Domain)

#Now, let calculate DF calibration estimator for variable Clothing when the frame
#sizes and the overlap domain size are known
CalDF(DatA$Clo, DatB$Clo, PiklA, PiklB, DatA$Domain, DatB$Domain,
N_A = 1735, N_B = 1191, N_ab = 601)

#Finally, let calculate DF calibration estimator and a 90% confidence interval
#for population total for variable Feeding, considering Income as auxiliary variable in
#frame A and Metres2 as auxiliary variable in frame B and with frame sizes and overlap
#domain size known.
CalDF(DatA$Feed, DatB$Feed, PiklA, PiklB, DatA$Domain, DatB$Domain,
N_A = 1735, N_B = 1191, N_ab = 601, xsAframeA = DatA$Inc, xsBframeA = DatB$Inc,
xsAframeB = DatA$M2, xsBframeB = DatB$M2, XA = 4300260, XB = 176553,
conf_level = 0.90)

```

---

CalSF

*SF calibration estimator*

---

## Description

Produces estimates for population totals and means using the SF calibration estimator from survey data obtained from a dual frame sampling design. Confidence intervals are also computed, if required.

## Usage

```

CalSF(ysA, ysB, pi_A, pi_B, pik_ab_B, pik_ba_A, domains_A, domains_B, N_A = NULL,
N_B = NULL, N_ab = NULL, xsAframeA = NULL, xsBframeA = NULL, xsAframeB = NULL,
xsBframeB = NULL, xsT = NULL, XA = NULL, XB = NULL, X = NULL, met = "linear",
conf_level = NULL)

```

## Arguments

<code>ysA</code>	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable(s) of interest from $s_A$ .
<code>ysB</code>	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable(s) of interest from $s_B$ .
<code>pi_A</code>	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .

pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	(Optional) A numeric value indicating the size of frame A
N_B	(Optional) A numeric value indicating the size of frame B
N_ab	(Optional) A numeric value indicating the size of the overlap domain
xsAFrameA	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsBFrameA	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.
xsAFrameB	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.
xsBFrameB	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
xsT	(Optional) A numeric vector of length $n$ or a numeric matrix or data frame of dimensions $n \times m_T$ , with $m_T$ the number of auxiliary variables in both frames, containing auxiliary information for all units in the entire sample $s = s_A \cup s_B$ .
XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
X	(Optional) A numeric value or vector of length $m_T$ , with $m_T$ the number of auxiliary variables in both frames, indicating the population totals for the auxiliary variables considered in both frames.
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".

conf\_level (Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

### Details

SF calibration estimator of population total is given by

$$\hat{Y}_{CalSF} = \hat{Y}_a + \hat{Y}_{ab} + \hat{Y}_b$$

where  $\hat{Y}_a = \sum_{i \in s_a} \tilde{d}_i y_i$ ,  $\hat{Y}_{ab} = \sum_{i \in (s_{ab} \cup s_{ba})} \tilde{d}_i y_i$  and  $\hat{Y}_b = \sum_{i \in s_b} \tilde{d}_i y_i$ , with  $\tilde{d}_i$  calibration weights which are calculated having into account a different set of constraints, depending on the case. For instance, if  $N_A$ ,  $N_B$  and  $N_{ab}$  are known and no other auxiliary information is available, calibration constraints are

$$\sum_{i \in s_a} \tilde{d}_i = N_a, \quad \sum_{i \in s_{ab} \cup s_{ba}} \tilde{d}_i = N_{ab}, \quad \sum_{i \in s_{ba}} \tilde{d}_i = N_{ba}$$

Function covers following scenarios:

- There is not any additional auxiliary variable
  - $N_A$ ,  $N_B$  and  $N_{ab}$  unknown
  - $N_{ab}$  known and  $N_A$  and  $N_B$  unknown
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_A$ ,  $N_B$  and  $N_{ab}$  known
- At least, information about one additional auxiliary variable is available
  - $N_{ab}$  known and  $N_A$  and  $N_B$  unknown
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_A$ ,  $N_B$  and  $N_{ab}$  known

To obtain an estimator of the variance for this estimator, one can use Deville's expression

$$\hat{V}(\hat{Y}_{CalSF}) = \frac{1}{1 - \sum_{k \in s} a_k^2} \sum_{k \in s} (1 - \pi_k) \left( \frac{e_k}{\pi_k} - \sum_{l \in s} a_l \frac{e_l}{\pi_l} \right)^2$$

where  $a_k = (1 - \pi_k) / \sum_{l \in s} (1 - \pi_l)$  and  $e_k$  are the residuals of the regression with auxiliary variables as regressors.

### Value

CalSF returns an object of class "EstimatorDF" which is a list with, at least, the following components:

Call the matched call.  
 Est total and mean estimation for main variable(s).  
 VarEst variance estimation for main variable(s).

If parameter conf\_level is different from NULL, object includes component

ConfInt total and mean estimation and confidence intervals for main variables(s).

In addition, components TotDomEst and MeanDomEst are available when estimator is based on estimators of the domains. Component Param shows value of parameters involved in calculation of the estimator (if any). By default, only Est component (or ConfInt component, if parameter conf\_level is different from NULL) is shown. It is possible to access to all the components of the objects by using function summary.

## References

Ranalli, M. G., Arcos, A., Rueda, M. and Teodoro, A. (2013) *Calibration estimation in dual frame surveys*. arXiv:1312.0761 [stat.ME]

Deville, J. C., Sarndal, C. E. (1992) *Calibration estimators in survey sampling*. Journal of the American Statistical Association, 87, 376 - 382

## See Also

[JackCalSF](#)

## Examples

```
data(DatA)
data(DatB)
data(Pik1A)
data(Pik1B)

#Let calculate SF calibration estimator for variable Clothing, without
#considering any auxiliary information
CalSF(DatA$Clo, DatB$Clo, Pik1A, Pik1B, DatA$ProbB, DatB$ProbA,
DatA$Domain, DatB$Domain)

#Now, let calculate SF calibration estimator for variable Leisure when the frame
#sizes and the overlap domain size are known
CalSF(DatA$Lei, DatB$Lei, Pik1A, Pik1B, DatA$ProbB, DatB$ProbA, DatA$Domain,
DatB$Domain, N_A = 1735, N_B = 1191, N_ab = 601)

#Finally, let calculate SF calibration estimator and a 90% confidence interval
#for population total for variable Feeding, considering Income and Metres2 as auxiliary
#variables and with frame sizes and overlap domain size known.
CalSF(DatA$Feed, DatB$Feed, Pik1A, Pik1B, DatA$ProbB, DatB$ProbA, DatA$Domain,
DatB$Domain, N_A = 1735, N_B = 1191, N_ab = 601, xsAFrameA = DatA$Inc,
xsBFrameA = DatB$Inc, xsAFrameB = DatA$M2, xsBFrameB = DatB$M2,
XA = 4300260, XB = 176553, conf_level = 0.90)
```

---

Compare

*Summary of estimators*

---

## Description

Returns all possible estimators that can be computed according to the information provided

**Usage**

```
Compare(ysA, ysB, pi_A, pi_B, domains_A, domains_B, pik_ab_B = NULL, pik_ba_A = NULL,
N_A = NULL, N_B = NULL, N_ab = NULL, xsAframeA = NULL, xsBframeA = NULL,
xsAframeB = NULL, xsBframeB = NULL, XA = NULL, XB = NULL, met = "linear",
conf_level = NULL)
```

**Arguments**

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable(s) of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable(s) of interest from $s_B$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of length $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of length $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
pik_ab_B	(Optional) A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	(Optional) A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
N_A	(Optional) A numeric value indicating the size of frame A.
N_B	(Optional) A numeric value indicating the size of frame B.
N_ab	(Optional) A numeric value indicating the size of the overlap domain.
xsAframeA	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsBframeA	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.
xsAframeB	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.
xsBframeB	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .

XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

### Examples

```
data(DataA)
data(DataB)
data(PiklA)
data(PiklB)
```

```
Compare(DataA$Feed, DataB$Feed, PiklA, PiklB, DataA$Domain, DataB$Domain)
```

---

CovHT

*Covariance estimator between two Horvitz - Thompson estimators*


---

### Description

Computes the covariance estimator between two Horvitz - Thompson estimators of population total from survey data obtained from a single stage sampling design

### Usage

```
CovHT(y, x, pikl)
```

### Arguments

y	A numeric vector of size n containing information about first variable of interest in the sample
x	A numeric vector of size n containing information about second variable of interest in the sample
pikl	A square numeric matrix of dimension n containing first and second order inclusion probabilities for units included in the sample

### Details

Covariance estimator between two Horvitz - Thompson estimators of population total is given by

$$\hat{Cov}(\hat{Y}_{HT}, \hat{X}_{HT}) = \sum_{k \in s} \sum_{l \in s} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k}{\pi_k} \frac{x_l}{\pi_l}$$

**Value**

A numeric value representing covariance estimator between two Horvitz - Thompson estimators for population total for considered values

**References**

Horvitz, D. G. and Thompson, D. J. (1952) *A generalization of sampling without replacement from a finite universe*. Journal of the American Statistical Association, 47, 663 - 685 @references Sarndal, C. E., Swensson, B. and Wretman, J. (1992) *Model Assisted Survey Sampling*. Springer-Verlag. New York.

**See Also**

[HT VarHT](#)

**Examples**

```
##### Example 1 #####
Indicators <- c(1, 2, 3, 4, 5)
X <- c(13, 18, 20, 14, 9)
Y <- c(2, 0.5, 1.2, 3.3, 2)
#Let draw two simple random samples without replacement of size 2
s <- sample(Indicators, 2)
sX <- X[s]
sY <- Y[s]
#Now, let calculate the associated probability matrix with first and
#second order inclusion probabilities
Ps <- matrix(c(0.4,0.2, 0.2,0.4), 2, 2)
CovHT(sX, sY, Ps)

##### Example 2 #####
data(DataA)
attach(DataA)
data(Pik1A)
#Let calculate Horvitz - Thompson estimator for total of variable Clothing in Frame A.
HT(Clo, ProbA)
#Let calculate Horvitz - Thompson estimator for total of variable Feeding in Frame A.
HT(Feed, ProbA)
#And now, let compute the covariance between the previous estimators
CovHT(Clo, Feed, Pik1A)
```

---

Dat

*Joint sample database*

---

**Description**

This dataset contains some variables coming from a real dual frame survey conducted in 2013 in Andalusia (Spain) by a scientific institute specialized in social topics. With this dataset it is intended to show how to properly split a joint dual frame sample into subsamples, so functions of Frame2 can be used.

**Usage**

Dat

**Format**

**Drawnby** Indicates whether individual was selected in the landline sample(1) or in the cell phone sample(2).

**Stratum** Indicates the stratum each individual belongs to. For individuals selected in cell phone sample, value of this variable is NA.

**Opinion** Response of the individual to the question: Do you think that immigrants currently living in Andalusia are quite a lot? 1 represents "yes" and 0 represents "no".

**Landline** Indicates whether individual has a landline (1) or not (0).

**Cell** Indicates whether individual has a cell phone(1) or not(0).

**ProbLandline** First order inclusion probability of reaching the individual by landline.

**ProbCell** First order inclusion probability of reaching the individual by cell phone.

**Income** Monthly income (in euros) of the individual.

**Details**

The survey was based on two frames: a landline frame and a cell phone frame. Landline frame was stratified by province and simple random sampling without replacement was considered in cell phone frame. The size of the whole sample was  $n = 2402$ . Total of the variable Income in the whole population is  $X_{Income} = 12686232063$ .

**Examples**

```
data(Dat)
attach(Dat)

#We are going to split dataset Dat into two new datasets, each
#one corresponding to a frame: frame containing individuals
#using landline and frame containing individuals using cell phone.

FrameLandline <- Dat[Landline == 1,]
FrameCell <- Dat[Cell == 1,]

#Equally, we can split the original dataset in three new different
#datasets, each one corresponding to one domain: first domain containing
#individuals using only landline, second domain containing individuals
#using only cell phone and the third domain containing individuals
#using both landline and cell phone.

DomainLandline <- Dat[Landline == 1 & Cell == 0,]
DomainCell <- Dat[Landline == 0 & Cell == 1,]
DomainBoth <- Dat[Landline == 1 & Cell == 1,]

#From the domain datasets, we can build frame datasets
```

```
FrameLandline <- rbind(DomainLandline, DomainBoth)
FrameCell <- rbind(DomainCell, DomainBoth)
```

---

 Data

*Database of household expenses for frame A*


---

### Description

This dataset contains some variables regarding household expenses for a sample of 105 households selected from a list of landline phones (let say, frame A) in a particular city in a specific month.

### Usage

Data

### Format

**Domain** A string indicating the domain each household belongs to. Possible values are "a" if household belongs to domain a or "ab" if household belongs to overlap domain.

**Feed** Feeding expenses (in euros) at the household

**Clo** Clothing expenses (in euros) at the household

**Lei** Leisure expenses (in euros) at the household

**Inc** Household income (in euros). Values for this variable are only available for households included in frame A. For households included in domain b, value of this variable is set to 0.

**Tax** Household municipal taxes (in euros) paid. Values for this variable are only available for households included in frame A. For households included in domain b, value of this variable is set to 0.

**M2** Square meters of the house. Values for this variable are only available for households included in frame B. For households included in domain a, value of this variable is set to 0.

**Size** Household size. Values for this variable are only available for households included in frame B. For households included in domain a, value of this variable is set to 0.

**ProbA** First order inclusion probability in frame A. This probability is 0 for households included in domain b.

**ProbB** First order inclusion probability in frame B. This probability is 0 for households included in domain a.

**Stratum** A numeric value indicating the stratum each household belongs to.

### Details

The sample, of size  $n_A = 105$ , has been drawn from a population of  $N_A = 1735$  households with landline phone according to a stratified random sampling. Population units were divided in 6 different strata. Population sizes of these strata are  $N_A^h = (727, 375, 113, 186, 115, 219)$ .  $N_{ab} = 601$  of the households composing the population have, also, mobile phone. On the other hand, frame totals for auxiliary variables in this frame are  $X_{Income}^A = 4300260$  and  $X_{Axes}^A = 215577$ .

**See Also**[Pik1A](#)**Examples**

```
data(DataA)
attach(DataA)
#Let perform a brief descriptive analysis for the three main variables
param <- data.frame(Feed, Clo, Lei)
summary (param)
hist (Feed)
hist (Clo)
hist (Lei)
```

---

DatB

*Database of household expenses for frame B*

---

**Description**

This dataset contains some variables regarding household expenses for a sample of 135 households selected from a list of mobile phones (let say, frame B) in a particular city in a specific month.

**Usage**

DatB

**Format**

- Domain** A string indicating the domain each household belongs to. Possible values are "b" if household belongs to domain b or "ba" if household belongs to overlap domain.
- Feed** Feeding expenses (in euros) at the household
- Clo** Clothing expenses (in euros) at the household
- Lei** Leisure expenses (in euros) at the household
- Inc** Household income (in euros). Values for this variable are only available for households included in frame A. For households included in domain b, value of this variable is set to 0.
- Tax** Household municipal taxes (in euros) paid. Values for this variable are only available for households included in frame A. For households included in domain b, value of this variable is set to 0.
- M2** Square meters of the house. Values for this variable are only available for households included in frame B. For households included in domain a, value of this variable is set to 0.
- Size** Household size. Values for this variable are only available for households included in frame B. For households included in domain a, value of this variable is set to 0.
- ProbA** First order inclusion probability in frame A. This probability is 0 for households included in domain b.
- ProbB** First order inclusion probability in frame B. This probability is 0 for households included in domain a.

**Details**

The sample, of size  $n_B = 135$ , has been drawn from a population of  $N_B = 1191$  households with mobile phone according to a simple random sampling without replacement design.  $N_{ab} = 601$  of these households have, also, landline phone. On the other hand, frame totals for auxiliary variables in this frame are  $X_{Metres2}^B = 176553$  and  $X_{Size}^B = 3529$

**See Also**

[Pik1B](#)

**Examples**

```
data(DatB)
attach(DatB)
#Let perform a brief descriptive analysis for the three main variables
param <- data.frame(Feed, Clo, Lei)
summary (param)
hist (Feed)
hist (Clo)
hist (Lei)
```

---

DatMA

*Database of students' program choice for frame A*

---

**Description**

This dataset contains some variables regarding the program choice for a sample of 180 students included in the sampling frame A.

**Usage**

DatMA

**Format**

**Id\_Pop** An integer from 1 to  $N$ , with  $N$  the number of students in the whole population, identifying the student within the population.

**Id\_Frame** An integer from 1 to  $N_A$ , with  $N_A$  the number of students in the frame, identifying the student within the frame.

**Prog** A factor with three categories (academic, general and vocation) indicating the program choice of the student.

**Ses** An ordinal factor with three categories (low, middle and high) indicating the socio-economical status of the student.

**Read** A number indicating the mark of the student in a reading test.

**Write** A number indicating the mark of the student in a writing test.

**Sch\_Size** A number indicating the size of the school the students belongs to.

**Domain** A string indicating the domain each student belongs to. Possible values are "a" if student belongs to domain a or "ab" if student belongs to overlap domain.

**ProbA** First order inclusion probability in frame A.

**ProbB** First order inclusion probability in frame B. This probability is 0 for students included in domain a.

### Details

The sample, of size  $n_A = 180$ , has been drawn from a population of  $N_A = 5500$  students according to a proportional-to-size sampling design according to the size of the school. So, students attending bigger schools have a higher probability of being selected in the sample.  $N_{ab} = 2000$  of the students composing the population belongs also to frame B.

### See Also

[DatPopM](#)

### Examples

```
data(DatMA)
attach(DatMA)
#Let perform a brief descriptive analysis for the main variable
summary (Prog)
#And let do the same for the numerical auxiliary variables Read and Write
summary(Read)
summary(Write)
```

---

DatMB

*Database of students' program choice for frame B*

---

### Description

This dataset contains some variables regarding the program choice for a sample of 232 students included in the sampling frame B.

### Usage

DatMB

### Format

**Id\_Pop** An integer from 1 to  $N$ , with  $N$  the number of students in the whole population, identifying the student within the population.

**Id\_Frame** An integer from 1 to  $N_B$ , with  $N_B$  the number of students in the frame, identifying the student within the frame.

**Prog** A factor with three categories (academic, general and vocation) indicating the program choice of the student.

**Ses** An ordinal factor with three categories (low, middle and high) indicating the socio-economical status of the student.

**Read** A number indicating the mark of the student in a reading test.

**Write** A number indicating the mark of the student in a writing test.

**Sch\_Size** A number indicating the size of the school the students belongs to.

**Domain** A string indicating the domain each student belongs to. Possible values are "b" if student belongs to domain b or "ba" if student belongs to overlap domain.

**ProbA** First order inclusion probability in frame A. This probability is 0 for students included in domain b.

**ProbB** First order inclusion probability in frame B.

### Details

The sample, of size  $n_B = 232$ , has been drawn from a population of  $N_B = 6500$  students according to a simple random sampling design.  $N_{ab} = 2000$  of the students composing the population belongs also to frame A.

### See Also

[DatPopM](#)

### Examples

```
data(DatMB)
attach(DatMB)
#Let perform a brief descriptive analysis for the main variable
summary (Prog)
#And let do the same for the numerical auxiliary variables Read and Write
summary(Read)
summary(Write)
```

---

DatPopM

*Database of auxiliary information for the whole population of students*

---

### Description

This dataset contains population information about the auxiliary variables of the population of students

### Usage

DatPopM

**Format**

**Ses** An ordinal factor with three categories (low, middle and high) indicating the socio-economical status of the student.

**Read** A number indicating the mark of the student in a reading test.

**Write** A number indicating the mark of the student in a writing test.

**Domain** A string indicating the domain each student belongs to. Possible values are "a" if student belongs to domain a, "b" if student belongs to domain b or "ab" if student belongs to overlap domain.

**Details**

The population size is  $N = 10000$ .

**See Also**

[DatMA](#) [DatMB](#)

**Examples**

```
data(DatPopM)
attach(DatPopM)
#Let perform a brief descriptive analysis for the three auxiliary variables
summary (Ses)
summary(Read)
summary(Write)
```

---

Domains

*Domains*

---

**Description**

Given a main vector, an auxiliary one and a value of the latter, identifies positions of the auxiliary vector corresponding to values other than the given one. Then, turns zero values of the main vector corresponding to these positions.

**Usage**

Domains (y, domains, value)

**Arguments**

y	A numeric main vector of size n
domains	A numeric/character/logic auxiliary vector of size n
value	A value of the auxiliary vector

**Value**

A numeric vector, copy of  $y$ , with some values turned zero depending on values of domains and value

**Examples**

```
##### Example 1 #####
U <- c(13, 18, 20, 14, 9)
#Let build an auxiliary vector indicating whether values in U are above or below the mean.
aux <- c("Below", "Above", "Above", "Below", "Below")
#Now, only values below the mean remain, the other ones are turned zero.
Domains (U, aux, "Below")

##### Example 2 #####
data(DataA)
attach(DataA)
#Let calculate total feeding expenses corresponding to households in domain a.
sum (Domains (Feed, Domain, "a"))
```

---

 FB

---

*Fuller-Burmeister estimator*


---

**Description**

Produces estimates for population totals and means using the Fuller - Burmeister estimator from survey data obtained from a dual frame sampling design. Confidence intervals are also computed, if required.

**Usage**

```
FB(ysA, ysB, pi_A, pi_B, domains_A, domains_B, conf_level = NULL)
```

**Arguments**

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable of interest from $s_B$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".

domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals.

### Details

Fuller-Burmeister estimator of population total is given by

$$\hat{Y}_{FB} = \hat{Y}_a^A + \hat{\beta}_1 \hat{Y}_{ab}^A + (1 - \hat{\beta}_1) \hat{Y}_{ab}^B + \hat{Y}_b^B + \hat{\beta}_2 (\hat{N}_{ab}^A - \hat{N}_{ab}^B)$$

where optimal values for  $\hat{\beta}$  to minimize variance of the estimator are:

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = - \begin{pmatrix} \hat{V}(\hat{Y}_{ab}^A - \hat{Y}_{ab}^B) & \widehat{Cov}(\hat{Y}_{ab}^A - \hat{Y}_{ab}^B, \hat{N}_{ab}^A - \hat{N}_{ab}^B) \\ \widehat{Cov}(\hat{Y}_{ab}^A - \hat{Y}_{ab}^B, \hat{N}_{ab}^A - \hat{N}_{ab}^B) & \hat{V}(\hat{N}_{ab}^A - \hat{N}_{ab}^B) \end{pmatrix}^{-1} \times \\ \begin{pmatrix} \widehat{Cov}(\hat{Y}_a^A + \hat{Y}_b^B + \hat{Y}_{ab}^B, \hat{Y}_{ab}^A - \hat{Y}_{ab}^B) \\ \widehat{Cov}(\hat{Y}_a^A + \hat{Y}_b^B + \hat{Y}_{ab}^B, \hat{N}_{ab}^A - \hat{N}_{ab}^B) \end{pmatrix}$$

Due to Fuller-Burmeister estimator is not defined for estimating population sizes, estimation of the mean is computed as  $\hat{Y}_{FB}/\hat{N}_H$ , where  $\hat{N}_H$  is the estimation of the population size using Hartley estimator. Estimated variance for the Fuller-Burmeister estimator can be obtained through expression

$$\begin{aligned} \hat{V}(\hat{Y}_{FB}) &= \hat{V}(\hat{Y}_a^A) + \hat{V}(\hat{Y}_b^B) + \hat{\beta}_1 [\widehat{Cov}(\hat{Y}_a^A, \hat{Y}_{ab}^A) - \widehat{Cov}(\hat{Y}_b^B, \hat{Y}_{ab}^B)] \\ &\quad + \hat{\beta}_2 [\widehat{Cov}(\hat{Y}_a^A, \hat{N}_{ab}^A) - \widehat{Cov}(\hat{Y}_b^B, \hat{N}_{ab}^B)] \end{aligned}$$

If both first and second order probabilities are known, variances and covariances involved in calculation of  $\hat{\beta}$  and  $\hat{V}(\hat{Y}_{FB})$  are estimated using functions VarHT and CovHT, respectively. If only first order probabilities are known, variances are estimated using Deville's method and covariances are estimated using following expression

$$\widehat{Cov}(\hat{X}, \hat{Y}) = \frac{\hat{V}(X + Y) - \hat{V}(X) - \hat{V}(Y)}{2}$$

### Value

FB returns an object of class "EstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	total and mean estimation for main variable(s).
VarEst	variance estimation for main variable(s).

If parameter conf\_level is different from NULL, object includes component

ConfInt	total and mean estimation and confidence intervals for main variables(s).
---------	---

In addition, components TotDomEst and MeanDomEst are available when estimator is based on estimators of the domains. Component Param shows value of parameters involved in calculation of the estimator (if any). By default, only Est component (or ConfInt component, if parameter conf\_level is different from NULL) is shown. It is possible to access to all the components of the objects by using function summary.

**References**

Fuller, W.A. and Burmeister, L.F. (1972). *Estimation for Samples Selected From Two Overlapping Frames* ASA Proceedings of the Social Statistics Sections, 245 - 249.

**See Also**

[Hartley JackFB](#)

**Examples**

```
data(Data)
data(DatB)
data(PiklA)
data(PiklB)

#Let calculate Fuller-Burmeister estimator for variable Clothing
FB(DatA$Clo, DatB$Clo, PiklA, PiklB, DatA$Domain, DatB$Domain)

#Now, let calculate Fuller-Burmeister estimator and a 90% confidence interval
#for variable Leisure, considering only first order inclusion probabilities
FB(DatA$Lei, DatB$Lei, DatA$ProbA, DatB$ProbB, DatA$Domain,
DatB$Domain, 0.90)
```

---

Hartley

*Hartley estimator*

---

**Description**

Produces estimates for population totals and means using Hartley estimator from survey data obtained from a dual frame sampling design. Confidence intervals are also computed, if required.

**Usage**

```
Hartley(ysA, ysB, pi_A, pi_B, domains_A, domains_B, conf_level = NULL)
```

**Arguments**

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable of interest from $s_B$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .

domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals.

### Details

Hartley estimator of population total is given by

$$\hat{Y}_H = \hat{Y}_a^A + \hat{\theta}\hat{Y}_{ab}^A + (1 - \hat{\theta})\hat{Y}_{ab}^B + \hat{Y}_b^B$$

where  $\hat{\theta} \in [0, 1]$ . Optimum value for  $\hat{\theta}$  to minimize variance of the estimator is

$$\hat{\theta}_{opt} = \frac{\hat{V}(\hat{Y}_{ab}^B) + \widehat{Cov}(\hat{Y}_b^B, \hat{Y}_{ab}^B) - \widehat{Cov}(\hat{Y}_a^A, \hat{Y}_{ab}^A)}{\hat{V}(\hat{Y}_{ab}^A) + \hat{V}(\hat{Y}_{ab}^B)}$$

Taking into account the independence between  $s_A$  and  $s_B$ , an estimator for the variance of the Hartley estimator can be obtained as follows:

$$\hat{V}(\hat{Y}_H) = \hat{V}(\hat{Y}_a^A + \hat{\theta}\hat{Y}_{ab}^A) + \hat{V}((1 - \hat{\theta})\hat{Y}_{ab}^B + \hat{Y}_b^B)$$

If both first and second order probabilities are known, variances and covariances involved in calculation of  $\hat{\theta}_{opt}$  and  $\hat{V}(\hat{Y}_H)$  are estimated using functions VarHT and CovHT, respectively. If only first order probabilities are known, variances are estimated using Deville's method and covariances are estimated using following expression

$$\widehat{Cov}(\hat{X}, \hat{Y}) = \frac{\hat{V}(X + Y) - \hat{V}(X) - \hat{V}(Y)}{2}$$

### Value

Hartley returns an object of class "EstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	total and mean estimation for main variable(s).
VarEst	variance estimation for main variable(s).

If parameter conf\_level is different from NULL, object includes component

ConfInt	total and mean estimation and confidence intervals for main variables(s).
---------	---

In addition, components TotDomEst and MeanDomEst are available when estimator is based on estimators of the domains. Component Param shows value of parameters involved in calculation of the estimator (if any). By default, only Est component (or ConfInt component, if parameter conf\_level is different from NULL) is shown. It is possible to access to all the components of the objects by using function summary.

## References

Hartley, H. O. (1962) *Multiple Frames Surveys*. Proceedings of the American Statistical Association, Social Statistics Sections, 203 - 206.

Hartley, H. O. (1974) *Multiple frame methodology and selected applications*. Sankhya C, Vol. 36, 99 - 118.

## See Also

[JackHartley](#)

## Examples

```
data(DatA)
data(DatB)
data(Pik1A)
data(Pik1B)

#Let calculate Hartley estimator for variable Feeding
Hartley(DatA$Feed, DatB$Feed, Pik1A, Pik1B, DatA$Domain, DatB$Domain)

#Now, let calculate Hartley estimator and a 90% confidence interval
#for variable Leisure, considering only first order inclusion probabilities
Hartley(DatA$Lei, DatB$Lei, DatA$ProbA, DatB$ProbB, DatA$Domain,
DatB$Domain, 0.90)
```

---

HT

*Horvitz - Thompson estimator*

---

## Description

Computes the Horvitz - Thompson estimator

## Usage

```
HT(y, pik)
```

## Arguments

y	A numeric vector of size n containing information about variable of interest
pik	A numeric vector of size n containing first order inclusion probabilities for units included in y

## Details

Horvitz - Thompson estimator of population total is given by

$$\hat{Y}_{HT} = \sum_{k \in s} \frac{y_k}{\pi_k}$$

**Value**

A numeric value representing Horvitz - Thompson estimator for population total for considered values

**References**

Horvitz, D. G. and Thompson, D. J. (1952) *A generalization of sampling without replacement from a finite universe*. Journal of the American Statistical Association, 47, 663 - 685

**See Also**

[VarHT](#)

**Examples**

```
##### Example 1 #####
U <- c(13, 18, 20, 14, 9)
#A simple random sample of size 2 without replacement is drawn from population
s <- sample(U, 2)
ps <- c(0.4, 0.4)
HT(s, ps)

##### Example 2 #####
data(DataA)
attach(DataA)
#Let estimate population total for variable Feeding in frame A
HT(Feed, ProbA)
```

---

JackBKA

*Confidence intervals for Bankier-Kalton-Anderson estimator based on jackknife method*

---

**Description**

Calculates confidence intervals for Bankier-Kalton-Anderson estimator using jackknife procedure

**Usage**

```
JackBKA(ysA, ysB, piA, piB, pik_ab_B, pik_ba_A, domainsA, domainsB,
conf_level, sdA = "srs", sdB = "srs", strA = NULL, strB = NULL, clusA = NULL,
clusB = NULL, fcpA = FALSE, fcpB = FALSE)
```

**Arguments**

ysA            A numeric vector of length  $n_A$  or a numeric matrix or data frame of dimensions  $n_A \times c$  containing information about variable of interest from  $s_A$ .

ysB            A numeric vector of length  $n_B$  or a numeric matrix or data frame of dimensions  $n_B \times c$  containing information about variable of interest from  $s_B$ .

piA	A numeric vector of length $nA$ or a square numeric matrix of dimension $nA$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
piB	A numeric vector of length $nB$ or a square numeric matrix of dimension $nB$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $nA$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $nB$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domainsA	A character vector of size $nA$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domainsB	A character vector of size $nB$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

## Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected. In this context, jackknife variance estimator of an estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th stratum from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA} / n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB} / n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

## Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

## References

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

## See Also

[BKA](#)

## Examples

```
data(DatA)
data(DatB)

#Let obtain a 95% jackknife confidence interval for variable Clothing,
#supposing a stratified sampling in frame A and a simple random sampling without
#replacement in frame B with no finite population correction factor in any frame.
JackBKA(DatA$Feed, DatB$Feed, DatA$ProbA, DatB$ProbB, DatA$ProbB,
DatB$ProbA, DatA$Domain, DatB$Domain, 0.95, "str", "srs",
strA = DatA$Stratum)

#Let check how interval estimation varies when a finite
#population correction factor is considered in both frames.
JackBKA(DatA$Feed, DatB$Feed, DatA$ProbA, DatB$ProbB, DatA$ProbB,
DatB$ProbA, DatA$Domain, DatB$Domain, 0.95, "str", "srs",
strA = DatA$Stratum, fcpA = TRUE, fcpB = TRUE)
```

---

JackCalDF	<i>Confidence intervals for dual frame calibration estimator based on jackknife method</i>
-----------	--

---

### Description

Calculates confidence intervals for dual frame calibration estimator using jackknife procedure

### Usage

```
JackCalDF(ysA, ysB, piA, piB, domainsA, domainsB, N_A = NULL, N_B = NULL,
N_ab = NULL, xsAframeA = NULL, xsBframeA = NULL, xsAframeB = NULL,
xsBframeB = NULL, xsT = NULL, XA = NULL, XB = NULL, X = NULL, met = "linear",
conf_level, sdA = "srs", sdB = "srs", stra = NULL, strB = NULL, clusA = NULL,
clusB = NULL, fcpA = FALSE, fcpB = FALSE)
```

### Arguments

ysA	A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times c$ containing information about variable of interest from $s_B$ .
piA	A numeric vector of length $nA$ or a square numeric matrix of dimension $nA$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
piB	A numeric vector of length $nB$ or a square numeric matrix of dimension $nB$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domainsA	A character vector of size $nA$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domainsB	A character vector of size $nB$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	(Optional) A numeric value indicating the size of frame A
N_B	(Optional) A numeric value indicating the size of frame B
N_ab	(Optional) A numeric value indicating the size of the overlap domain
xsAframeA	(Optional) A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsBframeA	(Optional) A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.

xsAFrameB	(Optional) A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.
xsBFrameB	(Optional) A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
xsT	(Optional) A numeric vector of length $n$ or a numeric matrix or data frame of dimensions $n \times m_T$ , with $m_T$ the number of auxiliary variables in both frames, containing auxiliary information for all units in the entire sample $s = s_A \cup s_B$ .
XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
X	(Optional) A numeric value or vector of length $m_T$ , with $m_T$ the number of auxiliary variables in both frames, indicating the population totals for the auxiliary variables considered in both frames.
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.

fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of a estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA} / n_A$  and  $\bar{\pi}_B = \sum_{j \in s_A} \pi_{jB} / n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

### See Also

[CalDF](#)

### Examples

```
data(DatA)
data(DatB)

#Let obtain a 95% jackknife confidence interval for variable Clothing,
#with frame sizes and overlap domain size known, supposing a stratified
#sampling in frame A and a simple random sampling without replacement
#in frame B with no finite population correction factor in any frame.
JackCalDF(DatA$Clo, DatB$Clo, DatA$ProbA, DatB$ProbB,
DatA$Domain, DatB$Domain, N_A = 1735, N_B = 1191, N_ab = 601, conf_level = 0.95,
sdA = "str", sdB = "srs", strA = DatA$Stratum)

#Finally, let consider a finite population correction factor in both frames.
JackCalDF(DatA$Clo, DatB$Clo, DatA$ProbA, DatB$ProbB,
```

```

DatA$Domain, DatB$Domain, N_A = 1735, N_B = 1191, N_ab = 601, conf_level = 0.95,
sdA = "str", sdB = "srs", strA = DatA$Stratum, fcpA = TRUE, fcpB = TRUE)

```

---

JackCalSF *Confidence intervals for SF calibration estimator based on jackknife method*

---

## Description

Produces estimates for variance of SF calibration estimator using Jackknife procedure

## Usage

```

JackCalSF(ysA, ysB, piA, piB, pik_ab_B, pik_ba_A, domainsA, domainsB,
N_A = NULL, N_B = NULL, N_ab = NULL, xsAframeA = NULL, xsBframeA = NULL,
xsAframeB = NULL, xsBframeB = NULL, xsT = NULL, XA = NULL, XB = NULL,
X = NULL, met = "linear", conf_level, sdA = "srs", sdB = "srs", strA = NULL,
strB = NULL, clusA = NULL, clusB = NULL, fcpA = FALSE, fcpB = FALSE)

```

## Arguments

ysA	A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times c$ containing information about variable of interest from $s_B$ .
piA	A numeric vector of length $nA$ or a square numeric matrix of dimension $nA$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
piB	A numeric vector of length $nB$ or a square numeric matrix of dimension $nB$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $nA$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $nB$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domainsA	A character vector of size $nA$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domainsB	A character vector of size $nB$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	(Optional) A numeric value indicating the size of frame A
N_B	(Optional) A numeric value indicating the size of frame B
N_ab	(Optional) A numeric value indicating the size of the overlap domain

xsAFrameA	(Optional) A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsBFrameA	(Optional) A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.
xsAFrameB	(Optional) A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.
xsBFrameB	(Optional) A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
xsT	(Optional) A numeric vector of length $n$ or a numeric matrix or data frame of dimensions $n \times m_T$ , with $m_T$ the number of auxiliary variables in both frames, containing auxiliary information for all units in the entire sample $s = s_A \cup s_B$ .
XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
X	(Optional) A numeric value or vector of length $m_T$ , with $m_T$ the number of auxiliary variables in both frames, indicating the population totals for the auxiliary variables considered in both frames.
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.

c1usA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
c1usB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of a estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ysA$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th from sample  $ysB$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA} / n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB} / n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

### See Also

[CalSF](#)

### Examples

```
data(DatA)
data(DatB)
```

```
#Let obtain a 95% jackknife confidence interval for variable Clothing,
#with frame sizes and overlap domain size known, supposing a stratified
```

```
#sampling in frame A and a simple random sampling without replacement
#in frame B with no finite population correction factor in any frame
JackCalsF(DatA$Clo, DatB$Clo, DatA$ProbA, DatB$ProbB,
DatA$ProbB, DatB$ProbA, DatA$Domain, DatB$Domain, N_A = 1735,
N_B = 1191, N_ab = 601, conf_level = 0.95, sdA = "str", sdB = "srs",
strA = DatA$Stratum)
```

JackFB

*Confidence intervals for Fuller-Burmeister estimator based on jackknife method*

## Description

Calculates confidence intervals for Fuller-Burmeister estimator using jackknife procedure

## Usage

```
JackFB(ysA, ysB, piA, piB, domainsA, domains_B, conf_level, sdA = "srs",
sdB = "srs", strA = NULL, strB = NULL, clusA = NULL, clusB = NULL, fcpA = FALSE,
fcpB = FALSE)
```

## Arguments

ysA	A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times c$ containing information about variable of interest from $s_B$ .
piA	A numeric vector of length $nA$ or a square numeric matrix of dimension $nA$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
piB	A numeric vector of length $nB$ or a square numeric matrix of dimension $nB$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domainsA	A character vector of size $nA$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $nB$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "streclu" (stratified cluster sampling). Default is "srs".

sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected. In this context, jackknife variance estimator of an estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA}/n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB}/n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

**See Also**[FB](#)**Examples**

```

data(DatA)
data(DatB)

#Let obtain a 95% jackknife confidence interval for variable Clothing,
#supposing a stratified sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction factor
#in any frame.
JackFB(DatA$Clo, DatB$Clo, DatA$ProbA, DatB$ProbB, DatA$Domain,
DatB$Domain, 0.95, "str", "srs", strA = DatA$Stratum)

#Let check how interval estimation varies when a finite
#population correction factor is considered in both frames.
JackFB(DatA$Clo, DatB$Clo, DatA$ProbA, DatB$ProbB, DatA$Domain,
DatB$Domain, 0.95, "str", "srs", strA = DatA$Stratum,
fcpA = TRUE, fcpB = TRUE)

```

JackHartley

*Confidence intervals for Hartley estimator based on jackknife method***Description**

Calculates confidence intervals for Hartley estimator using jackknife procedure

**Usage**

```

JackHartley(ysA, ysB, piA, piB, domainsA, domainsB, conf_level, sdA = "srs",
sdB = "srs", strA = NULL, strB = NULL, clusA = NULL, clusB = NULL, fcpA = FALSE,
fcpB = FALSE)

```

**Arguments**

ysA	A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times c$ containing information about variable of interest from $s_B$ .
piA	A numeric vector of length $nA$ or a square numeric matrix of dimension $nA$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
piB	A numeric vector of length $nB$ or a square numeric matrix of dimension $nB$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .

domainsA	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domainsB	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

## Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of a estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping j-th unit of l-th from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \pi_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or

$\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA}/n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB}/n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

### See Also

[Hartley](#)

### Examples

```
data(DatA)
data(DatB)

#Let obtain a 95% jackknife confidence interval for variable Feeding,
#supposing a stratified sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackHartley(DatA$Feed, DatB$Feed, DatA$ProbA, DatB$ProbB, DatA$Domain,
DatB$Domain, 0.95, "str", "srs", strA = DatA$Stratum)

#Let check how interval estimation varies when a finite
#population correction factor is considered in both frames.
JackHartley(DatA$Feed, DatB$Feed, DatA$ProbA, DatB$ProbB, DatA$Domain,
DatB$Domain, 0.95, "str", "srs", strA = DatA$Stratum, fcpA = TRUE,
fcpB = TRUE)
```

---

JackMLCDF

*Confidence intervals for MLCDF estimator based on jackknife method*

---

### Description

Calculates confidence intervals for MLCDF estimator using jackknife procedure

### Usage

```
JackMLCDF(ysA, ysB, pik_A, pik_B, domains_A, domains_B, xsA, xsB, xA, xB, ind_samA,
ind_samB, ind_domA, ind_domB, N, N_ab = NULL, met = "linear", conf_level, sdA = "srs",
sdB = "srs", strA = NULL, strB = NULL, clusA = NULL, clusB = NULL, fcpA = FALSE,
fcpB = FALSE)
```

**Arguments**

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xSA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
xA	A numeric vector or length $N_A$ or a numeric matrix or data frame of dimensions $N_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information for the units in frame A.
xB	A numeric vector or length $N_B$ or a numeric matrix or data frame of dimensions $N_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information for the units in frame B.
ind_samA	A numeric vector of length $n_A$ containing the identifiers of units of the frame A (from 1 to $N_A$ ) that belongs to $s_A$ .
ind_samB	A numeric vector of length $n_B$ containing the identifiers of units of the frame B (from 1 to $N_B$ ) that belongs to $s_B$ .
ind_domA	A character vector of length $N_A$ indicating the domain each unit from frame A belongs to. Possible values are "a" and "ab".
ind_domB	A character vector of length $N_B$ indicating the domain each unit from frame B belongs to. Possible values are "b" and "ba".
N	A numeric value indicating the size of the population.
N_ab	(Optional) A numeric value indicating the size of the overlap domain
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "streclu" (stratified cluster sampling). Default is "srs".

sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into  $L$  strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the  $l$ -th stratum is selected. In this context, jackknife variance estimator of a estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $y_{sA}$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th from sample  $y_{sB}$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA}/n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB}/n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

## References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

## See Also

[MLCDF](#)

## Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)

N <- nrow(DatPopM)
levels(DatPopM$Domain) <- c(levels(DatPopM$Domain), "ba")
DatPopMA <- subset(DatPopM, DatPopM$Domain == "a" | DatPopM$Domain == "ab", stringAsFactors = FALSE)
DatPopMB <- subset(DatPopM, DatPopM$Domain == "b" | DatPopM$Domain == "ab", stringAsFactors = FALSE)
DatPopMB[DatPopMB$Domain == "ab",]$Domain <- "ba"

#Let obtain a 95% jackknife confidence interval for variable Feeding,
#supposing a pps sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackMLCDF(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain,
DatMB$Domain, DatMA$Read, DatMB$Read, DatPopMA$Read, DatPopMB$Read,
DatMA$Id_Frame, DatMB$Id_Frame, DatPopMA$Domain, DatPopMB$Domain, N,
conf_level = 0.95, sdA = "pps", sdB = "srs")
```

---

JackMLCDW

*Confidence intervals for MLCDW estimator based on jackknife method*

---

## Description

Calculates confidence intervals for MLCDW estimator using jackknife procedure

## Usage

```
JackMLCDW (ysA, ysB, pik_A, pik_B, domains_A, domains_B, xsA, xsB, x,
ind_sam, N_A, N_B, N_ab = NULL, met = "linear", conf_level, sdA = "srs",
sdB = "srs", strA = NULL, strB = NULL, clusA = NULL, clusB = NULL,
fcpA = FALSE, fcpB = FALSE)
```

**Arguments**

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xSA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
x	A numeric vector or length $N$ or a numeric matrix or data frame of dimensions $N \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information for every unit in the population.
ind_sam	A numeric vector of length $n = n_A + n_B$ containing the identifiers of units of the population (from 1 to $N$ ) that belongs to $s_A$ or $s_B$
N_A	A numeric value indicating the size of frame A
N_B	A numeric value indicating the size of frame B
N_ab	(Optional) A numeric value indicating the size of the overlap domain
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.

strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of an estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ysA$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th from sample  $ysB$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA}/n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB}/n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

### See Also

[MLCDW](#)

**Examples**

```

data(DatMA)
data(DatMB)
data(DatPopM)

IndSample <- c(DatMA$Id_Pop, DatMB$Id_Pop)
N_FrameA <- nrow(DatPopM[DatPopM$Domain == "a" | DatPopM$Domain == "ab",])
N_FrameB <- nrow(DatPopM[DatPopM$Domain == "b" | DatPopM$Domain == "ab",])

#Let obtain a 95% jackknife confidence interval for variable Feeding,
#supposing a pps sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackMLCDW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain,
DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, N_FrameA,
N_FrameB, conf_level = 0.95, sdA = "pps", sdB = "srs")

```

JackMLCSW

*Confidence intervals for MLCSW estimator based on jackknife method***Description**

Calculates confidence intervals for MLCSW estimator using jackknife procedure

**Usage**

```

JackMLCSW (ysA, ysB, pik_A, pik_B, pik_ab_B, pik_ba_A, domains_A,
domains_B, xsA, xsB, x, ind_sam, N_A, N_B, N_ab = NULL, met = "linear",
conf_level, sdA = "srs", sdB = "srs", strA = NULL, strB = NULL, clusA = NULL,
clusB = NULL, fcpA = FALSE, fcpB = FALSE)

```

**Arguments**

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .

pik_ba_A	A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xSA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
x	A numeric vector or length $N$ or a numeric matrix or data frame of dimensions $N \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information for every unit in the population.
ind_sam	A numeric vector of length $n = n_A + n_B$ containing the identifiers of units of the population (from 1 to $N$ ) that belongs to $s_A$ or $s_B$
N_A	A numeric value indicating the size of frame A
N_B	A numeric value indicating the size of frame B
N_ab	(Optional) A numeric value indicating the size of the overlap domain
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.

c1usB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of an estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping j-th unit of l-th from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA} / n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB} / n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

### See Also

[MLCSW](#)

### Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)
```

```
IndSample <- c(DatMA$Id_Pop, DatMB$Id_Pop)
N_FrameA <- nrow(DatPopM[DatPopM$Domain == "a" | DatPopM$Domain == "ab",])
```

```

N_FrameB <- nrow(DatPopM[DatPopM$Domain == "b" | DatPopM$Domain == "ab",])

#Let obtain a 95% jackknife confidence interval for variable Feeding,
#supposing a pps sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackMLCSW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$ProbB,
DatMB$ProbA, DatMA$Domain, DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read,
IndSample, N_FrameA, N_FrameB, conf_level = 0.95, sdA = "pps", sdB = "srs")

```

---

JackMLDF

*Confidence intervals for MLDF estimator based on jackknife method*


---

### Description

Calculates confidence intervals for MLDF estimator using jackknife procedure

### Usage

```

JackMLDF (ysA, ysB, pik_A, pik_B, domains_A, domains_B, xsA, xsB, xA, xB, ind_samA,
ind_samB, ind_domA, ind_domB, N, conf_level, sdA = "srs", sdB = "srs", strA = NULL,
strB = NULL, clusA = NULL, clusB = NULL, fcpA = FALSE, fcpB = FALSE)

```

### Arguments

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xsA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .

<code>xA</code>	A numeric vector or length $N_A$ or a numeric matrix or data frame of dimensions $N_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information for the units in frame A.
<code>xB</code>	A numeric vector or length $N_B$ or a numeric matrix or data frame of dimensions $N_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information for the units in frame B.
<code>ind_samA</code>	A numeric vector of length $n_A$ containing the identifiers of units of the frame A (from 1 to $N_A$ ) that belongs to $s_A$ .
<code>ind_samB</code>	A numeric vector of length $n_B$ containing the identifiers of units of the frame B (from 1 to $N_B$ ) that belongs to $s_B$ .
<code>ind_domA</code>	A character vector of length $N_A$ indicating the domain each unit from frame A belongs to. Possible values are "a" and "ab".
<code>ind_domB</code>	A character vector of length $N_B$ indicating the domain each unit from frame B belongs to. Possible values are "b" and "ba".
<code>N</code>	A numeric value indicating the size of the population.
<code>conf_level</code>	A numeric value indicating the confidence level for the confidence intervals.
<code>sdA</code>	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
<code>sdB</code>	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
<code>strA</code>	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
<code>strB</code>	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
<code>clusA</code>	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
<code>clusB</code>	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
<code>fcpA</code>	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
<code>fcpB</code>	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

## Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of a estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA} / n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB} / n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

## Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

## References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

## See Also

[MLDF](#)

## Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)

N <- nrow(DatPopM)
levels(DatPopM$Domain) <- c(levels(DatPopM$Domain), "ba")
DatPopMA <- subset(DatPopM, DatPopM$Domain == "a" | DatPopM$Domain == "ab", stringAsFactors = FALSE)
DatPopMB <- subset(DatPopM, DatPopM$Domain == "b" | DatPopM$Domain == "ab", stringAsFactors = FALSE)
DatPopMB[DatPopMB$Domain == "ab",]$Domain <- "ba"

#Let obtain a 95% jackknife confidence interval for variable Feeding,
#supposing a pps sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackMLDF(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain,
```

```
DatMB$Domain, DatMA$Read, DatMB$Read, DatPopMA$Read, DatPopMB$Read,
DatMA$Id_Frame, DatMB$Id_Frame, DatPopMA$Domain, DatPopMB$Domain, N, 0.95,
"pps", "srs")
```

---

 JackMLDW

*Confidence intervals for MLDW estimator based on jackknife method*


---

## Description

Calculates confidence intervals for MLDW estimator using jackknife procedure

## Usage

```
JackMLDW (ysA, ysB, pik_A, pik_B, domains_A, domains_B, xsA, xsB, x,
ind_sam, conf_level, sdA = "srs", sdB = "srs", strA = NULL, strB = NULL,
clusA = NULL, clusB = NULL, fcpA = FALSE, fcpB = FALSE)
```

## Arguments

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xsA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
x	A numeric vector or length $N$ or a numeric matrix or data frame of dimensions $N \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information for every unit in the population.
ind_sam	A numeric vector of length $n = n_A + n_B$ containing the identifiers of units of the population (from 1 to $N$ ) that belongs to $s_A$ or $s_B$
conf_level	A numeric value indicating the confidence level for the confidence intervals.

sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of a estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from ysA and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping j-th unit of l-th from sample ysB and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA}/n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB}/n_B$  A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

## References

- Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.
- Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

## See Also

[MLDW](#)

## Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)

IndSample <- c(DatMA$Id_Pop, DatMB$Id_Pop)

#Let obtain a 95% jackknife confidence interval for variable Feeding,
#supposing a pps sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackMLDW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain,
DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, 0.95,
"pps", "srs")
```

---

JackMLSW

*Confidence intervals for MLSW estimator based on jackknife method*

---

## Description

Calculates confidence intervals for MLSW estimator using jackknife procedure

## Usage

```
JackMLSW (ysA, ysB, pik_A, pik_B, pik_ab_B, pik_ba_A, domains_A,
domains_B, xsA, xsB, x, ind_sam, conf_level, sdA = "srs", sdB = "srs",
strA = NULL, strB = NULL, clusA = NULL, clusB = NULL, fcpA = FALSE, fcpB = FALSE)
```

## Arguments

- |       |   |
|-------|---|
| ysA   | A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ . |
| ysB   | A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ . |
| pik_A | A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .       |

pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xSA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
x	A numeric vector or length $N$ or a numeric matrix or data frame of dimensions $N \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information for every unit in the population.
ind_sam	A numeric vector of length $n = n_A + n_B$ containing the identifiers of units of the population (from 1 to $N$ ) that belongs to $s_A$ or $s_B$
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.

c1usB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of a estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ysA$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping j-th unit of l-th from sample  $ysB$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA} / n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB} / n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

### See Also

[MLSW](#)

### Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)
```

```
IndSample <- c(DatMA$Id_Pop, DatMB$Id_Pop)
```

```
#Let obtain a 95% jackknife confidence interval for variable Feeding,
#supposing a pps sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackMLSW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$ProbB,
DatMB$ProbA, DatMA$Domain, DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read,
IndSample, 0.95, "pps", "srs")
```

JackPEL

*Confidence intervals for the pseudo empirical likelihood estimator  
based on jackknife method*

## Description

Calculates confidence intervals for pseudo empirical likelihood estimator using jackknife procedure

## Usage

```
JackPEL(ysA, ysB, piA, piB, domainsA, domainsB, N_A = NULL, N_B = NULL,
N_ab = NULL, xsAframeA = NULL, xsBframeA = NULL, xsAframeB = NULL, xsBframeB = NULL,
XA = NULL, XB = NULL, conf_level, sdA = "srs", sdB = "srs", strA = NULL, strB = NULL,
clusA = NULL, clusB = NULL, fcpA = FALSE, fcpB = FALSE)
```

## Arguments

ysA	A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times c$ containing information about variable of interest from $s_B$ .
piA	A numeric vector of length $nA$ or a square numeric matrix of dimension $nA$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
piB	A numeric vector of length $nB$ or a square numeric matrix of dimension $nB$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domainsA	A character vector of size $nA$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domainsB	A character vector of size $nB$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	(Optional) A numeric value indicating the size of frame A
N_B	(Optional) A numeric value indicating the size of frame B
N_ab	(Optional) A numeric value indicating the size of the overlap domain

xsAFrameA	(Optional) A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsBFrameA	(Optional) A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.
xsAFrameB	(Optional) A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.
xsBFrameB	(Optional) A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcPA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.

fcpB (Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected In this context, jackknife variance estimator of a estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping j-th unit of l-th from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA} / n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB} / n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

### See Also

[PEL](#)

---

JackPML	<i>Confidence intervals for the pseudo maximum likelihood estimator based on jackknife method</i>
---------	---

---

### Description

Calculates confidence intervals for pseudo maximum likelihood estimator using jackknife procedure

### Usage

```
JackPML(ysA, ysB, piA, piB, domainsA, domainsB, N_A, N_B, conf_level,
sdA = "srs", sdB = "srs", strA = NULL, strB = NULL, clusA = NULL, clusB = NULL,
fcpA = FALSE, fcpB = FALSE)
```

**Arguments**

ysA	A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times c$ containing information about variable of interest from $s_B$ .
piA	A numeric vector of length $nA$ or a square numeric matrix of dimension $nA$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
piB	A numeric vector of length $nB$ or a square numeric matrix of dimension $nB$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domainsA	A character vector of size $nA$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domainsB	A character vector of size $nB$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	A numeric value indicating the size of frame A
N_B	A numeric value indicating the size of frame B
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

## Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected. In this context, jackknife variance estimator of an estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA} / n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB} / n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

## Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

## References

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

## See Also

[PML](#)

## Examples

```
data(DatA)
data(DatB)

#Let obtain a 95% jackknife confidence interval for variable Leisure,
#supposing a stratified sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackPML(DatA$Lei, DatB$Lei, DatA$ProbA, DatB$ProbB, DatA$Domain,
DatB$Domain, 1735, 1191, 0.95, "str", "srs", strA = DatA$Stratum)

#Let check how interval estimation varies when a finite
#population correction factor is considered in both frames.
JackPML(DatA$Lei, DatB$Lei, DatA$ProbA, DatB$ProbB, DatA$Domain,
DatB$Domain, 1735, 1191, 0.95, "str", "srs", strA = DatA$Stratum,
fcPA = TRUE, fcPB = TRUE)
```

---

JackSFRR                      *Confidence intervals for raking ratio estimator based on jackknife method*

---

### Description

Calculates confidence intervals for raking ratio estimator using jackknife procedure

### Usage

```
JackSFRR(ysA, ysB, piA, piB, pik_ab_B, pik_ba_A, domainsA, domainsB, N_A,
N_B, conf_level, sdA = "srs", sdB = "srs", strA = NULL, strB = NULL, clusA = NULL,
clusB = NULL, fcpA = FALSE, fcpB = FALSE)
```

### Arguments

ysA	A numeric vector of length $nA$ or a numeric matrix or data frame of dimensions $nA \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $nB$ or a numeric matrix or data frame of dimensions $nB \times c$ containing information about variable of interest from $s_B$ .
piA	A numeric vector of length $nA$ or a square numeric matrix of dimension $nA$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
piB	A numeric vector of length $nB$ or a square numeric matrix of dimension $nB$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $nA$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $nB$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domainsA	A character vector of size $nA$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domainsB	A character vector of size $nB$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	A numeric value indicating the size of frame A
N_B	A numeric value indicating the size of frame B
conf_level	A numeric value indicating the confidence level for the confidence intervals.
sdA	(Optional) A character vector indicating the sampling design considered in frame A. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "streclu" (stratified cluster sampling). Default is "srs".

sdB	(Optional) A character vector indicating the sampling design considered in frame B. Possible values are "srs" (simple random sampling without replacement), "pps" (probabilities proportional to size sampling), "str" (stratified sampling), "clu" (cluster sampling) and "strclu" (stratified cluster sampling). Default is "srs".
strA	(Optional) A numeric vector indicating the stratum each unit in frame A belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame A.
strB	(Optional) A numeric vector indicating the stratum each unit in frame B belongs to, if a stratified sampling or a stratified cluster sampling has been considered in frame B.
clusA	(Optional) A numeric vector indicating the cluster each unit in frame A belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame A.
clusB	(Optional) A numeric vector indicating the cluster each unit in frame B belongs to, if a cluster sampling or a stratified cluster sampling has been considered in frame B.
fcpA	(Optional) A logic value indicating if a finite population correction factor should be considered in frame A. Default is FALSE.
fcpB	(Optional) A logic value indicating if a finite population correction factor should be considered in frame B. Default is FALSE.

### Details

Let suppose a non stratified sampling design in frame A and a stratified sampling design in frame B where frame has been divided into L strata and a sample of size  $n_{Bl}$  from the  $N_{Bl}$  composing the l-th stratum is selected. In this context, jackknife variance estimator of an estimator  $\hat{Y}_c$  is given by

$$v_J(\hat{Y}_c) = \frac{n_A - 1}{n_A} \sum_{i \in s_A} (\hat{Y}_c^A(i) - \bar{Y}_c^A)^2 + \sum_{l=1}^L \frac{n_{Bl} - 1}{n_{Bl}} \sum_{i \in s_{Bl}} (\hat{Y}_c^B(lj) - \bar{Y}_c^{Bl})^2$$

with  $\hat{Y}_c^A(i)$  the value of estimator  $\hat{Y}_c$  after dropping  $i$ -th unit from  $ys_A$  and  $\bar{Y}_c^A$  the mean of values  $\hat{Y}_c^A(i)$ . Similarly,  $\hat{Y}_c^B(lj)$  is the value taken by  $\hat{Y}_c$  after dropping  $j$ -th unit of  $l$ -th from sample  $ys_B$  and  $\bar{Y}_c^{Bl}$  is the mean of values  $\hat{Y}_c^B(lj)$ . If needed, a finite population correction factor can be included in frames by replacing  $\hat{Y}_c^A(i)$  or  $\hat{Y}_c^B(lj)$  with  $\hat{Y}_c^{A*}(i) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_A}(\hat{Y}_c^A(i) - \hat{Y}_c)$  or  $\hat{Y}_c^{B*}(lj) = \hat{Y}_c + \sqrt{1 - \bar{\pi}_B}(\hat{Y}_c^B(lj) - \hat{Y}_c)$ , where  $\bar{\pi}_A = \sum_{i \in s_A} \pi_{iA}/n_A$  and  $\bar{\pi}_B = \sum_{j \in s_B} \pi_{jB}/n_B$ . A confidence interval for any parameter of interest,  $Y$  can be calculated, then, using the pivotal method.

### Value

A numeric matrix containing estimations of population total and population mean and their corresponding confidence intervals obtained through jackknife method.

### References

Wolter, K. M. (2007) *Introduction to Variance Estimation*. 2nd Edition. Springer, Inc., New York.

**See Also**[SFRR](#)**Examples**

```

data(DataA)
data(DatB)

#Let obtain a 95% jackknife confidence interval for variable Leisure,
#supposing a stratified sampling in frame A and a simple random sampling
#without replacement in frame B with no finite population correction
#factor in any frame.
JackSFRR(DataA$Lei, DatB$Lei, DataA$ProbA, DatB$ProbB, DataA$ProbB,
DatB$ProbA, DataA$Domain, DatB$Domain, 1735, 1191, 0.95, "str", "srs",
strA = DataA$Stratum)

#Let check how interval estimation varies when a finite
#population correction factor is considered in both frames.
JackSFRR(DataA$Lei, DatB$Lei, DataA$ProbA, DatB$ProbB, DataA$ProbB,
DatB$ProbA, DataA$Domain, DatB$Domain, 1735, 1191, 0.95, "str", "srs",
strA = DataA$Stratum, fcpA = TRUE, fcpB = TRUE)

```

MLCDF

*Multinomial logistic calibration estimator under dual frame approach  
with auxiliary information from each frame*

**Description**

Produces estimates for class totals and proportions using multinomial logistic regression from survey data obtained from a dual frame sampling design using a model calibrated dual frame approach with a possibly different set of auxiliary variables for each frame. Confidence intervals are also computed, if required.

**Usage**

```
MLCDF (ysA, ysB, pik_A, pik_B, domains_A, domains_B, xsA, xsB, xA, xB, ind_samA,
ind_samB, ind_domA, ind_domB, N, N_ab = NULL, met = "linear", conf_level = NULL)
```

**Arguments**

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .

domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xSA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
xA	A numeric vector or length $N_A$ or a numeric matrix or data frame of dimensions $N_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information for the units in frame A.
xB	A numeric vector or length $N_B$ or a numeric matrix or data frame of dimensions $N_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information for the units in frame B.
ind_samA	A numeric vector of length $n_A$ containing the identifiers of units of the frame A (from 1 to $N_A$ ) that belongs to $s_A$ .
ind_samB	A numeric vector of length $n_B$ containing the identifiers of units of the frame B (from 1 to $N_B$ ) that belongs to $s_B$ .
ind_domA	A character vector of length $N_A$ indicating the domain each unit from frame A belongs to. Possible values are "a" and "ab".
ind_domB	A character vector of length $N_B$ indicating the domain each unit from frame B belongs to. Possible values are "b" and "ba".
N	A numeric value indicating the size of the population.
N_ab	(Optional) A numeric value indicating the size of the overlap domain
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

## Details

Multinomial logistic calibration estimator in dual frame using auxiliary information from each frame for a proportion is given by

$$\hat{P}_{MLC_i}^{DF} = \frac{1}{N} \left( \sum_{k \in s_A \cup s_B} w_k^\circ z_{ki} \right), \quad i = 1, \dots, m$$

with  $m$  the number of categories of the response variable,  $z_i$  the indicator variable for the  $i$ -th category of the response variable, and  $w^\circ$  calibration weights which are calculated having into account a different set of constraints, depending on the case. For instance, if  $N_A, N_B$  and  $N_{ab}$  are known, calibration constraints are

$$\sum_{k \in s_a} w_k^\circ = N_a, \quad \sum_{k \in s_{ab}} w_k^\circ = \eta N_{ab}, \quad \sum_{k \in s_{ba}} w_k^\circ = (1 - \eta) N_{ab}, \quad \sum_{k \in s_b} w_k^\circ = N_b,$$

$$\sum_{k \in s_A} w_k^{\circ} p_{ki}^A = \sum_{k \in U_a} p_{ki}^A + \eta \sum_{k \in U_{ab}} p_{ki}^A$$

and

$$\sum_{k \in s_B} w_k^{\circ} p_{ki}^B = \sum_{k \in U_b} p_{ki}^B + (1 - \eta) \sum_{k \in U_{ba}} p_{ki}^B$$

with  $\eta \in (0, 1)$  and

$$p_{ki}^A = \frac{\exp(x_k' \beta_i^A)}{\sum_{r=1}^m \exp(x_k' \beta_r^A)},$$

being  $\beta_i^A$  the maximum likelihood parameters of the multinomial logistic model considering original design weights  $d^A$ .  $p_{ki}^B$  can be defined similarly.

### Value

MLCDF returns an object of class "MultEstimatorDF" which is a list with, at least, the following components:

Call                    the matched call.  
Est                     class frequencies and proportions estimations for main variable(s).

### References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

### See Also

[JackMLCDF](#)

### Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)

N <- nrow(DatPopM)
levels(DatPopM$Domain) <- c(levels(DatPopM$Domain), "ba")
DatPopMA <- subset(DatPopM, DatPopM$Domain == "a" | DatPopM$Domain == "ab", stringAsFactors = FALSE)
DatPopMB <- subset(DatPopM, DatPopM$Domain == "b" | DatPopM$Domain == "ab", stringAsFactors = FALSE)
DatPopMB[DatPopMB$Domain == "ab",]$Domain <- "ba"

#Let calculate proportions of categories of variable Prog using MLCDF estimator
#using Read as auxiliary variable
MLCDF(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopMA$Read, DatPopMB$Read, DatMA$Id_Frame, DatMB$Id_Frame,
DatPopMA$Domain, DatPopMB$Domain, N)

#Let obtain 95% confidence intervals together with the estimations
MLCDF(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopMA$Read, DatPopMB$Read, DatMA$Id_Frame, DatMB$Id_Frame,
DatPopMA$Domain, DatPopMB$Domain, N, conf_level = 0.95)
```

---

MLCDW	<i>Multinomial logistic calibration estimator under dual frame approach with auxiliary information from the whole population</i>
-------	--

---

### Description

Produces estimates for class totals and proportions using multinomial logistic regression from survey data obtained from a dual frame sampling design using a model calibrated dual frame approach with auxiliary information from the whole population. Confidence intervals are also computed, if required.

### Usage

```
MLCDW (ysA, ysB, pik_A, pik_B, domains_A, domains_B, xSA, xSB, x, ind_sam, N_A,
      N_B, N_ab = NULL, met = "linear", conf_level = NULL)
```

### Arguments

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xSA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xSB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
x	A numeric vector or length $N$ or a numeric matrix or data frame of dimensions $N \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information for every unit in the population.
ind_sam	A numeric vector of length $n = n_A + n_B$ containing the identifiers of units of the population (from 1 to $N$ ) that belongs to $s_A$ or $s_B$
N_A	A numeric value indicating the size of frame A
N_B	A numeric value indicating the size of frame B

N_ab	(Optional) A numeric value indicating the size of the overlap domain
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

### Details

Multinomial logistic calibration estimator in dual frame using auxiliary information from the whole population for a proportion is given by

$$\hat{P}_{MLC_i}^{DW} = \frac{1}{N} \left( \sum_{k \in s_A \cup s_B} w_k^\circ z_{ki} \right), \quad i = 1, \dots, m$$

with  $m$  the number of categories of the response variable,  $z_i$  the indicator variable for the  $i$ -th category of the response variable, and  $w^\circ$  calibration weights which are calculated having into account a different set of constraints, depending on the case. For instance, if  $N_A, N_B$  and  $N_{ab}$  are known, calibration constraints are

$$\sum_{k \in s_a} w_k^\circ = N_a, \quad \sum_{k \in s_{ab}} w_k^\circ = \eta N_{ab}, \quad \sum_{k \in s_{ba}} w_k^\circ = (1 - \eta) N_{ab}, \quad \sum_{k \in s_b} w_k^\circ = N_b$$

and

$$\sum_{k \in s_A \cup s_B} w_k^\circ p_{ki}^\circ = \sum_{k \in U} p_{ki}^\circ$$

with  $\eta \in (0, 1)$  and

$$p_{ki}^\circ = \frac{\exp(x'_k \beta_i^\circ)}{\sum_{r=1}^m \exp(x'_k \beta_r^\circ)},$$

being  $\beta_i^\circ$  the maximum likelihood parameters of the multinomial logistic model considering weights

$$d_k^\circ = \begin{cases} d_k^A & \text{if } k \in a \\ \eta d_k^A & \text{if } k \in ab \\ (1 - \eta) d_k^B & \text{if } k \in ba \\ d_k^B & \text{if } k \in b \end{cases}.$$

### Value

MLCDW returns an object of class "MultEstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	class frequencies and proportions estimations for main variable(s).

### References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

**See Also**[JackMLCDW](#)**Examples**

```

data(DatMA)
data(DatMB)
data(DatPopM)

IndSample <- c(DatMA$Id_Pop, DatMB$Id_Pop)
N_FrameA <- nrow(DatPopM[DatPopM$Domain == "a" | DatPopM$Domain == "ab",])
N_FrameB <- nrow(DatPopM[DatPopM$Domain == "b" | DatPopM$Domain == "ab",])
N_Domainab <- nrow(DatPopM[DatPopM$Domain == "ab",])
#Let calculate proportions of categories of variable Prog using MLCDW estimator
#using Read as auxiliary variable
MLCDW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, N_FrameA, N_FrameB)

#Now, let suppose that the overlap domain size is known
MLCDW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, N_FrameA, N_FrameB, N_Domainab)

#Let obtain 95% confidence intervals together with the estimations
MLCDW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, N_FrameA, N_FrameB, N_Domainab,
conf_level = 0.95)

```

MLCSW

*Multinomial logistic calibration estimator under single frame approach with auxiliary information from the whole population*

**Description**

Produces estimates for class totals and proportions using multinomial logistic regression from survey data obtained from a dual frame sampling design using a model calibrated single frame approach with auxiliary information from the whole population. Confidence intervals are also computed, if required.

**Usage**

```
MLCSW(ysA, ysB, pik_A, pik_B, pik_ab_B, pik_ba_A, domains_A, domains_B, xsA, xsB,
x, ind_sam, N_A, N_B, N_ab = NULL, met = "linear", conf_level = NULL)
```

**Arguments**

**ysA** A data frame containing information about one or more factors, each one of dimension  $n_A$ , collected from  $s_A$ .

**ysB** A data frame containing information about one or more factors, each one of dimension  $n_B$ , collected from  $s_B$ .

pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xSA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
x	A numeric vector or length $N$ or a numeric matrix or data frame of dimensions $N \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information for every unit in the population.
ind_sam	A numeric vector of length $n = n_A + n_B$ containing the identifiers of units of the population (from 1 to $N$ ) that belongs to $s_A$ or $s_B$
N_A	A numeric value indicating the size of frame A
N_B	A numeric value indicating the size of frame B
N_ab	(Optional) A numeric value indicating the size of the overlap domain
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

### Details

Multinomial logistic calibration estimator in single frame using auxiliary information from the whole population for a proportion is given by

$$\hat{P}_{MLC_i}^{SW} = \frac{1}{N} \left( \sum_{k \in s_A \cup s_B} \tilde{w}_k z_{ki} \right) \quad i = 1, \dots, m$$

with  $m$  the number of categories of the response variable,  $z_i$  the indicator variable for the  $i$ -th category of the response variable, and  $\tilde{w}$  calibration weights which are calculated having into account

a different set of constraints, depending on the case. For instance, if  $N_A, N_B$  and  $N_{ab}$  are known, calibration constraints are

$$\sum_{k \in s_a} \tilde{w}_k = N_a, \quad \sum_{k \in s_{ab} \cup s_{ba}} \tilde{w}_k = N_{ab}, \quad \sum_{k \in s_{ba}} \tilde{w}_k = N_{ba}$$

and

$$\sum_{k \in s_A \cup s_B} \tilde{w}_k \tilde{p}_{ki} = \sum_{k \in U} \tilde{p}_{ki}$$

with

$$\tilde{p}_{ki} = \frac{\exp(x'_k \tilde{\beta}_i)}{\sum_{r=1}^m \exp(x'_k \tilde{\beta}_r)}$$

being  $\tilde{\beta}_i$  the maximum likelihood parameters of the multinomial logistic model considering weights

$$\tilde{d}_k = \begin{cases} d_k^A & \text{if } k \in a \\ (1/d_k^A + 1/d_k^B)^{-1} & \text{if } k \in ab \cup ba \\ d_k^B & \text{if } k \in b \end{cases} .$$

## Value

MLCSW returns an object of class "MultEstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	class frequencies and proportions estimations for main variable(s).

## References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

## See Also

[JackMLCSW](#)

## Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)

IndSample <- c(DatMA$Id_Pop, DatMB$Id_Pop)
N_FrameA <- nrow(DatPopM[DatPopM$Domain == "a" | DatPopM$Domain == "ab",])
N_FrameB <- nrow(DatPopM[DatPopM$Domain == "b" | DatPopM$Domain == "ab",])
N_Domainab <- nrow(DatPopM[DatPopM$Domain == "ab",])
#Let calculate proportions of categories of variable Prog using MLCSW estimator
#using Read as auxiliary variable
MLCSW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$ProbB, DatMB$ProbA,
DatMA$Domain, DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, N_FrameA,
N_FrameB)
```

```
#Now, let suppose that the overlap domian size is known
MLCSW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$ProbB, DatMB$ProbA,
DatMA$Domain, DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, N_FrameA,
N_FrameB, N_Domainab)

#Let obtain 95% confidence intervals together with the estimations
MLCSW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$ProbB, DatMB$ProbA,
DatMA$Domain, DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, N_FrameA,
N_FrameB, N_Domainab, conf_level = 0.95)
```

---

MLDF *Multinomial logistic estimator under dual frame approach with auxiliary information from each frame*

---

### Description

Produces estimates for class totals and proportions using multinomial logistic regression from survey data obtained from a dual frame sampling design using a model assisted approach with a possibly different set of auxiliary variables for each frame. Confidence intervals are also computed, if required.

### Usage

```
MLDF (ysA, ysB, pik_A, pik_B, domains_A, domains_B, xsA, xsB, xA, xB, ind_samA,
ind_samB, ind_domA, ind_domB, N, conf_level = NULL)
```

### Arguments

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xsA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .

xA	A numeric vector or length $N_A$ or a numeric matrix or data frame of dimensions $N_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information for the units in frame A.
xB	A numeric vector or length $N_B$ or a numeric matrix or data frame of dimensions $N_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information for the units in frame B.
ind_samA	A numeric vector of length $n_A$ containing the identifiers of units of the frame A (from 1 to $N_A$ ) that belongs to $s_A$ .
ind_samB	A numeric vector of length $n_B$ containing the identifiers of units of the frame B (from 1 to $N_B$ ) that belongs to $s_B$ .
ind_domA	A character vector of length $N_A$ indicating the domain each unit from frame A belongs to. Possible values are "a" and "ab".
ind_domB	A character vector of length $N_B$ indicating the domain each unit from frame B belongs to. Possible values are "b" and "ba".
N	A numeric value indicating the size of the population.
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

### Details

Multinomial logistic estimator in dual frame using auxiliary information from each frame for a proportion is given by

$$\hat{P}_{MLi}^{DF} = \frac{1}{N} \left( \sum_{k \in U_a} p_{ki}^A + \eta \sum_{k \in U_{ab}} p_{ki}^A + (1 - \eta) \sum_{k \in U_{ba}} p_{ki}^B + \sum_{k \in U_b} p_{ki}^B \right. \\ \left. + \sum_{k \in s_a} d_k^A (z_{ki} - p_{ki}^A) + \eta \sum_{k \in s_{ab}} d_k^A (z_{ki} - p_{ki}^A) \right. \\ \left. + (1 - \eta) \sum_{k \in s_{ba}} d_k^B (z_{ki} - p_{ki}^B) + \sum_{k \in s_b} d_k^B (z_{ki} - p_{ki}^B) \right), \quad i = 1, \dots, m$$

with  $\eta \in (0, 1)$ ,  $m$  the number of categories of the response variable,  $z_i$  the indicator variable for the  $i$ -th category of the response variable,  $d^A$  and  $d^B$  the design weights for each frame, defined as the inverse of the first order inclusion probabilities and

$$p_{ki}^A = \frac{\exp(x_k' \beta_i^A)}{\sum_{r=1}^m \exp(x_k' \beta_r^A)},$$

being  $\beta_i^A$  the maximum likelihood parameters of the multinomial logistic model considering weights  $d^A$ .  $p_{ki}^B$  can be defined similarly.

### Value

MLDF returns an object of class "MultEstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	class frequencies and proportions estimations for main variable(s).

## References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

Lehtonen, R. and Veijanen, A. (1998) *On multinomial logistic generalizaed regression estimators* Technical report 22, Department of Statistics, University of Jyvaskyla.

## See Also

[JackMLDF](#)

## Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)

N <- nrow(DatPopM)
levels(DatPopM$Domain) <- c(levels(DatPopM$Domain), "ba")
DatPopMA <- subset(DatPopM, DatPopM$Domain == "a" | DatPopM$Domain == "ab", stringAsFactors = FALSE)
DatPopMB <- subset(DatPopM, DatPopM$Domain == "b" | DatPopM$Domain == "ab", stringAsFactors = FALSE)
DatPopMB[DatPopMB$Domain == "ab",]$Domain <- "ba"

#Let calculate proportions of categories of variable Prog using MLDF estimator
#using Read as auxiliary variable
MLDF(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopMA$Read, DatPopMB$Read, DatMA$Id_Frame, DatMB$Id_Frame,
DatPopMA$Domain, DatPopMB$Domain, N)

#Let obtain 95% confidence intervals together with the estimations
MLDF(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopMA$Read, DatPopMB$Read, DatMA$Id_Frame, DatMB$Id_Frame,
DatPopMA$Domain, DatPopMB$Domain, N, conf_level = 0.95)
```

---

MLDW

*Multinomial logistic estimator under dual frame approach with auxiliary information from the whole population*

---

## Description

Produces estimates for class totals and proportions using multinomial logistic regression from survey data obtained from a dual frame sampling design using a dual frame model assisted approach. Confidence intervals are also computed, if required.

## Usage

```
MLDW (ysA, ysB, pik_A, pik_B, domains_A, domains_B, xsA, xsB, x, ind_sam,
conf_level = NULL)
```

**Arguments**

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xsA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
x	A numeric vector or length $N$ or a numeric matrix or data frame of dimensions $N \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information for every unit in the population.
ind_sam	A numeric vector of length $n = n_A + n_B$ containing the identifiers of units of the population (from 1 to $N$ ) that belongs to $s_A$ or $s_B$
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

**Details**

Multinomial logistic estimator in dual frame using auxiliary information from the whole population for a proportion is given by

$$\hat{P}_{MLi}^{DW} = \frac{1}{N} \left( \sum_{k \in U} p_{ki}^{\circ} + \sum_{k \in s} d_k^{\circ} (z_{ki} - p_{ki}^{\circ}) \right) \quad i = 1, \dots, m$$

with  $m$  the number of categories of the response variable,  $z_i$  the indicator variable for the  $i$ -th

$$\text{category of the response variable, } d_k^{\circ} = \begin{cases} d_k^A & \text{if } k \in a \\ \eta d_k^A & \text{if } k \in ab \\ (1 - \eta) d_k^B & \text{if } k \in ba \\ d_k^B & \text{if } k \in b \end{cases} \quad \text{with } \eta \in (0, 1) \text{ and}$$

$$p_{ki}^{\circ} = \frac{\exp(x'_k \beta_i^{\circ})}{\sum_{r=1}^m \exp(x'_k \beta_r^{\circ})},$$

being  $\beta_i^{\circ}$  the maximum likelihood parameters of the multinomial logistic model considering the weights  $d^{\circ}$ .

**Value**

MLDW returns an object of class "MultEstimatorDF" which is a list with, at least, the following components:

Call                    the matched call.  
 Est                    class frequencies and proportions estimations for main variable(s).

**References**

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

Lehtonen, R. and Veijanen, A. (1998) *On multinomial logistic generalizaed regression estimators* Technical report 22, Department of Statistics, University of Jyväskylä.

**See Also**

[JackMLDW](#)

**Examples**

```
data(DatMA)
data(DatMB)
data(DatPopM)

IndSample <- c(DatMA$Id_Pop, DatMB$Id_Pop)
#Let calculate proportions of categories of variable Prog using MLDW estimator
#using Read as auxiliary variable
MLDW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopM$Read, IndSample)

#Let obtain 95% confidence intervals together with the estimations
MLDW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$Domain, DatMB$Domain,
DatMA$Read, DatMB$Read, DatPopM$Read, IndSample, 0.95)
```

---

MLSW

*Multinomial logistic estimator under single frame approach with auxiliary information from the whole population*

---

**Description**

Produces estimates for class totals and proportions using multinomial logistic regression from survey data obtained from a dual frame sampling design with the same set of auxiliary variables for the whole population. Confidence intervals are also computed, if required.

**Usage**

```
MLSW(ysA, ysB, pik_A, pik_B, pik_ab_B, pik_ba_A, domains_A, domains_B, xsA, xsB,
x, ind_sam, conf_level = NULL)
```

**Arguments**

ysA	A data frame containing information about one or more factors, each one of dimension $n_A$ , collected from $s_A$ .
ysB	A data frame containing information about one or more factors, each one of dimension $n_B$ , collected from $s_B$ .
pik_A	A numeric vector of length $n_A$ containing first order inclusion probabilities for units included in $s_A$ .
pik_B	A numeric vector of length $n_B$ containing first order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
xsA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame A for units included in $s_A$ .
xsB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information in frame B for units included in $s_B$ .
x	A numeric vector or length $N$ or a numeric matrix or data frame of dimensions $N \times m$ , with $m$ the number of auxiliary variables, containing auxiliary information for every unit in the population.
ind_sam	A numeric vector of length $n = n_A + n_B$ containing the identifiers of units of the population (from 1 to $N$ ) that belongs to $s_A$ or $s_B$
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

**Details**

Multinomial logistic estimator in single frame using auxiliary information from the whole population for a proportion is given by

$$\hat{P}_{MLi}^{SW} = \frac{1}{N} \left( \sum_{k \in U} \tilde{p}_{ki} + \sum_{k \in s} \tilde{d}_k (z_{ki} - \tilde{p}_{ki}) \right) \quad i = 1, \dots, m$$

with  $m$  the number of categories of the response variable,  $z_i$  the indicator variable for the  $i$ -th category of the response variable,  $\tilde{d}_k = \begin{cases} d_k^A & \text{if } k \in a \\ (1/d_k^A + 1/d_k^B)^{-1} & \text{if } k \in ab \cup ba \\ d_k^B & \text{if } k \in b \end{cases}$  and

$$\tilde{p}_{ki} = \frac{\exp(x'_k \tilde{\beta}_i)}{\sum_{r=1}^m \exp(x'_k \tilde{\beta}_r)},$$

being  $\tilde{\beta}_i$  the maximum likelihood parameters of the multinomial logistic model considering weights  $\tilde{d}$ .

### Value

PMLSW returns an object of class "MultEstimatorDF" which is a list with, at least, the following components:

Call                    the matched call.  
Est                     class frequencies and proportions estimations for main variable(s).

### References

Molina, D., Rueda, M., Arcos, A. and Ranalli, M. G. (2015) *Multinomial logistic estimation in dual frame surveys* Statistics and Operations Research Transactions (SORT). To be printed.

Lehtonen, R. and Veijanen, A. (1998) *On multinomial logistic generalised regression estimators* Technical report 22, Department of Statistics, University of Jyväskylä.

### See Also

[JackMLSW](#)

### Examples

```
data(DatMA)
data(DatMB)
data(DatPopM)

IndSample <- c(DatMA$Id_Pop, DatMB$Id_Pop)
#Let calculate proportions of categories of variable Prog using MLSW estimator
#using Read as auxiliary variable
MLSW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$ProbB, DatMB$ProbA,
DatMA$Domain, DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read, IndSample)

#Let obtain 95% confidence intervals together with the estimations
MLSW(DatMA$Prog, DatMB$Prog, DatMA$ProbA, DatMB$ProbB, DatMA$ProbB, DatMB$ProbA,
DatMA$Domain, DatMB$Domain, DatMA$Read, DatMB$Read, DatPopM$Read, IndSample,
conf_level = 0.95)
```

---

 PEL

*Pseudo empirical likelihood estimator*


---

### Description

Produces estimates for population totals using the pseudo empirical likelihood estimator from survey data obtained from a dual frame sampling design. Confidence intervals for the population total are also computed, if required.

### Usage

```
PEL(ysA, ysB, pi_A, pi_B, domains_A, domains_B, N_A = NULL, N_B = NULL,
    N_ab = NULL, xsAframeA = NULL, xsBframeA = NULL, xsAframeB = NULL, xsBframeB = NULL,
    XA = NULL, XB = NULL, conf_level = NULL)
```

### Arguments

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable(s) of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable(s) of interest from $s_B$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	(Optional) A numeric value indicating the size of frame A.
N_B	(Optional) A numeric value indicating the size of frame B.
N_ab	(Optional) A numeric value indicating the size of the overlap domain.
xsAframeA	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsBframeA	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.
xsAframeB	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.

xsBFrameB	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

### Details

Pseudo empirical likelihood estimator for the population mean is computed as

$$\hat{Y}_{PEL} = \frac{N_a}{N} \hat{Y}_a + \frac{\eta N_{ab}}{N} \hat{Y}_{ab}^A + \frac{(1-\eta)N_{ab}}{N} \hat{Y}_{ab}^B + \frac{N_b}{N} \hat{Y}_b$$

where  $\hat{Y}_a = \sum_{k \in s_a} \hat{p}_{ak} y_k$ ,  $\hat{Y}_{ab} = \sum_{k \in s_{ab}^A} \hat{p}_{abk}^A y_k$ ,  $\hat{Y}_{ab}^B = \sum_{k \in s_{ab}^B} \hat{p}_{abk}^B y_k$  and  $\hat{Y}_b = \sum_{k \in s_b} \hat{p}_{bk} y_k$  with  $\hat{p}_{ak}$ ,  $\hat{p}_{abk}^A$ ,  $\hat{p}_{abk}^B$  and  $\hat{p}_{bk}$  the weights resulting of applying the pseudo empirical likelihood procedure to a determined function under a determined set of constraints, depending on the case. Furthermore,  $\eta \in (0, 1)$ . In this case,  $N_A$ ,  $N_B$  and  $N_{ab}$  have been supposed known and no additional auxiliary variables have been considered. This is not happening in some cases. Function covers following scenarios:

- There is not any additional auxiliary variable
  - $N_A$ ,  $N_B$  and  $N_{ab}$  unknown
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_A$ ,  $N_B$  and  $N_{ab}$  known
- At least, one additional auxiliary variable is available
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_A$ ,  $N_B$  and  $N_{ab}$  known

Explicit variance of this estimator is not easy to obtain. Instead, confidence intervals can be computed through the bi-section method. This method constructs intervals in the form  $\{\theta | r_{ns}(\theta) < \chi_1^2(\alpha)\}$ , where  $\chi_1^2(\alpha)$  is the  $1 - \alpha$  quantile from a  $\chi^2$  distribution with one degree of freedom and  $r_{ns}(\theta)$  represents the so called pseudo empirical log likelihood ratio statistic, which can be obtained as a difference of two pseudo empirical likelihood functions.

### Value

PEL returns an object of class "EstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	total and mean estimation for main variable(s).
VarEst	variance estimation for main variable(s).

If parameter `conf_level` is different from `NULL`, object includes component

`ConfInt`            total and mean estimation and confidence intervals for main variables(s).

In addition, components `TotDomEst` and `MeanDomEst` are available when estimator is based on estimators of the domains. Component `Param` shows value of parameters involved in calculation of the estimator (if any). By default, only `Est` component (or `ConfInt` component, if parameter `conf_level` is different from `NULL`) is shown. It is possible to access to all the components of the objects by using function `summary`.

## References

Rao, J. N. K. and Wu, C. (2010) *Pseudo Empirical Likelihood Inference for Multiple Frame Surveys*. Journal of the American Statistical Association, 105, 1494 - 1503.

Wu, C. (2005) *Algorithms and R codes for the pseudo empirical likelihood methods in survey sampling*. Survey Methodology, Vol. 31, 2, pp. 239 - 243.

## See Also

[JackPEL](#)

## Examples

```
data(Data)
data(DatB)
data(Pik1A)
data(Pik1B)

#Let calculate pseudo empirical likelihood estimator for variable Feeding, without
#considering any auxiliary information
PEL(Data$Feed, DatB$Feed, Pik1A, Pik1B, Data$Domain, DatB$Domain)

#Now, let calculate pseudo empirical estimator for variable Clothing when the frame
#sizes and the overlap domain size are known
PEL(Data$Clo, DatB$Clo, Pik1A, Pik1B, Data$Domain, DatB$Domain,
N_A = 1735, N_B = 1191, N_ab = 601)

#Finally, let calculate pseudo empirical likelihood estimator and a 90% confidence interval
#for population total for variable Feeding, considering Income and Metres2 as auxiliary
#variables and with frame sizes and overlap domain size known.
PEL(Data$Feed, DatB$Feed, Pik1A, Pik1B, Data$Domain, DatB$Domain,
N_A = 1735, N_B = 1191, N_ab = 601, xsAFrameA = Data$Inc, xsBFrameA = DatB$Inc,
xsAFrameB = Data$M2, xsBFrameB = DatB$M2, XA = 4300260, XB = 176553,
conf_level = 0.90)
```

---

PiklA	<i>Matrix of inclusion probabilities for units selected in sample from frame A</i>
-------	--

---

**Description**

This dataset consists of a square matrix of dimension 105 with the first and second order inclusion probabilities for the units included in sample  $s_A$ , which has been drawn from a population of size  $N_A = 1735$  according to a stratified random sampling with population strata sizes  $N_A^h = (727, 375, 113, 186, 115, 219)$

**Usage**

```
PiklA
```

**See Also**

[DatA](#)

**Examples**

```
data(PiklA)
#Let choose the submatrix of inclusion probabilities for the first 5 units sA.
PiklA[1:5, 1:5]
#Now, let select only the first order inclusion probabilities
diag(PiklA)
```

---

PiklB	<i>Matrix of inclusion probabilities for units selected in sample from frame B</i>
-------	--

---

**Description**

This dataset consists of a square matrix of dimension 135 with the first and second order inclusion probabilities for the units included in  $s_B$ , which has been drawn from a population of size  $N_B = 1191$  according to a simple random sampling without replacement.

**Usage**

```
PiklB
```

**See Also**

[DatB](#)

**Examples**

```

data(PiklB)
#Let choose the submatrix of inclusion probabilities for the first 5 units in sB.
PiklB[1:5, 1:5]
#Now, let select the first order inclusion probabilities
diag(PiklB)

```

---

PML

*Pseudo Maximum Likelihood estimator*


---

**Description**

Produces estimates for population totals and means using PML estimator from survey data obtained from a dual frame sampling design. Confidence intervals are also computed, if required.

**Usage**

```
PML(ysA, ysB, pi_A, pi_B, domains_A, domains_B, N_A, N_B, conf_level = NULL)
```

**Arguments**

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable of interest from $s_B$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	A numeric value indicating the size of frame A
N_B	A numeric value indicating the size of frame B
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

### Details

Pseudo Maximum Likelihood estimator of population total is given by

$$\hat{Y}_{PML}(\hat{\theta}) = \frac{N_A - \hat{N}_{ab,PML}}{\hat{N}_a} \hat{Y}_a^A + \frac{N_B - \hat{N}_{ab,PML}}{\hat{N}_b} \hat{Y}_b^B + \frac{\hat{N}_{ab,PML}}{\hat{\theta} \hat{N}_{ab}^A + (1 - \hat{\theta}) \hat{N}_{ab}^B} [\hat{\theta} \hat{Y}_{ab}^A + (1 - \hat{\theta}) \hat{Y}_{ab}^B]$$

where  $\hat{\theta} \in [0, 1]$  and  $\hat{N}_{ab,PML}$  is the smaller of the roots of the quadratic equation

$$[\hat{\theta}/N_B + (1 - \hat{\theta})/N_A]x^2 - [1 + \hat{\theta}\hat{N}_{ab}^A/N_B + (1 - \hat{\theta})\hat{N}_{ab}^B/N_A]x + \hat{\theta}\hat{N}_{ab}^A + (1 - \hat{\theta})\hat{N}_{ab}^B = 0.$$

Optimal value for  $\hat{\theta}$  is  $\frac{\hat{N}_a N_B \hat{V}(\hat{N}_{ab}^B)}{\hat{N}_a N_B \hat{V}(\hat{N}_{ab}^B) + \hat{N}_b N_A \hat{V}(\hat{N}_{ab}^A)}$ . Variance is estimated according to following expression

$$\hat{V}(\hat{Y}_{PML}(\hat{\theta})) = \hat{V}\left(\sum_{i \in s_A} \tilde{z}_i^A\right) + \hat{V}\left(\sum_{i \in s_B} \tilde{z}_i^B\right)$$

where,  $\tilde{z}_i^A = y_i - \frac{\hat{Y}_a}{\hat{N}_a}$  if  $i \in a$  and  $\tilde{z}_i^A = \hat{\gamma}_{opt}(y_i - \frac{\hat{Y}_a}{\hat{N}_a}) + \hat{\lambda}\hat{\phi}$  if  $i \in ab$  with

$$\hat{\gamma}_{opt} = \frac{\hat{N}_a N_B \hat{V}(\hat{N}_{ab}^B)}{\hat{N}_a N_B \hat{V}(\hat{N}_{ab}^B) + \hat{N}_b + N_A + \hat{V}(\hat{N}_{ab}^A)}$$

$$\hat{\lambda} = \frac{n_A/N_A \hat{Y}_{ab}^A + n_B/N_B \hat{Y}_{ab}^B}{n_A/N_A \hat{N}_{ab}^A + n_B/N_B \hat{N}_{ab}^B} - \frac{\hat{Y}_a}{\hat{N}_a} - \frac{\hat{Y}_b}{\hat{N}_b}$$

$$\hat{\phi} = \frac{n_A \hat{N}_b}{n_A \hat{N}_b + n_B \hat{N}_a}$$

Similarly, we define  $\tilde{z}_i^B = y_i - \frac{\hat{Y}_b}{\hat{N}_b}$  if  $i \in b$  and  $\tilde{z}_i^B = (1 - \hat{\gamma}_{opt})(y_i - \frac{\hat{Y}_b}{\hat{N}_b}) + \hat{\lambda}(1 - \hat{\phi})$  if  $i \in ba$

### Value

PML returns an object of class "EstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	total and mean estimation for main variable(s).
VarEst	variance estimation for main variable(s).

If parameter `conf_level` is different from NULL, object includes component

ConfInt	total and mean estimation and confidence intervals for main variables(s).
---------	---

In addition, components `TotDomEst` and `MeanDomEst` are available when estimator is based on estimators of the domains. Component `Param` shows value of parameters involded in calculation of the estimator (if any). By default, only `Est` component (or `ConfInt` component, if parameter `conf_level` is different from NULL) is shown. It is possible to access to all the components of the objects by using function `summary`.

### References

Skinner, C. J. and Rao, J. N. K. (1996) *Estimation in Dual Frame Surveys with Complex Designs*. Journal of the American Statistical Association, Vol. 91, 433, 349 - 356.

**Examples**

```

data(DatA)
data(DatB)
data(PiklA)
data(PiklB)

#Let calculate Pseudo Maximum Likelihood estimator for population total for variable Clothing
PML(DatA$Clo, DatB$Clo, PiklA, PiklB, DatA$Domain, DatB$Domain,
N_A = 1735, N_B = 1191)

#Now, let calculate Pseudo Maximum Likelihood estimator for population total for variable
#Feeding, using first order inclusion probabilities
PML(DatA$Feed, DatB$Feed, DatA$ProbA, DatB$ProbB, DatA$Domain, DatB$Domain,
N_A = 1735, N_B = 1191)

#Finally, let calculate Pseudo Maximum Likelihood estimator and a 90% confidence interval for
#population total for variable Leisure
PML(DatA$Lei, DatB$Lei, PiklA, PiklB, DatA$Domain, DatB$Domain,
N_A = 1735, N_B = 1191, 0.90)

```

---

SFRR

*Raking ratio estimator*


---

**Description**

Produces estimates for population total and mean using the raking ratio estimator from survey data obtained from a dual frame sampling desing. Confidence intervals are also computed, if required.

**Usage**

```
SFRR(ysA, ysB, pi_A, pi_B, pik_ab_B, pik_ba_A, domains_A, domains_B, N_A, N_B,
conf_level = NULL)
```

**Arguments**

<code>ysA</code>	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable of interest from $s_A$ .
<code>ysB</code>	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable of interest from $s_B$ .
<code>pi_A</code>	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
<code>pi_B</code>	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
<code>pik_ab_B</code>	A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling desing in frame B for units belonging to overlap domain that have been selected in $s_A$ .

pik_ba_A	A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_A$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_A$ belongs to. Possible values are "b" and "ba".
N_A	A numeric value indicating the size of frame A
N_B	A numeric value indicating the size of frame B
conf_level	(Optional) A numeric value indicating the confidence level for the confidence intervals, if desired.

### Details

Raking ratio estimator of population total is given by

$$\hat{Y}_{SFRR} = \frac{N_A - \hat{N}_{ab,rake}}{\hat{N}_a^A} \hat{Y}_a^A + \frac{N_B - \hat{N}_{ab,rake}}{\hat{N}_b^B} \hat{Y}_b^B + \frac{\hat{N}_{ab,rake}}{\hat{N}_{abS}} \hat{Y}_{abS}$$

where  $\hat{Y}_{abS} = \sum_{i \in s_{ab}^A} \tilde{d}_i^A y_i + \sum_{i \in s_{ab}^B} \tilde{d}_i^B y_i$ ,  $\hat{N}_{abS} = \sum_{i \in s_{ab}^A} \tilde{d}_i^A + \sum_{i \in s_{ab}^B} \tilde{d}_i^B$  and  $\hat{N}_{ab,rake}$  is the smallest root of the quadratic equation  $\hat{N}_{ab,rake} x^2 - [\hat{N}_{ab,rake}(N_A + N_B) + \hat{N}_{aS} \hat{N}_{bS}] x + \hat{N}_{ab,rake} N_A N_B = 0$ , with  $\hat{N}_{aS} = \sum_{s_a^A} \tilde{d}_i^B$  and  $\hat{N}_{bS} = \sum_{s_b^B} \tilde{d}_i^B$ . Weights  $\tilde{d}_i^A$  and  $\tilde{d}_i^B$  are obtained as follows  $\tilde{d}_i^A = \begin{cases} d_i^A & \text{if } i \in a \\ (1/d_i^A + 1/d_i^B)^{-1} & \text{if } i \in ab \end{cases}$  and  $\tilde{d}_i^B = \begin{cases} d_i^B & \text{if } i \in b \\ (1/d_i^A + 1/d_i^B)^{-1} & \text{if } i \in ba \end{cases}$  being  $d_i^A$  and  $d_i^B$  the design weights, obtained as the inverse of the first order inclusion probabilities, that is  $d_i^A = 1/\pi_i^A$  and  $d_i^B = 1/\pi_i^B$ .

To obtain an estimator of the variance for this estimator, one has taken into account that raking ratio estimator coincides with SF calibration estimator when frame sizes are known and "raking" method is used. So, one can use here Deville's expression to calculate an estimator for the variance of the raking ratio estimator

$$\hat{V}(\hat{Y}_{SFRR}) = \frac{1}{1 - \sum_{k \in s} a_k^2} \sum_{k \in s} (1 - \pi_k) \left( \frac{e_k}{\pi_k} - \sum_{l \in s} a_l \frac{e_l}{\pi_l} \right)^2$$

where  $a_k = (1 - \pi_k) / \sum_{l \in s} (1 - \pi_l)$  and  $e_k$  are the residuals of the regression with auxiliary variables as regressors.

### Value

SFRR returns an object of class "EstimatorDF" which is a list with, at least, the following components:

Call	the matched call.
Est	total and mean estimation for main variable(s).
VarEst	variance estimation for main variable(s).

If parameter `conf_level` is different from `NULL`, object includes component

`ConfInt`            total and mean estimation and confidence intervals for main variables(s).

In addition, components `TotDomEst` and `MeanDomEst` are available when estimator is based on estimators of the domains. Component `Param` shows value of parameters involved in calculation of the estimator (if any). By default, only `Est` component (or `ConfInt` component, if parameter `conf_level` is different from `NULL`) is shown. It is possible to access to all the components of the objects by using function `summary`.

## References

- Lohr, S. and Rao, J.N.K. (2000). *Inference in Dual Frame Surveys*. Journal of the American Statistical Association, Vol. 95, 271 - 280.
- Rao, J.N.K. and Skinner, C.J. (1996). *Estimation in Dual Frame Surveys with Complex Designs*. Proceedings of the Survey Method Section, Statistical Society of Canada, 63 - 68.
- Skinner, C.J. and Rao J.N.K. (1996). *Estimation in Dual Frame Surveys with Complex Designs*. Journal of the American Statistical Association, Vol. 91, 443, 349 - 356.
- Skinner, C.J. (1991). *On the Efficiency of Raking Ratio Estimation for Multiple Frame Surveys*. Journal of the American Statistical Association, Vol. 86, 779 - 784.

## See Also

[JackSFRR](#)

## Examples

```
data(Data)
data(DatB)
data(Pik1A)
data(Pik1B)

#Let calculate raking ratio estimator for population total for variable Clothing
SFRR(Data$Clo, DatB$Clo, Pik1A, Pik1B, Data$ProbB, DatB$ProbA, Data$Domain,
DatB$Domain, 1735, 1191)

#Now, let calculate raking ratio estimator and a 90% confidence interval for
#population total for variable Feeding, considering only first order inclusion probabilities
SFRR(Data$Feed, DatB$Feed, Data$ProbA, DatB$ProbB, Data$ProbB, DatB$ProbA,
Data$Domain, DatB$Domain, 1735, 1191, 0.90)
```

---

VarHT

*Variance estimator of Horvitz - Thompson estimator*

---

## Description

Computes the variance estimator of Horvitz - Thompson estimator of population total

**Usage**

```
VarHT(y, pikl)
```

**Arguments**

**y** A numeric vector of size n containing information about variable of interest

**pikl** A square numeric matrix of dimension n containing first and second order inclusion probabilities for units included in y

**Details**

Variance estimator of Horvitz - Thompson estimator of population total is given by

$$\widehat{Var}(\widehat{Y}_{HT}) = \sum_{k \in s} \frac{y_k^2}{\pi_k^2} (1 - \pi_k) + \sum_{k \in s} \sum_{l \in s, l \neq k} \frac{y_k y_l}{\pi_k \pi_l} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}}$$

**Value**

A numeric value representing variance estimator of Horvitz - Thompson estimator for population total for considered values

**References**

Horvitz, D. G. and Thompson, D. J. (1952) *A generalization of sampling without replacement from a finite universe*. Journal of the American Statistical Association, 47, 663 - 685

Sarndal, C. E., Swensson, B. and Wretman, J. (1992) *Model Assisted Survey Sampling*. Springer-Verlag. New York.

**See Also**

[HT CovHT](#)

**Examples**

```
##### Example 1 #####
U <- c(13, 18, 20, 14, 9)
#A simple random sample of size 2 without replacement is drawn from population
s <- sample(U, 2)
#Horvitz - Thompson estimator of population total is calculated.
ps <- c(0.4, 0.4)
HT(s, ps)
#Now, we calculate variance estimator of the Horvitz - Thompson estimator.
Ps <- matrix(c(0.4,0.1, 0.1,0.4), 2 ,2)
VarHT(s, Ps)

##### Example 2 #####
data(DataA)
attach(DataA)
data(PiklA)
```

```
#Let calculate Horvitz - Thompson estimator for total of variable Clothing in Frame A.
HT(Clo, ProbA)
#And now, let compute the variance of the previous estimator
VarHT(Clo, Pik1A)
```

---

WeightsCalDF

*g-weights for the dual frame calibration estimator*


---

## Description

Computes the g-weights for the dual frame calibration estimator.

## Usage

```
WeightsCalDF(ysA, ysB, pi_A, pi_B, domains_A, domains_B, N_A = NULL, N_B = NULL,
N_ab = NULL, xsAFrameA = NULL, xsBFrameA = NULL, xsAFrameB = NULL, xsBFrameB = NULL,
xsT = NULL, XA = NULL, XB = NULL, X = NULL, met = "linear")
```

## Arguments

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable(s) of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable(s) of interest from $s_B$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .
pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
domains_A	A character vector of length $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of length $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	(Optional) A numeric value indicating the size of frame A.
N_B	(Optional) A numeric value indicating the size of frame B.
N_ab	(Optional) A numeric value indicating the size of the overlap domain.
xsAFrameA	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsBFrameA	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.

xsAFrameB	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.
xsBFrameB	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
xsT	(Optional) A numeric vector of length $n$ or a numeric matrix or data frame of dimensions $n \times m_T$ , with $m_T$ the number of auxiliary variables in both frames, containing auxiliary information for all units in the entire sample $s = s_A \cup s_B$ .
XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
X	(Optional) A numeric value or vector of length $m_T$ , with $m_T$ the number of auxiliary variables in both frames, indicating the population totals for the auxiliary variables considered in both frames.
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".

### Details

Function provides g-weights in following scenarios:

- There is not any additional auxiliary variable
  - $N_A, N_B$  and  $N_{ab}$  unknown
  - $N_{ab}$  known and  $N_A$  and  $N_B$  unknown
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_A, N_B$  and  $N_{ab}$  known
- At least, one additional auxiliary variable is available
  - $N_{ab}$  known and  $N_A$  and  $N_B$  unknown
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_A, N_B$  and  $N_{ab}$  known

### Value

A numeric vector containing the g-weights for the dual frame calibration estimator.

### References

- Ranalli, M. G., Arcos, A., Rueda, M. and Teodoro, A. (2013) *Calibration estimationonn in dual frame surveys*. arXiv:1312.0761 [stat.ME]
- Deville, J. C., Sv"arndal, C. E. (1992) *Calibration estimators in survey sampling*. Journal of the American Statistical Association, 87, 376 - 382

**Examples**

```

data(Data)
data(DatB)
data(PiklA)
data(PiklB)

#Let calculate g-weights for the dual frame calibration estimator for variable Feeding,
#without considering any auxiliary information
WeightsCalDF(Data$Feed, DatB$Feed, PiklA, PiklB, DatA$Domain, DatB$Domain)

#Now, let calculate g-weights for the dual frame calibration estimator for variable Clothing
#when the frame sizes and the overlap domain size are known
WeightsCalDF(Data$Clo, DatB$Clo, PiklA, PiklB, DatA$Domain, DatB$Domain,
N_A = 1735, N_B = 1191, N_ab = 601)

#Finally, let calculate g-weights for the dual frame calibration estimator
#for variable Feeding, considering Income as auxiliary variable in frame A
#and Metres2 as auxiliary variable in frame B and with frame sizes and overlap
#domain size known.
WeightsCalDF(Data$Feed, DatB$Feed, PiklA, PiklB, DatA$Domain, DatB$Domain,
N_A = 1735, N_B = 1191, N_ab = 601, xsAFrameA = DatA$Inc, xsBFrameA = DatB$Inc,
xsAFrameB = DatA$M2, xsBFrameB = DatB$M2, XA = 4300260, XB = 176553)

```

---

WeightsCalSF

*g-weights for the SF calibration estimator*


---

**Description**

Computes the g-weights for the SF calibration estimator.

**Usage**

```

WeightsCalSF(ysA, ysB, pi_A, pi_B, pik_ab_B, pik_ba_A, domains_A, domains_B,
N_A = NULL, N_B = NULL, N_ab = NULL, xsAFrameA = NULL, xsBFrameA = NULL,
xsAFrameB = NULL, xsBFrameB = NULL, xsT = NULL, XA = NULL, XB = NULL, X = NULL,
met = "linear")

```

**Arguments**

ysA	A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times c$ containing information about variable(s) of interest from $s_A$ .
ysB	A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times c$ containing information about variable(s) of interest from $s_A$ .
pi_A	A numeric vector of length $n_A$ or a square numeric matrix of dimension $n_A$ containing first order or first and second order inclusion probabilities for units included in $s_A$ .

pi_B	A numeric vector of length $n_B$ or a square numeric matrix of dimension $n_B$ containing first order or first and second order inclusion probabilities for units included in $s_B$ .
pik_ab_B	A numeric vector of size $n_A$ containing first order inclusion probabilities according to sampling design in frame B for units belonging to overlap domain that have been selected in $s_A$ .
pik_ba_A	A numeric vector of size $n_B$ containing first order inclusion probabilities according to sampling design in frame A for units belonging to overlap domain that have been selected in $s_B$ .
domains_A	A character vector of size $n_A$ indicating the domain each unit from $s_A$ belongs to. Possible values are "a" and "ab".
domains_B	A character vector of size $n_B$ indicating the domain each unit from $s_B$ belongs to. Possible values are "b" and "ba".
N_A	(Optional) A numeric value indicating the size of frame A
N_B	(Optional) A numeric value indicating the size of frame B
N_ab	(Optional) A numeric value indicating the size of the overlap domain
xsAFrameA	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_A$ .
xsBFrameA	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_A$ , with $m_A$ the number of auxiliary variables in frame A, containing auxiliary information in frame A for units included in $s_B$ . For units in domain $b$ , these values are 0.
xsAFrameB	(Optional) A numeric vector of length $n_A$ or a numeric matrix or data frame of dimensions $n_A \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_A$ . For units in domain $a$ , these values are 0.
xsBFrameB	(Optional) A numeric vector of length $n_B$ or a numeric matrix or data frame of dimensions $n_B \times m_B$ , with $m_B$ the number of auxiliary variables in frame B, containing auxiliary information in frame B for units included in $s_B$ .
xsT	(Optional) A numeric vector of length $n$ or a numeric matrix or data frame of dimensions $n \times m_T$ , with $m_T$ the number of auxiliary variables in both frames, containing auxiliary information for all units in the entire sample $s = s_A \cup s_B$ .
XA	(Optional) A numeric value or vector of length $m_A$ , with $m_A$ the number of auxiliary variables in frame A, indicating the population totals for the auxiliary variables considered in frame A.
XB	(Optional) A numeric value or vector of length $m_B$ , with $m_B$ the number of auxiliary variables in frame B, indicating the population totals for the auxiliary variables considered in frame B.
X	(Optional) A numeric value or vector of length $m_T$ , with $m_T$ the number of auxiliary variables in both frames, indicating the population totals for the auxiliary variables considered in both frames.
met	(Optional) A character vector indicating the distance that must be used in calibration process. Possible values are "linear", "raking" and "logit". Default is "linear".

## Details

Function provides g-weights in following scenarios:

- There is not any additional auxiliary variable
  - $N_A$ ,  $N_B$  and  $N_{ab}$  unknown
  - $N_{ab}$  known and  $N_A$  and  $N_B$  unknown
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_A$ ,  $N_B$  and  $N_{ab}$  known
- At least, one additional auxiliary variable is available
  - $N_{ab}$  known and  $N_A$  and  $N_B$  unknown
  - $N_A$  and  $N_B$  known and  $N_{ab}$  unknown
  - $N_A$ ,  $N_B$  and  $N_{ab}$  known

## Value

A numeric vector containing the g-weights for the SF calibration estimator.

## References

Ranalli, M. G., Arcos, A., Rueda, M. and Teodoro, A. (2013) *Calibration estimationonn in dual frame surveys*. arXiv:1312.0761 [stat.ME]

Deville, J. C., SV"arndal, C. E. (1992) *Calibration estimators in survey sampling*. Journal of the American Statistical Association, 87, 376 - 382

## Examples

```
data(Data)
data(DatB)
data(PiklA)
data(PiklB)

#Let calculate g-weights for the SF calibration estimator for variable Clothing,
#without considering any auxiliary information
WeightsCalSF(DatA$Clo, DatB$Clo, PiklA, PiklB, DatA$ProbB, DatB$ProbA,
Data$Domain, DatB$Domain)

#Now, let calculate g-weights for the SF calibration estimator for variable Leisure
#when the frame sizes and the overlap domain size are known
WeightsCalSF(DatA$Lei, DatB$Lei, PiklA, PiklB, DatA$ProbB, DatB$ProbA,
Data$Domain, DatB$Domain, N_A = 1735, N_B = 1191, N_ab = 601)

#Finally, let calculate g-weights for the SF calibration estimator
#for variable Feeding, considering Income and Metres2 as auxiliary
#variables and with frame sizes and overlap domain size known.
WeightsCalSF(DatA$Feed, DatB$Feed, PiklA, PiklB, DatA$ProbB, DatB$ProbA,
Data$Domain, DatB$Domain, N_A = 1735, N_B = 1191, N_ab = 601, xsAFrameA = DatA$Inc,
xsBFrameA = DatB$Inc, xsAFrameB = DatA$M2, xsBFrameB = DatB$M2,
XA = 4300260, XB = 176553)
```

# Index

## \* datasets

- Dat, [14](#)
  - DatA, [16](#)
  - DatB, [17](#)
  - DatMA, [18](#)
  - DatMB, [19](#)
  - DatPopM, [20](#)
  - Pik1A, [82](#)
  - Pik1B, [82](#)
- BKA, [2](#), [29](#)
- Ca1DF, [5](#), [32](#)
- Ca1SF, [8](#), [35](#)
- Compare, [11](#)
- CovHT, [13](#), [88](#)
- Dat, [14](#)
- DatA, [16](#), [82](#)
- DatB, [17](#), [82](#)
- DatMA, [18](#), [21](#)
- DatMB, [19](#), [21](#)
- DatPopM, [19](#), [20](#), [20](#)
- Domains, [21](#)
- FB, [22](#), [38](#)
- Hartley, [24](#), [24](#), [40](#)
- HT, [14](#), [26](#), [88](#)
- JackBKA, [4](#), [27](#)
- JackCa1DF, [7](#), [30](#)
- JackCa1SF, [11](#), [33](#)
- JackFB, [24](#), [36](#)
- JackHartley, [26](#), [38](#)
- JackMLCDF, [40](#), [66](#)
- JackMLCDW, [43](#), [69](#)
- JackMLCSW, [46](#), [71](#)
- JackMLDF, [49](#), [74](#)
- JackMLDW, [52](#), [76](#)
- JackMLSW, [54](#), [78](#)
- JackPEL, [57](#), [81](#)
- JackPML, [59](#)
- JackSFRR, [62](#), [87](#)
- MLCDF, [43](#), [64](#)
- MLCDW, [45](#), [67](#)
- MLCSW, [48](#), [69](#)
- MLDF, [51](#), [72](#)
- MLDW, [54](#), [74](#)
- MLSW, [56](#), [76](#)
- PEL, [59](#), [79](#)
- Pik1A, [17](#), [82](#)
- Pik1B, [18](#), [82](#)
- PML, [61](#), [83](#)
- SFRR, [64](#), [85](#)
- VarHT, [14](#), [27](#), [87](#)
- WeightsCa1DF, [89](#)
- WeightsCa1SF, [91](#)