# Package 'OBPSAE'

August 16, 2017

**Title** Observed best predictors for small area estimation

**Version** 1.0.0

**Description** Functions for OBP under Fay-Herriot model.

**Depends** R (>= 3.1.1)

**License** UCDavis

**LazyData** true

**RoxygenNote** 6.0.1

**Roxygen** list(markdown = TRUE)

**Author** Rohosen Bandyopadhyay [aut, cre]

**Maintainer** Rohosen Bandyopadhyay <rohosen.stat@gmail.com>

## R topics documented:

bpeFH *Best predictive estimator for Fay-Herriot model.*

## Description

This function computes the Best Predictive Estimators (BPE) of the unknown parameters for Fay-Herriot model. It computes BPE of the regression coefficients of the fixed effect (beta). The variance of the random effect can be specified by the user in which case that will be used to calculate the BPE of beta. Otherwise the function will calculate BPE of the variance component of the random effect and use that to calculate the BPE of beta.

## Usage

```
bpeFH(formula, data, errorvar, randvar = NULL, maxiter = 100,
  precision = 1e-04)
```

## Arguments

| | |
|---|---|
| formula | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| data | optional data frame containing the variable names in formula. |
| errorvar | vector containing the variances of the random errors for all small areas. |
| randvar | variance of the random effect. If not supplied, it is estimated by BPE. |
| maxiter | maximum number of iterations used in estimating randvar. |
| precision | covergence tolerance limit for estimating randvar. |

## Details

If randvar is not provided, it is first estimated by its BPE.

formula is specified in the form response ~ predictors where the predictors are separated by +. formula has an implied intercept term. To remove the intercept term, use either y ~ x - 1 or y ~ 0 + x.

## Value

The function will return a list with the following objects.

| | |
|---|---|
| A.BPE | BPE of variance of the random effect (if not specified by the user). |
| beta.BPE | BPE of the regression coefficient of the fixed effects. |

## References

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

`fitSpline.cv` *Fitting the optimal spline by cross validation method.*

## Description

This function fits an optimal spline function to the given data set after choosing optimal degree of the polynomial and optimal number of knots by cross-validation.

## Usage

```
fitSpline.cv(x, y)
```

## Arguments

| | |
|---|---|
| `x` | independent variable values. |
| `y` | dependent variable values |

## Details

The equation of a spline with degree $p$ and $q$ knots $\xi_1, \xi_2, \cdots, \xi_q$ is

$$y = \beta_0 + \sum_{k=1}^{p} \beta_k x^k + \sum_{l=1}^{q} \gamma_l (x - \xi_l)_+^p.$$

where $u_+ = uI(u > 0)$.

## Value

A list containing the following objects.

| | |
|---|---|
| `f.hat` | fitted values. |
| `p` | degree of the fitted polynomial. |
| `q` | number of knots. |
| `beta` | estimated coefficients of the polynomial. |
| `gamma` | estimated coefficients of the truncated polynomial. |
| `knots` | location of the knots. |

`Hospital` *Hospital data*

## Description

Morris and Christiansen (1995) presented a dataset involving 23 hospitals (out of a total of 219 hospitals) that had at least 50 kidney transplants during a 27-month period.

## Usage

```
Hospital
```

### Format

A data frame with 23 rows and 4 variables:

**Sample Size** Number of kidney transplants in each hospital during a 27 months period (at least 50)

**Graft Failure Rate** Graft failure rates for kidney transplant operations i.e. number of graft failures/sample size

**Severity Index** Severity index at each hospital which is measured by the average fraction of females, blacks, children, and extremely ill kidney recipients at that hospital

**SE** The standard error for the graft failure rate. The variance for the graft failure rate $D_i$, is approximated by $(0.2)(0.8)/n_i$, where 0.2 is the observed failure rate for all hospitals. Thus, $D_i$ is assumed known.

---

## MSPE_boot *Bootstrap MSPE Estimator of OBP.*

---

### Description

This function computes the bootstrap MSPE estimator of the Observed Best Predictor (OBP).

### Usage

```
MSPE_boot(theta.OBP, D, x, L = 200)
```

### Arguments

| | |
|---|---|
| theta.OBP | A vector of OBPs. |
| D | Scalar, Vector (or diagonal matrix) of random error variances. |
| x | Design matrix without intercept. |
| L | Number of bootstrap samples. the default is 200. |

### Value

This function will return a vector of the Bootstrap MSPE estimator of OBP.

### References

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

`MSPE_boot_adjusted` *Bootstrap MSPE estimator of Adjusted Observed Best Predictor (Adjusted OBP).*

## Description

This function computes the bootstrap MSPE estimator of benchmarked adjusted OBP.

## Usage

```
MSPE_boot_adjusted(theta.OBP.adjusted, D, x, weight, L = 200)
```

## Arguments

| | |
|---|---|
| `theta.OBP.adjusted` | A vector of adjusted OBPs. |
| `D` | Scalar, Vector (or diagonal matrix) of random error variances. |
| `x` | Design matrix without intercept term. |
| `weight` | Weight of each small area for calculating the benchmarked OBP (generally sampling weights of each small area). |
| `L` | is the number of bootstrap samples. the default is 200. |

## Value

This function will return a vector of the Bootstrap MSPE estimator of Adjusted OBP.

## References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

`MSPE_boot_augmented` *Bootstrap MSPE estimator of Augmented Observed Best Predictor (Augmented OBP).*

## Description

This function computes the bootstrap MSPE estimator of benchmarked augmented OBP.

## Usage

```
MSPE_boot_augmented(theta.OBP.augmented, D, x, weight, L = 200)
```

### Arguments

| | |
|---|---|
| theta.OBP.augmented | A vector of augmented OBPs. |
| D | Scalar, Vector (or diagonal matrix) of random error variances. |
| x | Design matrix without intercept term. |
| weight | Weight of each small area for calculating the benchmarked OBP (generally sampling weights of each small area). |
| L | Number of bootstrap samples. the default is 200. |

### Value

This function will return a vector of the Bootstrap MSPE estimator of Augmented OBP.

### References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

---

## MSPE_JNR *JNR MSPE Estimator of Observed Best Predictor.*

---

### Description

This function computes the MSPE estimator of Observed Best Predictor (OBP) proposed by Jiang, Nguyen and Rao (2011).

### Usage

```
MSPE_JNR(formula, data, errorvar, theta.OBP, A.BPE, beta.BPE)
```

### Arguments

| | |
|---|---|
| formula | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| data | optional data frame containing the variable names in formula. |
| errorvar | vector containing the variances of the random errors for all small areas. |
| theta.OBP | an optional vector of OBP values. See details. |
| A.BPE | optional BPE estimate of variance of random effects or the true value, if known. See details. |
| beta.BPE | optional BPE estimate of fixed effects coefficients. See details. |

## Details

formula is specified in the form response ~ predictors where the predictors are separated by +. formula has an implied intercept term. To remove the intercept term, use either y ~ x - 1 or y ~ 0 + x.

theta.OBP, A.BPE and beta.BPE are optional arguments. If any of them is missing, all three are computed from the data.

## Value

This function will return a vector of the JNR MSPE estimator of the OBP.

## References

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

---

# MSPE_McJack *McJack MSPE estimator*

---

## Description

This function computes the McJack MSPE estimator of the Observed Best Predictor (OBP).

## Usage

```
MSPE_McJack(formula, data, errorvar, A.BPE, K = 1000, returnMcSpline = TRUE)
```

## Arguments

| | |
|---|---|
| formula | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| data | optional data frame containing the variable names in formula. |
| errorvar | vector containing the variances of the random errors for all small areas. |
| A.BPE | optional BPE estimate of random effects variance. |
| K | number of Monte Carlo simulations. Default is 1000. |
| returnMcSpline | logical. Returns McSpline estimator with McJack (default). |

## Details

formula is specified in the form response ~ predictor where the predictor is univariate. formula has an implied intercept term. To remove the intercept term, use either y ~ x - 1 or y ~ 0 + x.

If A.BPE is missing, the function computes the BPE from data. User can also provide the true A instead if that is known.

## Value

The function will return a list with the following objects.

| | |
|---|---|
| McJack | McJack estimator of the MSPE of OBP. |
| McSpline | McSpline estimator of the MSPE of OBP. This is returned by default. To turn this feature off, set returnMcSpline = FALSE. |

## References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

---

# MSPE_McJack_Benchmark *McJack MSPE estimator of benchmarked OBP*

---

## Description

This function computes the McJack MSPE estimator of the benchmarked Observed Best Predictor.

## Usage

```
MSPE_McJack_Benchmark(formula, data, errorvar, weight, A.BPE, K = 1000,
  returnMcSpline = TRUE)
```

## Arguments

| | |
|---|---|
| formula | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| data | optional data frame containing the variable names in formula. |
| errorvar | vector containing the variances of the random errors for all small areas. |
| weight | vector containing the sampling weights of small areas. If sum of the weights is not 1, the weights are normalized. |
| A.BPE | optional BPE estimate of random effects variance. |
| K | number of Monte Carlo simulations. Default is 1000. |
| returnMcSpline | logical. Returns McSpline MSPE benchmark estimators with McJack (default). |

## Details

formula is specified in the form response ~ predictor where the predictor is univariate. formula has an implied intercept term. To remove the intercept term, use either y ~ x - 1 or y ~ 0 + x.

If A.BPE is missing, the function computes the BPE from data. User can also provide the true A instead if that is known.

## Value

The function will return a list with the following objects.

`McJack_Adj_bench`
: McJack MSPE estimator of the adjusted OBP.

`McJack_Aug_bench`
: McJack MSPE estimator of the augmented OBP.

`McSpline_Adj_bench`
: McSpline MSPE estimator of the adjusted OBP. This is returned by default. To turn this feature off, set `returnMcSpline = FALSE`

`McSpline_Aug_bench`
: McSpline MSPE estimator of the augmented OBP. This is returned by default. To turn this feature off, set `returnMcSpline = FALSE`

## References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

---

# `MSPE_McSpline` *McSpline MSPE estimator*

---

## Description

This function computes the McSPline MSPE estimator of the Observed Best Predictor (OBP).

## Usage

```
MSPE_McSpline(formula, data, errorvar, A.BPE, K = 1000)
```

## Arguments

| | |
|---|---|
| `formula` | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| `data` | optional data frame containing the variable names in `formula`. |
| `errorvar` | vector containing the variances of the random errors for all small areas. |
| `A.BPE` | optional BPE estimate of random effects variance. |
| `K` | number of Monte Carlo simulations. Default is 1000. |

## Details

`formula` is specified in the form `response ~ predictor` where the predictor is univariate. `formula` has an implied intercept term. To remove the intercept term, use either `y ~ x - 1` or `y ~ 0 + x`.

If `A.BPE` is missing, the function computes the BPE from data. User can also provide the true A instead if that is known.

## Value

The function will return a list with the following object.

| | |
|---|---|
| McSpline | McSpline estimator of the MSPE of OBP. |

## References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

---

# MSPE_McSpline_Benchmark

*McSpline MSPE estimator of benchmarked OBP*

---

## Description

This function computes the McSpline MSPE estimator of the benchmarked Observed Best Predictor.

## Usage

```
MSPE_McSpline_Benchmark(formula, data, errorvar, weight, A.BPE, K = 1000)
```

## Arguments

| | |
|---|---|
| formula | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| data | optional data frame containing the variable names in formula. |
| errorvar | vector containing the variances of the random errors for all small areas. |
| weight | vector containing the sampling weights of small areas. If sum of the weights is not 1, the weights are normalized. |
| A.BPE | optional BPE estimate of random effects variance. |
| K | number of Monte Carlo simulations. Default is 1000. |

## Details

formula is specified in the form response ~ predictor where the predictor is univariate. formula has an implied intercept term. To remove the intercept term, use either y ~ x - 1 or y ~ 0 + x.

If A.BPE is missing, the function computes its BPE from data. User can also provide the true A instead if that is known.

## Value

The function will return a list with the following objects.

`McSpline_Adj_bench`
: McSpline MSPE estimator of the adjusted OBP.

`McSpline_Aug_bench`
: McSpline MSPE estimator of the augmented OBP.

## References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

---

`MSPE_naive` *Naive MSPE Estimator of Observed Best Predictor (OBP).*

---

## Description

This function computes the naive MSPE estimator of OBP.

## Usage

```
MSPE_naive(y, D, theta.OBP, A.BPE)
```

## Arguments

| | |
|---|---|
| `y` | A vector of the response variable. |
| `D` | Scalar, Vector (or diagonal matrix) of random error variances. |
| `theta.OBP` | A vector of OBP values |
| `A.BPE` | BPE estimate of variance of random effects or the true value, if known. |

## Value

This function will return the Naive MSPE estimator of the OBP.

## References

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

obpFH *Observed best predictor for Fay-Herriot model.*

## Description

This function computes the Observed Best Predictor (OBP) for Fay-Herriot model. The variance of the random error can be specified by the user. Otherwise the function will calculate its Best Predictive Estimator (BPE). In the process of of computing OBP it also calculates the BPE of the regression coefficients of the fixed effect.

## Usage

```
obpFH(formula, data, errorvar, randvar = NULL, maxiter = 100,
  precision = 1e-04)
```

## Arguments

| | |
|---|---|
| formula | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| data | optional data frame containing the variable names in formula. |
| errorvar | vector containing the variances of the random errors for all small areas. |
| randvar | variance of the random effect. If not supplied, it is estimated by BPE. |
| maxiter | maximum number of iterations used in estimating randvar. |
| precision | covergence tolerance limit for estimating randvar. |

## Details

If randvar is not provided, it is first estimated by its BPE.

formula is specified in the form response ~ predictors where the predictors are separated by +. formula has an implied intercept term. To remove the intercept term, use either y ~ x - 1 or y ~ 0 + x.

## Value

The function will return a list with the following objects

| | |
|---|---|
| theta.OBP | OBP of the small area mean. |
| A.BPE | BPE of variance of the random effect (if not specified by the user). |
| beta.BPE | BPE of the regression coefficients of the fixed effect. |

## References

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

obpFHbenchmark *Benchmarked observed best predictor for Fay-Herriot model.*

## Description

This function computes the benchmarked Observed Best Predictor (OBP) for Fay-Herriot model. Depending on the method specified by the user it computes the Adjusted OBP or Augmented OBP or both.

## Usage

```
obpFHbenchmark(formula, data, errorvar, weight, method = c("adjusted",
  "augmented"), randvar = NULL, maxiter = 100, precision = 1e-04)
```

## Arguments

| | |
|---|---|
| formula | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| data | optional data frame containing the variable names in formula. |
| errorvar | vector containing the variances of the random errors for all small areas. |
| weight | vector containing the sampling weights of small areas. If sum of the weights is not 1, the weights are normalized. |
| method | string specifying the benchmarking method. Options are "adjusted" and "augmented". Computes both if not specified. See Details for more usage information. |
| randvar | variance of the random effect. If not supplied, BPE is estimated. |
| maxiter | maximum number of iterations used in estimating randvar. |
| precision | covergence tolerance limit for estimating randvar. |

## Details

If method is set to "adjusted", only obpAdjusted is returned.

If method is set to "augmented", obpAugmented, A.BPE.aug and beta.BPE.aug are returned.

The variance of the random effect can be specified by the user. Otherwise the function will calculate its Best Predictive Estimator (BPE). In the process of of computing OBP it also calculates the BPE of the regression coefficients of the fixed effect.

formula is specified in the form response ~ predictors where the predictors are separated by +. formula has an implied intercept term. To remove the intercept term, use either y ~ x - 1 or y ~ 0 + x.

## Value

The function will return a list with all the following objects by default.

| | |
|---|---|
| obpAdjusted | a vector of adjusted OBP values. |
| obpAugmented | a vector of augmented OBP values. |
| A.BPE.aug | BPE of variance component of random effects under the augmented model (if not provided by the user). |
| beta.BPE.aug | BPE of fixed effects regression coefficients under the augmented model. |

## References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

---

# obpFH_adjusted *Adjusted observed best predictor for Fay-Herriot Model.*

---

## Description

This function computes the Adjusted Observed Best Predictor (OBP) for Fay-Herriot model. The variance of the random error can be specified by the user. Otherwise the function will calculate its Best Predictive Estimator (BPE). In the process of of computing Adjusted OBP it also calculates the BPE of the regression coefficients of the fixed effect.

## Usage

```
obpFH_adjusted(formula, data, errorvar, weight, randvar = NULL,
  maxiter = 100, precision = 1e-04)
```

## Arguments

| | |
|---|---|
| formula | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| data | data frame containing the variable names in formula and errorvar. |
| errorvar | vector containing the variances of the random errors for all small areas. |
| weight | vector containing the sampling weights of small areas. If sum of the weights is not 1, the weights are normalized. |
| randvar | variance of the random effect. If not supplied, BPE is estimated. |
| maxiter | maximum number of iterations used in estimating randvar. |
| precision | covergence tolerance limit for estimating randvar. |

## Details

The variance of the random effect can be specified by the user. Otherwise the function will calculate its Best Predictive Estimator (BPE). In the process of of computing Adjusted OBP it also calculates the BPE of the regression coefficients of the fixed effect.

`formula` is specified in the form `response ~ predictors` where the predictors are separated by `+`. `formula` has an implied intercept term. To remove the intercept term, use either `y ~ x - 1` or `y ~ 0 + x`.

## Value

The function will return a list containing the Adjusted OBP as follows:

`obpAdjusted` a vector of adjusted OBP values.

## References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

---

`obpFH_augmented` *Augmented observed best predictor for Fay-Herriot model.*

---

## Description

This function computes the Augmented observed best predictor (OBP) for Fay-Herriot model. The variance of the random error can be specified by the user. Otherwise the function will calculate its Best Predictive Estimator (BPE) under the augmented model. In the process of of computing augmented OBP it also calculates the BPE of the regression coefficients of the fixed effect.

## Usage

```
obpFH_augmented(formula, data, errorvar, weight, randvar = NULL,
  maxiter = 100, precision = 1e-04)
```

## Arguments

| | |
|---|---|
| `formula` | an object of class formula (or one that can be coerced to that class): a symbolic description of the model to be fitted. The variables included in formula must have a length equal to the number of small areas. More about the model specification are given under Details. |
| `data` | data frame containing the variable names in formula and errorvar. |
| `errorvar` | vector containing the variances of the random errors for all small areas. |
| `weight` | vector containing the sampling weights of small areas. If sum of the weights is not 1, the weights are normalized. |
| `randvar` | variance of the random effect. If not supplied, BPE is estimated. |
| `maxiter` | maximum number of iterations used in estimating randvar. |
| `precision` | covergence tolerance limit for estimating randvar. |

## Details

The variance of the random effect can be specified by the user. Otherwise the function will calculate its Best Predictive Estimator (BPE) under the augmented model. In the process of of computing Adjusted OBP it also calculates the BPE of the regression coefficients of the fixed effect.

`formula` is specified in the form `response ~ predictors` where the predictors are separated by `+`. `formula` has an implied intercept term. To remove the intercept term, use either `y ~ x - 1` or `y ~ 0 + x`.

## Value

The function will return a list with all the following objects by default.

| | |
|---|---|
| `obpAugmented` | a vector of augmented OBP values. |
| `A.BPE.aug` | BPE of variance component of random effects under the augmented model (if not provided by the user). |
| `beta.BPE.aug` | BPE of fixed effects regression coefficients under the augmented model. |

## References

Bandyopadhyay R, Jiang J (2017) "Benchmarking the Observed Best Predictor"

Jiang J, Nguyen T, and Rao J. S. (2011), "Best Predictive Small Area Estimation", Journal of the American Statistical Association.

# Index