Package ‘ERP’

December 14, 2015

Type Package

Title Significance Analysis of Event-Related Potentials Data

Version 1.1

Date 2015-12-11

Author David Causeur (Agrocampus, Rennes, France) and Ching-Fan Sheu (National Cheng-Kung University, Tainan, Taiwan)

Maintainer David Causeur <david.causeur@agrocampus-ouest.fr>

Description The functions provided in the package ERP are designed for the significance analysis of ERP data in a linear model framework. The possible procedures are either the collection of FDR or FWER controlling methods available in the generic function p.adjust, the same collection combined with a factor modeling of the time dependence among tests (see Sheu, Perthame, Lee and Causeur, 2016) and the Guthrie and Buchwald (1991) test.

License GPL (>= 2)

Imports fdrtool,mnormt

NeedsCompilation no

Repository CRAN

Date/Publication 2015-12-14 17:19:08

R topics documented:

   ERP-package ..... 2
   erpavetest ..... 4
   erpcz ..... 6
   erpfatest ..... 7
   erpplot ..... 10
   erptest ..... 11
   gbtest ..... 13
   simerp ..... 15

Index 17
**Description**

The package provides multiple testing procedures designed for Event-Related Potentials (ERP) data in a linear model framework. These procedures are reviewed and compared in Sheu, Perthame, Lee and Causeur (2016). Some of the methods gathered in the package are the classical FDR- and FWER-controlling procedures, also available using function p.adjust. The package also implements the Guthrie-Buchwald procedure (Guthrie and Buchwald, 1991), which accounts for the auto-correlation among t-tests to control erroneous detections of short intervals. The Adaptive Factor-Adjustment method is an extension of the method described in Causeur, Chu, Hsieh and Sheu (2012). It assumes a factor model for the correlation among tests and combines adaptively the estimation of the signal and the updating of the dependence modelling (see Sheu et al., 2016 for further details).

**Details**

<table>
<thead>
<tr>
<th>Package:</th>
<th>ERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>Version:</td>
<td>1.0</td>
</tr>
<tr>
<td>Date:</td>
<td>2014-04-01</td>
</tr>
<tr>
<td>License:</td>
<td>GPL (&gt;=2)</td>
</tr>
</tbody>
</table>

Apart from the function erpplot, which is just a wrapper for matplot to display the ERP curves, all the functions in the package implement multiple testing procedures of ERP data in a linear model framework (F-tests for the comparison of two nested models). The function gbtest implements the Guthrie-Buchwald procedure (see Guthrie and Buchwald, 1991). The function erptest can be used for the classical FDR- and FWER-controlling multiple testing of ERP data: especially the Benjamini-Hochberg (see Benjamini and Hochberg, 1995) and Benjamini-Yekutieli (see Benjamini and Yekutieli, 2001) procedures, with the possible extension proposed by Storey and Tibshirani (2003) including a non-parametric estimation of the proportion of true nulls. The function erpavetest first partition of the entire interval of ERP observations into a predetermined number equally-spaced intervals before performing significance testing using the averaged ERPs.

**Author(s)**

David Causeur, (Agrocampus, Rennes, France), Ching-Fan Sheu (National Cheng Kung University, Tainan, Taiwan), and Mei-Chen Chu (National Cheng Kung University, Tainan, Taiwan)

References


Examples

```r
### simualted ERPs for 20 subjects (rows) and 251 ERPs measured at
### every 4 milliseconds plus a behavior score (columns)

data(simerp)

### Plot raw ERP curves

erpplot(simerp[,1:251],frames=seq(0,1001,4),xlab="Time (ms)",
        ylab=expression(ERP),col="black",main="Simulated ERP")

### Test of averaged ERPs over a pre-determined number of equal intervals

frames = seq(0,1001,4)

tests = erpavetest(simerp[,1:251],design=model.matrix(~y,data=simerp))

plot(frmes,sign(tests$signal)*sqrt(tests$r2),type="l",
```

Significance testing of averaged ERPs. The entire ERP recording time is first partitioned into a pre-determined number of equal intervals. The averaged ERPs for each time intervals are the input for analysis.

Description

The function first calculates averaged ERP values within a predetermined number of equally-spaced intervals then tests for significance of the relationship between averaged ERPs and covariates in a linear model framework.

Usage

```
erpavetest(dta, design, design0 = NULL, nintervals = 10, method = "none", alpha = 0.05)
```
Arguments

dta Data frame containing the ERP curves: each column corresponds to a time frame and each row to a curve.
design Design matrix of the full model for the relationship between the ERP and the experimental variables. Typically the output of the function model.matrix
design0 Design matrix of the null model. Typically a submodel of the full model, obtained by removing columns from design. Default is NULL, corresponding to the model with no covariates.
nintervals Number of intervals in the partition of the whole interval of observation. Default is 10.
method FDR- or FWER- controlling multiple testing procedures as available in the function p.adjust. Default is "none".
alpha The FDR or FWER control level. Default is 0.05

Value

pval p-values of the tests.
correctedpval Corrected p-values, for the multiplicity of tests. Depends on the multiple testing method (see function p.adjust).
significant Indices of the time points for which the test is positive.
segments Factor giving the membership of timepoints to each interval in the partition.
breaks Breakpoints of the partition.
test F-statistics.
df1 Residual degrees of freedom for the full model.
df0 Residual degrees of freedom for the null model.
signal Estimated signal: a pxT matrix, where p is the difference between the numbers of parameters in the full and null models and T the number of frames.
r2 R-squared values for each of the T linear models.

Author(s)

David Causeur - david.causeur@agrocampus-ouest.fr

See Also

erptest, erpfatest, gbtest, p.adjust, pval.estimate.eta0

Examples

require(mnormt)
require(fdrtool)
data(erpcz)
data(simerp)
# Paired t-tests for the comparison of ERP curves between two groups

tests = erpavetest(erpacz[,1:251], design=model.matrix(~Subject+Instruction, data=erpacz),
  design0=model.matrix(~Subject, data=erpacz))

frames = seq(0,1000,4)
plot(frames, tests$signal, type="l", xlab="Time (ms)", ylab="Difference ERP curves")
points(frames[tests$significant], rep(0, length(tests$significant)), pch=16, col="blue")
abline(v=frames[tests$breaks], lty=2, col="darkgray")
title("Paired comparison at electrode CZ")

# Tests for significance of correlations

tests = erpavetest(simerp[,1:251], design=model.matrix(~y, data=simerp))

plot(frames, sign(tests$signal)*sqrt(tests$r2), type="l",
xlab="Time (ms)", ylab="Correlation", ylim=c(-1,1))
points(frames[tests$significant], rep(-1, length(tests$significant)),
pch=16, col="blue")
abline(v=frames[tests$breaks], lty=2, col="darkgray")
title("Simulation")

---

erpacz

**ERP dataset from a directed forgetting experiment as presented in detail in Sheu, Perthame, Lee and Causeur (2016)**

---

**Description**

The experiment was conducted using an experimental paradigm similar to that described in Lee, Lee and Fawcett (2013). The objective of the study was to examine the electrophysiological process associated with directed forgetting.

**Usage**

data(erpacz)

**Format**

40 ERP curves obtained at electrode CZ for 20 subjects in two experimental conditions (TBR and TBF), each curve lasting for one thousand milliseconds (ms) with one record per 4ms. The last two columns, Subject and Instruction, provide the subject identifier and the group membership respectively.

**Details**

In the study phase of the experiment, 20 participants were instructed (with a ‘+’ or ‘X’ cue, respectively) to either remember or forget a stimulus word that was presented on a computer screen. ERPs were recorded throughout each of 90 trials - half for to-be-remembered (TBR) and half for
to-be-forgotten (TBF). ERPs from nine electrode positions - 3 each from frontal, central and posterior regions - during the study phase were analyzed. The current dataset provides the ERP curves at electrode CZ (central on the mline). The corresponding experimental covariates (subject and instruction) are provided in the last two columns.

Source

References

Examples
data(erpCZ)
erpplot(erpCZ[,1:251], frames=seq(0,1001,4), xlab="Time (ms)", ylab="ERP", main="ERP at electrode CZ")

derpfatest  Adaptive Factor-Adjustment for multiple testing of ERP data

Description
An adaptive factor-adjusted FDR- and FWER-controlling multiple testing procedures for ERP data. The procedure is described in detail in Sheu, Perthame, Lee, and Causeur (2016).

Usage
derpfatest(dta, design, design0 = NULL, method = "BH", nbf = NULL, nbfmax = 15, alpha = 0.05, pi0 = 1, wantplot = FALSE, s0 = NULL, min.err = 0.001, maxiter = 5, verbose = FALSE)

Arguments
dta Data frame containing the ERP measurements: each column corresponds to a time frame and each row to a curve.
design Design matrix of the full model for the relationship between the ERP and the experimental variables. Typically the output of the function model.matrix
design0 Design matrix of the null model. Typically a submodel of the full model, obtained by removing columns from design. Default is NULL, corresponding to the model with no covariates.
### Description

The method is described in Sheu et al. (2016). It combines a decorrelation step based on a regression factor model as in Leek and Storey (2008), Friguet et al. (2009) or Sun et al. (2012) with an adaptive estimation of the ERP signal. The multiple testing corrections of the p-values are described in the help documentation of the function `p.adjust` of package stats.

### Value

- **pval**: p-values of the Adaptive Factor-Adjusted tests.
- **correctedpval**: Corrected p-values, for the multiplicity of tests. Depends on the multiple testing method (see function `p.adjust`).
- **significant**: Indices of the time points for which the test is positive.
- **pi0**: Value for pi0: if the input argument pi0 is NULL, the output is the estimated proportion of true null hypotheses using the method by Storey and Tibshirani (2003).
- **test**: Factor-Adjusted F-statistics.

### Arguments

- **method**: FDR- or FWER- controlling multiple testing procedures as available in the function `p.adjust`. Default is "BH", for the Benjamini-Hochberg procedure (see Benjamini and Hochberg, 1995).
- **nbf**: Number of factors in the residual covariance model. Default is NULL: the number of factors is determined by minimization of the variance inflation criterion as suggested in Friguet et al. (2009).
- **nbfmax**: Only required if nbf=NULL. The largest possible number of factors.
- **alpha**: The FDR or FWER control level. Default is 0.05
- **pi0**: An estimate of the proportion of true null hypotheses, which can be plugged into an FDR controlling multiple testing procedure to improve its efficiency. Default is 1, corresponding to the classical FDR controlling procedures. If NULL, the proportion is estimated using the function `pval.estimate.eta0` of package `fdrtool`, with the method proposed by Storey and Tibshirani (2003).
- **wantplot**: If TRUE, a diagnostic plot is produced to help choosing the number of factors. Only active if nbf=NULL.
- **s0**: Prior knowledge of the time frames for which no signal is expected. For example, `s0=c(1:50, 226:251)` specifies that the first 50 time frames and time frames between 226 and 251 are known not to contain ERP signal. `s0` can also be specified by giving the lower and upper fraction of the entire time interval in which the signal is to be searched for. For example: `s0=c(0.2, 0.9)` means that ERP signals are not expected for for the first 20 percent and last 10 percent of the time frames measured. Default is NULL and it initiates a data-driven determination of `s0`.
- **min.err**: Control parameter for convergence of the iterative algorithm. Default is 1e-03.
- **maxiter**: Maximum number of iterations in algorithms. Default is 5.
- **verbose**: If TRUE, details are printed along the iterations of the algorithm. Default is FALSE.
df1  Residual degrees of freedom for the full model.
df0  Residual degrees of freedom for the null model.
nbf  Number of factors for the residual covariance model.
signal  Estimated signal: a pxT matrix, where p is the difference between the numbers of parameters in the full and null models and T the number of frames.
r2  R-squared values for each of the T linear models.

Author(s)
David Causeur - david.causeur@agrocampus-ouest.fr

References

See Also
erpavetest, erptest, gptest, p.adjust, pval.estimate.eta0

Examples
```r
require(mnormt)
require(fdrtool)

data(erpacz)
data(simerp)

# Paired t-tests for the comparison of ERP curves between two groups

tests = erpfatest(erpacz[,1:251],design=model.matrix(~Subject+Instruction,data=erpacz),
  design0=model.matrix(~Subject,data=erpacz),nbf=3,s=0=c(1:50,226:251),min.err=1e-01)

frames = seq(0,1001,4)
plot(frames,tests$signal,type="l",xlab="Time (ms)",ylab="Difference ERP curves")
points(frames[tests$significant],rep(0,length(tests$significant)),pch=16,col="blue")
```
title("Paired comparison at electrode CZ")

# Tests for significance of correlations

tests = erpfatest(simerp[,1:251],design=model.matrix(~y,data=simerp),
                 nbf=5,min.err=1e-01,maxiter=10)
plot(frames,sign(tests$signal)*sqrt(tests$r2),type="l",
     xlab="Time (ms)",ylab="Correlation",ylim=c(-1,1))
points(frames[tests$significant],rep(-1,length(tests$significant)),
       pch=16,col="blue")
title("Simulation")

erpplot  

**Plot of ERP curves.**

**Description**

Wrapper for matplot (package graphics) to display ERP curves.

**Usage**

```r
erpplot(dta, frames = NULL, ylim = NULL, ...)
```

**Arguments**

- **dta**
  Data frame containing the ERP curves: each column corresponds to a time frame and each row to a curve.

- **frames**
  Sequence of time frames. Default is NULL, in which case frames is just the sequence of integers between one and the total number of frames.

- **ylim**
  Limits for the y-axis. Default is NULL (set into the function).

- **...**
  Graphical parameters (see par) and any further arguments of plot, typically plot.default, may also be supplied as arguments to this function. Hence, the high-level graphics control arguments described under par and the arguments to title may be supplied to this function.

**Value**

The function generates a plot, but does not return any other numerical outputs.

**Author(s)**

David Causeur (Agrocampus, Rennes, France), Ching-Fan Sheu (National Cheng-Kung University, Tainan, Taiwan) and Mei-Chen Chu (National Cheng-Kung University, Tainan, Taiwan)

**See Also**

- matplot
Erptest

Examples

data(erpCz)
erpplot(erpCz[,1:251], frames = seq(0,1001,4), xlab = "Time (ms)",
ylab = "ERP", main = "ERP at electrode CZ")

data(simERP)
erpplot(simERP[,1:251], frames = seq(0,1001,4), xlab = "Time (ms)",
ylab = "ERP", main = "Simulated ERP")

Erptest  FDR- and FWER-controlling Multiple testing of ERP data

Description

Classical FDR- and FWER-controlling multiple testing procedures for ERP data in a linear model framework.

Usage

erptest(dta, design, design0 = NULL, method = "BH", alpha = 0.05, pi0 = 1)

Arguments

dta  Data frame containing the ERP curves: each column corresponds to a time frame and each row to a curve.
design  Design matrix of the full model for the relationship between the ERP and the experimental variables. Typically the output of the function model.matrix
design0  Design matrix of the null model. Typically a submodel of the full model, obtained by removing columns from design. Default is NULL, corresponding to the model with no covariates.
method  FDR- or FWER- controlling multiple testing procedures as available in the function p.adjust. Default is "BH", for the Benjamini-Hochberg procedure (see Benjamini and Hochberg, 1995).
alpha  The FDR or FWER control level. Default is 0.05
pi0  An estimate of the proportion of true null hypotheses, which can be plugged into an FDR controlling multiple testing procedure to improve its efficiency. Default is 1, corresponding to the classical FDR controlling procedures. If NULL, the proportion is estimated using the function pval.estimate.eta0 of package fdrtool, with the method proposed by Storey and Tibshirani (2003).

Details

The multiple testing corrections of the p-values are described in the help documentation of the function p.adjust du package stats.
Value

pval  
p-values of the tests.
correctedpval  
Corrected p-values, for the multiplicity of tests. Depends on the multiple testing method (see function p.adjust).
significant  
Indices of the time points for which the test is positive.
pi0  
Value for pi0: if the input argument pi0 is NULL, the output is the estimated proportion of true null hypotheses using the method by Storey and Tibshirani (2003).
test  
F-statistics.
df1  
Residual degrees of freedom for the full model.
df0  
Residual degrees of freedom for the null model.
signal  
Estimated signal: a pxT matrix, where p is the difference between the numbers of parameters in the full and null models and T the number of frames.
r2  
R-squared values for each of the T linear models.

Author(s)

David Causeur - david.causeur@agrocampus-ouest.fr

References


See Also

erpavetest, erpfatest, gbtest, p.adjust, pval.estimate.eta0
Examples

```r
require(mnormt)
require(fdrtool)
data(erpcz)
data(simerp)

# Paired comparison of ERP curves

tests = erptest(erpcz[,1:251], design=model.matrix(~Subject+Instruction, data=erpcz),
                 design0=model.matrix(~Subject, data=erpcz))
frames = seq(0,1001,4)
plot(frames, tests$signal, type="l", xlab="Time (ms)",
     ylab="Difference ERP curves")
points(frames[tests$significant], rep(0, length(tests$significant)),
       pch=16, col="blue")
title("Paired comparison at electrode CZ")

# Independent two-group comparison of ERP curves

tests = erptest(erpcz[,1:251], design=model.matrix(~Instruction, data=erpcz))
frames = seq(0,1001,4)
plot(frames, tests$signal, type="l", xlab="Time (ms)",
     ylab="Difference ERP curves")
points(frames[tests$significant], rep(0, length(tests$significant)),
       pch=16, col="blue")
title("Independent comparison at electrode CZ")

# Tests for significance of correlations

tests = erptest(simerp[,1:251], design=model.matrix(~y, data=simerp))
plot(frames, sign(tests$signal)*sqrt(tests$r2), type="l", xlab="Time (ms)",
     ylab="Correlation", ylim=c(-1,1))
points(frames[tests$significant], rep(-1, length(tests$significant)),
       pch=16, col="blue")
title("Simulation")
```

---

gbtest  
The Guthrie-Buchwald procedure for significance analysis of ERP data

Description

Monte-Carlo implementation of the Guthrie-Buchwald procedure (see Guthrie and Buchwald, 1991) which accounts for the auto-correlation among test statistics to control erroneous detections of short intervals.
Usage

gbtest(dta, design, design0 = NULL, graphthresh = 0.05, nbsamples = 1000)

Arguments

da
Data frame containing the ERP curves: each column corresponds to a time frame and each row to a curve.
design
Design matrix of the full model for the relationship between the ERP and the experimental variables. Typically the output of the function model.matrix
design0
Design matrix of the null model. Typically a submodel of the full model, obtained by removing columns from design. Default is NULL, corresponding to the model with no covariates.
graphthresh
Graphical threshold (see Guthrie and Buchwald, 1991). Default is 0.05. As the FDR control level, the smaller is the graphical threshold, the more conservative is the procedure.
nbsamples
Number of samples in the Monte-Carlo method to estimate the residual covariance. Default is 1000.

Details

The Guthrie-Buchwald method starts from a preliminary estimation of r, the lag-1 autocorrelation, among test statistics. Then, the null distribution of the lengths of the intervals $I_{alpha} = t : pvalue_t <= alpha$, where alpha is the so-called graphical threshold parameter of the method, is obtained using simulations of p-values $p_t$ associated to auto-regressive t-test process of order 1 with mean 0 and auto-correlation r. Such an interval $I_{alpha}$ is declared significant if its length exceeds the $(1-alpha)$-quantile of the null distribution. Note that the former method is designed to control erroneous detections of short significant intervals but not to control any type-I error rate.

Value

nbsignifintervals
Number of significant intervals.
intervals
List of length nbsignifintervals which components give the indices of each significant intervals.
significant
Indices of the time points for which the test is positive.
signal
Estimated signal: a pxT matrix, where p is the difference between the numbers of parameters in the full and null models and T the number of frames.
rho
Estimated lag-1 auto-correlation.
r2
R-squared values for each of the T linear models.

Author(s)

David Causeur - david.causeur@agrocampus-ouest.fr and Mei-Chen Chu (National Cheng-Kung University, Tainan, Taiwan)
simerp

References


See Also

erpavetest, erpfatest, erptest

Examples

require(mnormt)
data(erpcz)
data(simerp)

# Paired comparison of ERP curves
tests = gbtest(erpcz[,1:251],design=model.matrix(~Subject+Instruction,data=erpcz),
   design0=model.matrix(~Subject,data=erpcz),nbsamples=500)
frames = seq(0,1000,4)
plot(frames,tests$signal,type="l",xlab="Time (ms)",
   ylab="Difference ERP curves")
points(frames[tests$significant],rep(0,length(tests$significant)),
   pch=16,col="blue")
title("Paired comparison at electrode CZ")

# Tests for significance of correlations
tests = gbtest(simerp[,1:251],design=model.matrix(~y,data=simerp),nbsamples=500)
plot(frames,sign(tests$signal)*sqrt(tests$r2),type="l",xlab="Time (ms)",
   ylab="Correlation",ylim=c(-1,1))
points(frames[tests$significant],rep(-1,length(tests$significant)),
   pch=16,col="blue")
title("Simulation")

simerp

ERP dataset from a simulation study presented in detail in Sheu, Perthame, Lee and Causeur (2016)

Description

The dataset contains 20 ERP curves (time frames every 4 ms from 0 to 1000 ms) and a variable y for behavioral score.
Usage

data(simerp)

Format

A sample of 20 simulated ERP curves, each curve lasting for one thousand milliseconds (ms) with one record per 4ms. The last column gives a behavioral score for the directed forgetting experiment described in Sheu, Perthame, Lee and Causeur (2016).

Details

The dataset is simulated under assumption of a multivariate regression model with normal residuals. The covariate, y, is a recognition score, as observed in a directed forgetting experiment described in Sheu, Perthame, Lee and Causeur (2016). Only the slope coefficients corresponding to time points in the interval [450 ms ; 550 ms] are true non-nulls, with a peak at 500 ms. The residual covariance of the model mimics the sample residual covariance of the ERP datasets collected in the directed forgetting experiment.

Source


Examples

data(simerp)
erpplot(simerp[,1:251],frames=seq(0,1001,4),xlab="Time (ms)",
         ylab="ERP",main="Simulated ERP")
Index

*Topic ERP data
  erpavetest, 4
  erpfatetst, 7
  erptest, 11

*Topic ERP
  ERP-package, 2
  gbttest, 13

*Topic FDR
  erpavetest, 4
  erpfatetst, 7
  erptest, 11

*Topic Factor-adjustment
  erpfatetst, 7

*Topic Guthrie-Buchwald procedure
  gbttest, 13

*Topic Multiple testing
  erpavetest, 4
  erpfatetst, 7
  erptest, 11
  gbttest, 13

*Topic datasets
  erpcz, 6
  simerp, 15

*Topic package
  ERP-package, 2

*Topic significance analysis
  ERP-package, 2

ERP (ERP-package), 2
ERP-package, 2
erpavetest, 4, 9, 12, 15
erpcz, 6
erpfatetest, 5, 7, 12, 15
erpplot, 10
erptest, 5, 9, 11, 15

gbttest, 5, 9, 12, 13
matplot, 10
p.adjust, 5, 9, 12

pval.estimate.eta0, 5, 9, 12
simerp, 15