Package 'BSPM'

April 14, 2020

Version 0.1.0

Title Bayesian Statistical Process Monitoring

Author Dimitrios Kiagias [aut, cre, cph],
Konstantinos Bourazas [aut, cph],
Panayiotis Tsiamyrtzis [aut, cph]

Maintainer Dimitrios Kiagias <d.kiagias@sheffield.ac.uk>

Description Predictive Control Charts (PCC) for
members of the regular exponential family.

License GPL (>=2)

Depends R (>= 3.5.0)

Imports plotrix,
grDevices,
stats,
ggplot2,
grid,
gridExtra

LazyData true

R topics documented:

---

aPTT *Dataset for PCC process for Normal with both parameters unknown*

Description

The activated Partial Thromboplastin Time (aPTT) dataset is a data frame that contains two columns. The first column (aPTT_current) repsesents 30 sequentially gathered current observations from the daily routine of a medical laboratory. The values refer to the activated Partial Thromboplastin Time (aPTT), measured in seconds. The second column (aPTT_historical) refers to 30 historical In Control measurements of the same variable. APTT is a blood test that characterizes coagulation of the blood. More specifically, it is the time that until the coagulation when reagents are added to plasma.

Usage

```
data("aPTT")
```

Examples

```
# Loading data
attach(aPTT)

# Plotting data
plot(1:length(aPTT_current), aPTT_current, type = "b", lty = 1, xlab = "", ylab = "",
     ylim = c(27.3, 33.4), xlim = c(-30, 30), lwd = 1, pch = 16, axes = FALSE, yaxs = "i",
      main = "aPTT dataset")

# x - axis for historical and current data
pastx <- c(-30, -20, -10, 0) ; currentx <- c(0, 10, 20, 30)
# Adding points
points(-length(aPTT_historical):(-1), aPTT_historical,
       type = "b", lty = 2, xlab = "", ylab = "", lwd = 1, pch = 21, col = "gray55")
# Adding axis with names
axis(2)
mtext("Current Data", side = 1, at = 15, line = 2.2, cex = 1.1)
mtext("Historical Data", side = 1, at = -15, line = 2.2, cex = 1.1, col = "gray55")
mtext("aPTT [sec]", side = 2, line = 2.2, cex = 1.1)
axis(1, at = currentx, labels = currentx)
axis(1, at = pastx, labels = pastx, col.axis = "gray55", col = "gray55", lty = 2)
segments(0, 27.5, 0, 33.5, lwd = 1)
```

ECE *ECE dataset for the PCC process for Poisson with rate parameter unknown*

Description

The Electrical Complex Equipment (ECE) dataset repsesents 25 paired observations in a data frame. The first column (defect_counts) refers the number of defects per inspected number of units, encountered in a complex electrical equipment of an assembly line. The second column (inspected_units) contains the corresponding inspected number of units for the counting of defects. The data were also analyzed by Hansen and Ghare (1987) and Bayarri and Garcia-Donato (2005).

Usage

```
data("ECE")
```

References

[1] Bayarri, M. J., and Garcia-Donato, G. (2005), "A Bayesian Sequential Look at u-Control Charts", Technometrics, 47, 2, pp. 142-151

[2] Hansen, B., and Ghare, P. (1987), "Quality Control and Application, Prentice-Hall", Englewood Cliffs, NJ

Examples

```
# Loading data
attach(ECE)

# Plotting data
par( oma = c(1,3,2,3) )
plot(1:length(defect_counts), defect_counts/inspected_units, type = "b", lty = 1,
     xlab = "Observations", ylab = "", xlim = c(0, 25), ylim = c(1.5, 10.5),
     lwd = 1, pch = 16, axes = FALSE, yaxs = "i", main = "ECE dataset")

# Adding points
points(1:length(defect_counts), inspected_units, type = "b", lty = 2, lwd = 1, pch = 21, col = "gray55")
# Adding legend
legend("topleft", legend=c(expression(paste(s[i])), expression(paste(x[i]/s[i])) ), bty = "n",
        cex = 0.8, lty = c(2, 1), lwd = 1, col = c ("gray55", "black") , pch = c(21, 16))
# Adding axis with names
axis(1) ; axis(2) ; axis(4, col.axis = "gray55", col = "gray55")
mtext("Number of Defects \n per unit", side = 2,  line = 2.2, cex = 0.9)
mtext("Inspected units", side = 4,  line = 2, cex = 0.9, col = "gray55")
```

HPRD_Normal_BothUnknown

HPRD for Normal process with both parameters unknown

Description

HPRD_Normal_BothUnknown is used to calculate the Highest Predictive Density (HPrD) for the PCC process for Normal data, asssuming both the mean and the variance unknown.

Usage

```
HPRD_Normal_BothUnknown(far, Mpr, Lp, Ap, Bp)
```

Arguments

far	scalar (between 0 and 1); the false alarm rate, which reflects to an 100*(1-far)% coverage for the HPrD. The default value is 1/370.4.
Mpr	scalar; posterior predictive parameter.
Lp	scalar (positive); posterior predictive parameter.
Ap	scalar (positive); posterior predictive parameter.
Bp	scalar (positive); posterior predictive parameter.

Details

HPRD_Normal_BothUnknown provides (1-far)% the HPrD region of the PCC process for Normal data, assumming both the mean and the variance unknown.

More specifically the function provides the HPrD of a non-standardized Student's t-distribution with location parameter Mpr and scale parameter sqrt(Bp * (Lp+1) / (Ap * Lp)).

Examples

```
HPRD_Normal_BothUnknown(far = 0.0017, Mpr = 30, Lp = 12, Ap = 6, Bp = 5)
```

`HPRD_Poisson` *HPRD for Poisson process*

Description

HPRD_Poisson is used to calculate the Highest Predictive Density (HPrD) for the PCC process for Poisson data, asssuming the rate parameter unknown.

Usage

```
HPRD_Poisson(far, R, P)
```

Arguments

`far`	scalar (between 0 and 1); the false alarm rate, which reflects to an 100*(1-far)% coverage for the HPrD. The default value is 1/370.4.
`R`	scalar (positive); posterior predictive parameter.
`P`	scalar (between 0 and 1); posterior predictive parameter.

Details

HPRD_Poisson provides 100*(1-far)% the HPrD region of the PCC process for Poisson data, assumming the rate parameter unknown.

More specifically the function provides the HPrD of a negative Binomial distribution with number of failures R and failure probability P.

Examples

```
HPRD_Poisson(far = 0.0017, R = 10, P = 0.9)
```

PCC_Normal_BothUnknown

PCC for Normal with both parameters unknown

Description

PCC_Normal_BothUnknown is used to fit the PCC process for Normal data, assuming both the mean and the variance unknown.

Usage

```
PCC_Normal_BothUnknown(data = NULL, historical_data = NULL,
                       mu0 = 0, l0 = 0, a0 = -1/2, b0 = 0, alpha_0 = NULL,
                       ARL_0 = 370.4, FAP = NULL, FIR = FALSE, fFIR = .99,
                       aFIR=1/8, summary_list = TRUE, PCC_PLOT = TRUE,
                       historical_data_PLOT = FALSE, pdf_report = FALSE,
                      xlab = "Observations", ylab = "Values", main="PCC Normal")
```

Arguments

data	vector; a dataset for PCC implementation. Data needs to be in a vector form.
historical_data	vector; an optional dataset of historical data. Historical data needs to be in a vector form.
mu0	scalar; prior hyperparameter. It needs to be a number. The default is 0 and it refers to the initial reference prior.
l0	scalar (positive); prior hyperparameter. It needs to be a number. The default is 0 and it refers to the initial reference prior.
a0	scalar (positive); prior hyperparameter. It needs to be a number. The default is -1/2 and it refers to the initial reference prior.
b0	scalar (positive); prior hyperparameter. It needs to be a number. The default is 0 and it refers to the initial reference prior.
alpha_0	scalar (non negative); It is a power prior parameter controls the influence of the historical data on the posterior distribution. The default is 1/n_0, where n_0 is the size of the historical data.
ARL_0	scalar (positive); In Control (IC) Average Run Length (ARL). It is average number of IC data points that we will plot in the PCC before a false alarm occurs. The default value is 370.4
FAP	scalar (between 0 and 1); False Alarm Probability (FAP). It is the probability of raising at least one false alarm out of a pre-determined number of N hypothesis tests and it can be used instead of ARL_0. It is based on the Sidak's correction.
FIR	logical; If TRUE, then the Fast Initial Response (FIR) PCC is applied, which is an adjustment for the fisrt tests by narrowing the PCC limits.
fFIR	a number between 0 and 1; It is used if FIR=TRUE. The default value is 0.99 and represents the proportion of the adjusted PCC region over the initial one for the first test.
aFIR	non-negative number; It is used if FIR=TRUE. The default value is 0.125 and it is a smoothing parameter for the FIR adjustment.

summary_list logical; If it is TRUE, then a data frame is provided, containing the data sequence, the PCC limits and the occurence of an alarm. It is TRUE by default.

PCC_PLOT logical; if TRUE, the PCC plot is displayed. It is TRUE by default.

historical_data_PLOT
logical; if TRUE, the historical data are plotted along with the PCC plot is displayed. It is FALSE by default.

pdf_report logical; if TRUE then the summary list and the PCC plot reported in a pdf file.

xlab, ylab, main
The titles of the x-axis, y-axis and the overall title for the PCC plot.

Details

PCC_Normal_BothUnknown provides the PCC process for Normal data, assumming both the mean and the variance unknown. The PCC process is based on the sequential testing of the future observable against the Highest Predictive Density (HPrD), which is obtained by the posterior predictive distribution. The first test is available for the second observation.

The joint initial prior for the unknown parameters is a NIG(mu0,l0,a0,b0). Furthermore, the direct use of historical data is possible via the power prior, if they are available. In this case, the default value for the power prior parameter alpha_0 is the reciprocal of the length of the historical data, which conveys the weight of a single observation to the prior information. The default prior is the non-informative reference prior NIG(0,0,-1/2,0), without the use of historical data. In this case, the first test is available for the third observation.

A FIR option is available by narrowing the first few control limits. The metrics that can be used for the false alarms tolerance are ARL_0 and FAP.

Examples

```
# 30 Normal observations introducing an outlier of 3*sd at the 15th observation
set.seed(1234)
out <- rnorm(30)
out[15] <- out[15] + 3
PCC_Normal_BothUnknown(out)

# Real data application
attach(aPTT)
PCC_Normal_BothUnknown(data = aPTT_current, historical_data = aPTT_historical)
```

PCC_Poisson *PCC for Poisson with rate parameter unknown*

Description

PCC_Poisson is used to fit the PCC process for Poisson data, asssuming the rate parameter unknown.

Usage

```
PCC_Poisson(data = NULL, s = NULL, historical_data = NULL,
            historical_s = NULL, c = 1/2, d = 0, alpha_0 = NULL,
            ARL_0 = 370.4, FAP = NULL, FIR = FALSE, fFIR = .99,
            aFIR=1/8, summary_list = TRUE, PCC_PLOT = TRUE,
            historical_data_PLOT = FALSE, pdf_report = FALSE,
            xlab = "Observations", ylab = "Values", main="PCC Poisson")
```

Arguments

data	vector; a dataset for PCC implementation. Data needs to be in a vector form.
s	vector; refers to the period of time or grouping of counts, if data are rate dataset. It needs to be in a vector form.
historical_data	vector; an optional dataset of historical data. Historical data needs to be in a vector form.
historical_s	vector; refers to the period of time or grouping of counts, if historical_data are rate dataset. It needs to be in a vector form.
c	scalar (positive); prior hyperparameter. It needs to be a number. The default is 1/2 and it refers to the initial reference prior.
d	scalar (positive); prior hyperparameter. It needs to be a number. The default is 0 and it refers to the initial reference prior.
alpha_0	scalar (non negative); It is a power prior parameter controls the influence of the historical data on the posterior distribution. The default is 1/n_0, where n_0 is the size of the historical data.
ARL_0	scalar (positive); In Control (IC) Average Run Length (ARL). It is average number of IC data points that we will plot in the PCC before a false alarm occurs. The default value is 370.4
FAP	scalar (between 0 and 1); False Alarm Probability (FAP). It is the probability of raising at least one false alarm out of a pre-determined number of N hypothesis tests and it can be used instead of ARL_0. It is based on the Sidak's correction.
FIR	logical; If TRUE, then the Fast Initial Response (FIR) PCC is applied, which is an adjustment for the fisrt tests by narrowing the PCC limits.
fFIR	a number between 0 and 1; It is used if FIR=TRUE. The default value is 0.99 and represents the proportion of the adjusted PCC region over the initial one for the first test.
aFIR	non-negative number; It is used if FIR=TRUE. The default value is 0.125 and it is a smoothing parameter for the FIR adjustment.
summary_list	logical; If it is TRUE, then a data frame is provided, containing the data sequence, the PCC limits and the occurence of an alarm. It is TRUE by default.
PCC_PLOT	logical; if TRUE, the PCC plot is displayed. It is TRUE by default.
historical_data_PLOT	logical; if TRUE, the historical data are plotted along with the PCC plot is displayed. It is FALSE by default.
pdf_report	logical; if TRUE then the summary list and the PCC plot reported in a pdf file.
xlab, ylab, main	The titles of the x-axis, y-axis and the overall title for the PCC plot.

Details

PCC_Poisson provides the PCC process for Poisson data, either count or rate, assumming both the rate parameter unknown. The PCC process is based on the sequential testing of the future observable against the Highest Predictive Density (HPrD), which is obtained by the posterior predictive distribution. The first test is available for the second observation.

The initial prior for the unknown parameters is a G(c,d), where d ia the rate parameter. Furthermore, the direct use of historical data is possible via the power prior, if they are available. In this case, the

default value for the power prior parameter alpha_0 is the reciprocal of the length of the historical data, which conveys the weight of a single observation to the prior information. The default prior is the non-informative reference prior G(1/2,0), without the use of historical data.

A FIR option is available by narrowing the first few control limits. The metrics that can be used for the false alarms tolerance are ARL_0 and FAP.

Examples

```
# 30 Poisson observations introducing an outlier of 3*sd at the 15th observation
set.seed(1111)
out <- rpois(30, 4)
out[15] <- out[15] + 6
PCC_Poisson(out)

# Real data application
attach(ECE)
PCC_Poisson(data = defect_counts, s = inspected_units)
```

Index