Predictive Control Charts (PCC):
Bayesian Approach in Online Monitoring of Short Runs

K. Bourazas[1] D. Kiagias[2] P. Tsiamyrtzis[3]

[1]Dept. of Statistics, Athens University of Economics and Business, Greece

kbourazas@aueb.gr

[2]School of Mathematics and Statistics, University of Sheffield, UK

d.kiagias@sheffield.ac.uk

[3]Dept. of Mechanical Engineering, Politecnico di Milano, Italy

panagiotis.tsiamyrtzis@polimi.it

ENBIS-21 Spring Meeting, 05/18

Predictive Control Chart (PCC)

- We will propose a self-starting Bayesian method, named **Predictive Control Chart (PCC)**. We will focus in detecting outlying observations in short runs, while performing online monitoring.

Predictive Control Chart (PCC)

- We will propose a self-starting Bayesian method, named **Predictive Control Chart (PCC)**. We will focus in detecting outlying observations in short runs, while performing online monitoring.
- PCC will utilize the available prior information (or adopt an objective Bayesian approach in case of prior ignorance) and it will be formed as a sequentially updated region based on the predictive distribution.

Predictive Control Chart (PCC)

- We will propose a self-starting Bayesian method, named **Predictive Control Chart (PCC)**. We will focus in detecting outlying observations in short runs, while performing online monitoring.
- PCC will utilize the available prior information (or adopt an objective Bayesian approach in case of prior ignorance) and it will be formed as a sequentially updated region based on the predictive distribution.
- PCC will be introduced in a general form, allowing to handle data of any univariate (discrete or continuous) distribution, as long as this distribution is a member of the k-Parameter Regular Exponential Family (k-PREF).

PCC: statistical modeling

- From a process under study, we sequentially obtain the data $\boldsymbol{X} = (x_1, \ldots, x_n)$, which we consider to be a random sample from the distribution $X_j|\boldsymbol{\theta}$, a member of the k-PREF:

PCC: statistical modeling

- From a process under study, we sequentially obtain the data $\boldsymbol{X} = (x_1, \ldots, x_n)$, which we consider to be a random sample from the distribution $X_j|\boldsymbol{\theta}$, a member of the k-PREF:

$$f(\boldsymbol{X}|\boldsymbol{\theta}) = \left[\prod_{j=1}^{n} g(x_j)\right] [c(\boldsymbol{\theta})]^n \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta}) \sum_{j=1}^{n} h_i(x_j)\right\}$$

PCC: statistical modeling

- From a process under study, we sequentially obtain the data $\boldsymbol{X} = (x_1, \ldots, x_n)$, which we consider to be a random sample from the distribution $X_j|\boldsymbol{\theta}$, a member of the k-PREF:

$$f(\boldsymbol{X}|\boldsymbol{\theta}) = \left[\prod_{j=1}^{n} g(x_j)\right] [c(\boldsymbol{\theta})]^n \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta}) \sum_{j=1}^{n} h_i(x_j)\right\}$$

- Among the most prominent representatives in SPC/M are:

PCC: statistical modeling

- From a process under study, we sequentially obtain the data $\boldsymbol{X} = (x_1, \ldots, x_n)$, which we consider to be a random sample from the distribution $X_j|\boldsymbol{\theta}$, a member of the k-PREF:

$$f(\boldsymbol{X}|\boldsymbol{\theta}) = \left[\prod_{j=1}^{n} g(x_j)\right] [c(\boldsymbol{\theta})]^n \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta}) \sum_{j=1}^{n} h_i(x_j)\right\}$$

- Among the most prominent representatives in SPC/M are:
  - $X_j|\theta \sim Binomial(N_j, \theta)$,

PCC: statistical modeling

- From a process under study, we sequentially obtain the data $\boldsymbol{X} = (x_1, \ldots, x_n)$, which we consider to be a random sample from the distribution $X_j|\boldsymbol{\theta}$, a member of the k-PREF:

$$f(\boldsymbol{X}|\boldsymbol{\theta}) = \left[\prod_{j=1}^{n} g(x_j)\right] [c(\boldsymbol{\theta})]^n \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta}) \sum_{j=1}^{n} h_i(x_j)\right\}$$

- Among the most prominent representatives in SPC/M are:
  - $X_j|\theta \sim Binomial(N_j, \theta)$,
  - $X_j|\theta \sim Poisson(\theta)$,

PCC: statistical modeling

- From a process under study, we sequentially obtain the data $\boldsymbol{X} = (x_1, \ldots, x_n)$, which we consider to be a random sample from the distribution $X_j|\boldsymbol{\theta}$, a member of the k-PREF:

$$f(\boldsymbol{X}|\boldsymbol{\theta}) = \left[\prod_{j=1}^{n} g(x_j)\right] [c(\boldsymbol{\theta})]^n \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta}) \sum_{j=1}^{n} h_i(x_j)\right\}$$

- Among the most prominent representatives in SPC/M are:
  - $X_j|\theta \sim \mathit{Binomial}(N_j, \theta)$,
  - $X_j|\theta \sim \mathit{Poisson}(\theta)$,
  - $X_j|\boldsymbol{\theta} \sim \mathit{Normal}(\theta_1, \theta_2^2)$ etc.

PCC: statistical modeling

- From a process under study, we sequentially obtain the data $\boldsymbol{X} = (x_1, \ldots, x_n)$, which we consider to be a random sample from the distribution $X_j|\boldsymbol{\theta}$, a member of the k-PREF:

$$f(\boldsymbol{X}|\boldsymbol{\theta}) = \left[\prod_{j=1}^{n} g(x_j)\right] [c(\boldsymbol{\theta})]^n \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta}) \sum_{j=1}^{n} h_i(x_j)\right\}$$

- Among the most prominent representatives in SPC/M are:
  - $X_j|\theta \sim Binomial(N_j, \theta)$,
  - $X_j|\theta \sim Poisson(\theta)$,
  - $X_j|\boldsymbol{\theta} \sim Normal(\theta_1, \theta_2^2)$ etc.
- Our main interest is in detecting in an online fashion and without employing a phase I exercise, the presence of large transient shifts (outliers) on the unknown parameter(s) $\boldsymbol{\theta}$.

PCC: statistical modeling

- For the prior, since in practice, historical data (of the same or a similar process, not to be confused with phase I data) are typically available, we recommend the use of **power priors** (Ibrahim and Chen, 2000):

PCC: statistical modeling

- For the prior, since in practice, historical data (of the same or a similar process, not to be confused with phase I data) are typically available, we recommend the use of **power priors** (Ibrahim and Chen, 2000):

$$\pi(\boldsymbol{\theta}|\boldsymbol{Y}, \alpha_0, \boldsymbol{\tau}) \propto f(\boldsymbol{Y}|\boldsymbol{\theta})^{\alpha_0} \pi_0(\boldsymbol{\theta}|\boldsymbol{\tau})$$

where $\boldsymbol{Y} = (y_1, \ldots, y_{n_0})$ are the historical data, $0 \leq \alpha_0 \leq 1$ is a scalar parameter, $\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau})$ is the **initial prior** for the unknown parameter(s) and $\boldsymbol{\tau}$ is the vector of the initial prior hyperparameters.

PCC: statistical modeling

- For the prior, since in practice, historical data (of the same or a similar process, not to be confused with phase I data) are typically available, we recommend the use of **power priors** (Ibrahim and Chen, 2000):

$$\pi\left(\boldsymbol{\theta}|\boldsymbol{Y},\alpha_0,\boldsymbol{\tau}\right) \propto f\left(\boldsymbol{Y}|\boldsymbol{\theta}\right)^{\alpha_0}\pi_0\left(\boldsymbol{\theta}|\boldsymbol{\tau}\right)$$

  where $\boldsymbol{Y} = (y_1, \ldots, y_{n_0})$ are the historical data, $0 \leq \alpha_0 \leq 1$ is a scalar parameter, $\pi_0\left(\boldsymbol{\theta}|\boldsymbol{\tau}\right)$ is the **initial prior** for the unknown parameter(s) and $\boldsymbol{\tau}$ is the vector of the initial prior hyperparameters.

- From a subjective Bayesian point of view, $\pi_0(\cdot)$ should reflect all available information regarding the unknown parameter(s) **before** the data become available, using prior knowledge, expert's opinion etc.

PCC: statistical modeling

- For the prior, since in practice, historical data (of the same or a similar process, not to be confused with phase I data) are typically available, we recommend the use of **power priors** (Ibrahim and Chen, 2000):

$$\pi\left(\boldsymbol{\theta} | \boldsymbol{Y}, \alpha_0, \boldsymbol{\tau}\right) \propto f\left(\boldsymbol{Y} | \boldsymbol{\theta}\right)^{\alpha_0} \pi_0\left(\boldsymbol{\theta} | \boldsymbol{\tau}\right)$$

  where $\boldsymbol{Y} = (y_1, \ldots, y_{n_0})$ are the historical data, $0 \leq \alpha_0 \leq 1$ is a scalar parameter, $\pi_0\left(\boldsymbol{\theta} | \boldsymbol{\tau}\right)$ is the **initial prior** for the unknown parameter(s) and $\boldsymbol{\tau}$ is the vector of the initial prior hyperparameters.

- From a subjective Bayesian point of view, $\pi_0(\cdot)$ should reflect all available information regarding the unknown parameter(s) **before** the data become available, using prior knowledge, expert's opinion etc.

- From an objective Bayesian point of view we can adopt for $\pi_0(\cdot)$ a weakly informative or even non-informative initial prior, such as flat/Jeffreys/reference/... prior.

PCC: statistical modeling

- To preserve closed form solutions, we suggest a conjugate prior for $\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau})$, which always exists for any likelihood that is a member of the k-PREF and its form is given by:

PCC: statistical modeling

- To preserve closed form solutions, we suggest a conjugate prior for $\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau})$, which always exists for any likelihood that is a member of the k-PREF and its form is given by:

$$\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau}) = [K(\boldsymbol{\tau})]^{-1} [c(\boldsymbol{\theta})]^{\tau_0} \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta})\tau_i\right\}$$

where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ (parameter space) and $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_k)$ is the $(k+1)$-dimensional vector of the initial prior hyperparameters (reflect the prior knowledge, ranging from highly informative to vague and even non-informative choices).

PCC: statistical modeling

- To preserve closed form solutions, we suggest a conjugate prior for $\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau})$, which always exists for any likelihood that is a member of the k-PREF and its form is given by:

$$\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau}) = [K(\boldsymbol{\tau})]^{-1} [c(\boldsymbol{\theta})]^{\tau_0} \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta})\tau_i\right\}$$

  where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ (parameter space) and $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_k)$ is the $(k+1)$-dimensional vector of the initial prior hyperparameters (reflect the prior knowledge, ranging from highly informative to vague and even non-informative choices).

- A conjugate $\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau})$ will lead to a conjugate power prior of the form:

PCC: statistical modeling

- To preserve closed form solutions, we suggest a conjugate prior for $\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau})$, which always exists for any likelihood that is a member of the k-PREF and its form is given by:

$$\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau}) = [K(\boldsymbol{\tau})]^{-1} [c(\boldsymbol{\theta})]^{\tau_0} \exp\left\{\sum_{i=1}^{k} \eta_i(\boldsymbol{\theta})\tau_i\right\}$$

  where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ (parameter space) and $\boldsymbol{\tau} = (\tau_0, \tau_1, \ldots, \tau_k)$ is the $(k+1)$-dimensional vector of the initial prior hyperparameters (reflect the prior knowledge, ranging from highly informative to vague and even non-informative choices).

- A conjugate $\pi_0(\boldsymbol{\theta}|\boldsymbol{\tau})$ will lead to a conjugate power prior of the form:

$$\pi(\boldsymbol{\theta}|\boldsymbol{Y}, \alpha_0, \boldsymbol{\tau}) \propto \pi_0(\boldsymbol{\theta}|\boldsymbol{\tau} + \alpha_0 \boldsymbol{t}_{n_0}(\boldsymbol{Y}))$$

PCC: statistical modeling

- Then the posterior distribution of the unknown parameter(s) $\boldsymbol{\theta}$ will be:

PCC: statistical modeling

- Then the posterior distribution of the unknown parameter(s) $\boldsymbol{\theta}$ will be:

$$p\left(\boldsymbol{\theta} | \boldsymbol{X}, \boldsymbol{Y}, \alpha_0, \boldsymbol{\tau}\right) = \pi_0\left(\boldsymbol{\theta} | \boldsymbol{\tau} + \alpha_0 \boldsymbol{t}_{n_0}(\boldsymbol{Y}) + \boldsymbol{t}_n(\boldsymbol{X})\right)$$

PCC: statistical modeling

- Then the posterior distribution of the unknown parameter(s) $\boldsymbol{\theta}$ will be:

$$p\left(\boldsymbol{\theta}|\boldsymbol{X},\boldsymbol{Y},\alpha_0,\boldsymbol{\tau}\right)=\pi_0\left(\boldsymbol{\theta}|\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})\right)$$

while the predictive distribution of the single future observable X_{n+1} will be:

$$f\left(X_{n+1}|\boldsymbol{X},\boldsymbol{Y},\alpha_0,\boldsymbol{\tau}\right)=\frac{K\left(\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})+\boldsymbol{t}_1(X_{n+1})\right)}{K\left(\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})\right)}g(X_{n+1})$$

PCC: statistical modeling

- Then the posterior distribution of the unknown parameter(s) $\boldsymbol{\theta}$ will be:

$$p\left(\boldsymbol{\theta}|\boldsymbol{X},\boldsymbol{Y},\alpha_0,\boldsymbol{\tau}\right)=\pi_0\left(\boldsymbol{\theta}|\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})\right)$$

while the predictive distribution of the single future observable X_{n+1} will be:

$$f\left(X_{n+1}|\boldsymbol{X},\boldsymbol{Y},\alpha_0,\boldsymbol{\tau}\right)=\frac{K\left(\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})+\boldsymbol{t}_1(X_{n+1})\right)}{K\left(\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})\right)}g(X_{n+1})$$

- PCC construction will be based on the predictive distribution and it can start as soon as $n=2$ and is based on the sequentially updated form of the predictive distribution.

PCC: statistical modeling

- Then the posterior distribution of the unknown parameter(s) $\boldsymbol{\theta}$ will be:

$$p\left(\boldsymbol{\theta}|\boldsymbol{X},\boldsymbol{Y},\alpha_0,\boldsymbol{\tau}\right)=\pi_0\left(\boldsymbol{\theta}|\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})\right)$$

while the predictive distribution of the single future observable X_{n+1} will be:

$$f\left(X_{n+1}|\boldsymbol{X},\boldsymbol{Y},\alpha_0,\boldsymbol{\tau}\right)=\frac{K\left(\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})+\boldsymbol{t}_1(X_{n+1})\right)}{K\left(\boldsymbol{\tau}+\alpha_0\boldsymbol{t}_{n_0}(\boldsymbol{Y})+\boldsymbol{t}_n(\boldsymbol{X})\right)}g(X_{n+1})$$

- PCC construction will be based on the predictive distribution and it can start as soon as $n=2$ and is based on the sequentially updated form of the predictive distribution.

- Precisely, we will determine an IC region, R_{n+1}, where the future observable (X_{n+1}) will most likely be, as long as the process is stable (i.e. no changes occurred).

PCC: Highest Predictive Density/Mass (HPrD/M)

- Assume the set R^c which contains the values of the predictive density (or mass) function, which are greater than a threshold c, i.e.:

  $$R^c = \{x_{n+1} : f(x_{n+1} | \boldsymbol{X}, \boldsymbol{Y}, \alpha_0, \boldsymbol{\tau}) \geq c\}$$

  The HPrD/M region will be given by minimizing the absolute difference of a highest predictive probability from a significance level $1 - \alpha$, for all the possible values of c. Specifically:

  $$R_{n+1} = \min_{R^c} \left| \int_{R^c} f(x_{n+1} | \boldsymbol{X}, \boldsymbol{Y}, \alpha_0, \boldsymbol{\tau}) - (1 - \alpha) \right|,$$

  for the discrete case, we replace the integral sign by summation.

PCC: Highest Predictive Density/Mass (HPrD/M)

- Assume the set R^c which contains the values of the predictive density (or mass) function, which are greater than a threshold c, i.e.:

$$R^c = \{x_{n+1} : f(x_{n+1}|\boldsymbol{X}, \boldsymbol{Y}, \alpha_0, \boldsymbol{\tau}) \geq c\}$$

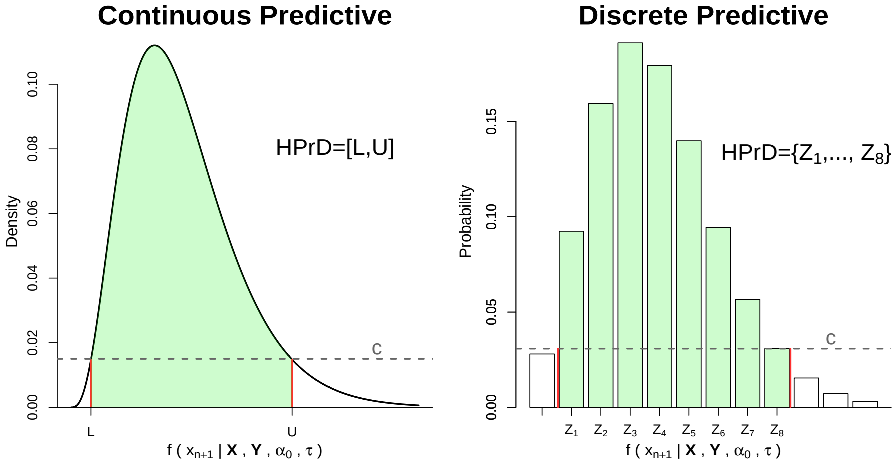
The HPrD/M region will be given by minimizing the absolute difference of a highest predictive probability from a significance level $1 - \alpha$, for all the possible values of c. Specifically:

$$R_{n+1} = \min_{R^c} \left| \int_{R^c} f(x_{n+1}|\boldsymbol{X}, \boldsymbol{Y}, \alpha_0, \boldsymbol{\tau}) - (1 - \alpha) \right|,$$

for the discrete case, we replace the integral sign by summation.

- R_{n+1} will be the shortest region with the smallest absolute difference from the probability $1 - \alpha$.

PCC: Highest Predictive Density/Mass (HPrD/M)



PCC: False Alarm Tolerance

- The (predetermined) parameter $0 < \alpha < 1$ will reflect our tolerance to false alarms and detection power.

PCC: False Alarm Tolerance

- The (predetermined) parameter $0 < \alpha < 1$ will reflect our tolerance to false alarms and detection power.
- The proposed PCC is a sequential (multiple) hypothesis testing procedure, where at each time point n, the probability to raise a false alarm is: $P\left(X_{n+1} \notin R_{n+1} | IC\right) = \alpha$.

PCC: False Alarm Tolerance

- The (predetermined) parameter $0 < \alpha < 1$ will reflect our tolerance to false alarms and detection power.
- The proposed PCC is a sequential (multiple) hypothesis testing procedure, where at each time point n, the probability to raise a false alarm is: $P\left(X_{n+1} \notin R_{n+1} | IC\right) = \alpha$.
- We suggest two metrics in selecting α:

PCC: False Alarm Tolerance

- The (predetermined) parameter $0 < \alpha < 1$ will reflect our tolerance to false alarms and detection power.
- The proposed PCC is a sequential (multiple) hypothesis testing procedure, where at each time point n, the probability to raise a false alarm is: $P\left(X_{n+1} \notin R_{n+1} | IC\right) = \alpha$.
- We suggest two metrics in selecting α:
  1. If we have a (known) fixed horizon of N data points and we wish to bound the Family Wise Error Rate (FWER) by Šidák's correction:

$$\alpha = 1 - (1 - FWER)^{\frac{1}{N-1}}$$

PCC: False Alarm Tolerance

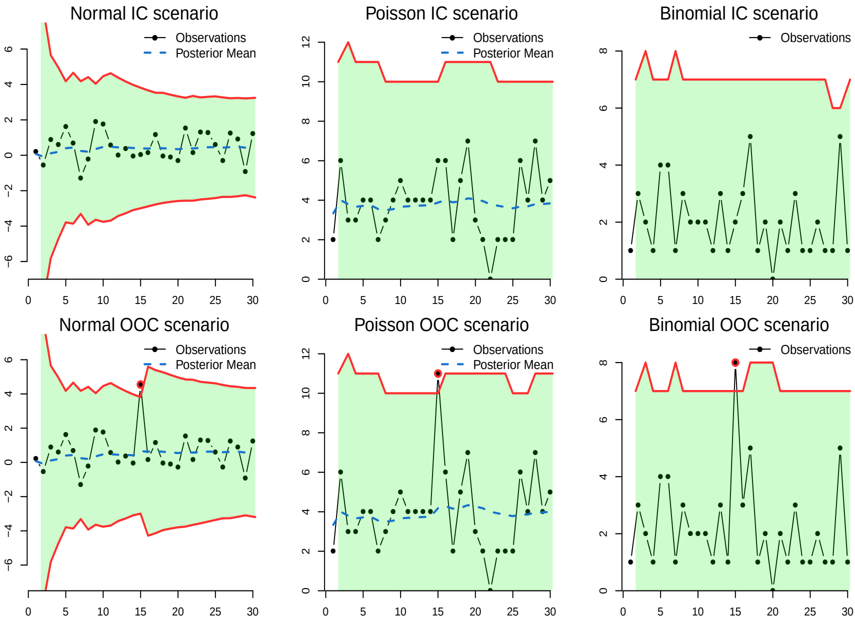
- The (predetermined) parameter $0 < \alpha < 1$ will reflect our tolerance to false alarms and detection power.
- The proposed PCC is a sequential (multiple) hypothesis testing procedure, where at each time point n, the probability to raise a false alarm is: $P\left(X_{n+1} \notin R_{n+1} | IC\right) = \alpha$.
- We suggest two metrics in selecting α:
  1. If we have a (known) fixed horizon of N data points and we wish to bound the Family Wise Error Rate (FWER) by Šidák's correction:

  $$\alpha = 1 - (1 - FWER)^{\frac{1}{N-1}}$$

  2. If N is either unknown in advance or it is too large, then we suggest to derive α using the metric of IC Average Run Length (ARL_0):

  $$\alpha \approx \frac{1}{ARL_0}$$

PCC: Illustration and Decision Making



PCC: Competing Methods and Sensitivity Analysis

- We will compare the performance of the suggested **PCC** (Bayesian parametric) against **Q-chart**.

PCC: Competing Methods and Sensitivity Analysis

- We will compare the performance of the suggested **PCC** (Bayesian parametric) against **Q-chart**.

- We generate 100,000 iterations of $N = 30$ data points from the IC states, which are: $N(0, 1)$, $P(2)$ and $Bin(20, 0.1)$

PCC: Competing Methods and Sensitivity Analysis

- We will compare the performance of the suggested **PCC** (Bayesian parametric) against **Q-chart**.
- We generate 100,000 iterations of $N = 30$ data points from the IC states, which are: $N(0,1)$, $P(2)$ and $Bin(20, 0.1)$
- The OOC states represent isolated shifts of size $\{2.5 \text{ or } 3\} \times \text{sd}$, which are introduced to the IC sequences at one of the locations: $\{5, \text{ or } 15 \text{ or } 25\}$.

PCC: Competing Methods and Sensitivity Analysis

- We will also examine the sensitivity of the PCC performance for various prior settings.

PCC: Competing Methods and Sensitivity Analysis

- We will also examine the sensitivity of the PCC performance for various prior settings.

- For each setup, we will make use of two initial priors (reference/objective and weakly informative) with the absence or the presence of $n_0 = 10$ historical data $\boldsymbol{Y}$. Thus, we will have four versions of PCC (with/without prior knowledge, with/without historical data). The initial priors $\pi_0(\cdot|\boldsymbol{\tau})$, which we will employ are:

PCC: Competing Methods and Sensitivity Analysis

- We will also examine the sensitivity of the PCC performance for various prior settings.

- For each setup, we will make use of two initial priors (reference/objective and weakly informative) with the absence or the presence of $n_0 = 10$ historical data $\boldsymbol{Y}$. Thus, we will have four versions of PCC (with/without prior knowledge, with/without historical data). The initial priors $\pi_0(\cdot|\boldsymbol{\tau})$, which we will employ are:
    - Normal: reference prior $\pi_0\left(\theta_1, \theta_2^2\right) \propto 1/\theta_2^2 \equiv NIG(0, 0, -1/2, 0)$ or the weakly informative $NIG(0, 2, 1, 0.8)$.

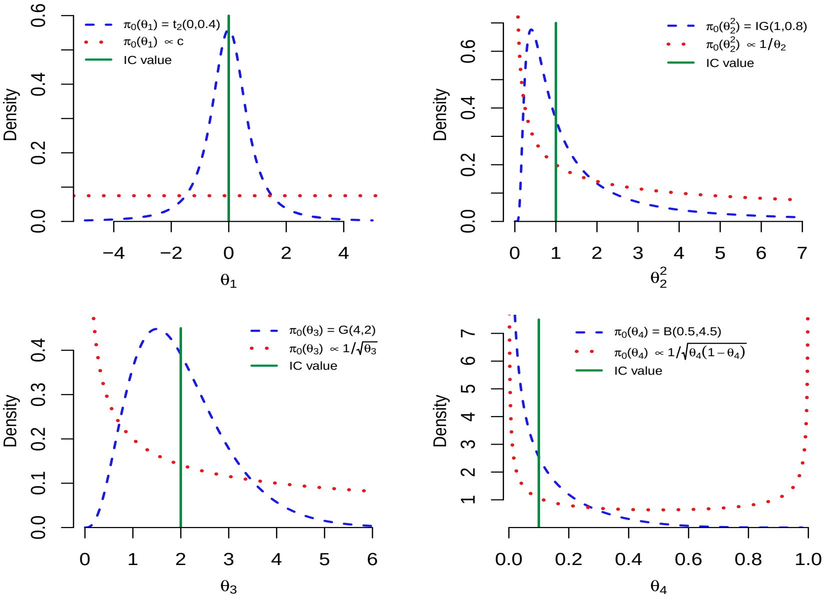
PCC: Competing Methods and Sensitivity Analysis

- We will also examine the sensitivity of the PCC performance for various prior settings.
- For each setup, we will make use of two initial priors (reference/objective and weakly informative) with the absence or the presence of $n_0 = 10$ historical data $\boldsymbol{Y}$. Thus, we will have four versions of PCC (with/without prior knowledge, with/without historical data). The initial priors $\pi_0(\cdot|\boldsymbol{\tau})$, which we will employ are:
  - Normal: reference prior $\pi_0\left(\theta_1, \theta_2^2\right) \propto 1/\theta_2^2 \equiv NIG(0, 0, -1/2, 0)$ or the weakly informative $NIG(0, 2, 1, 0.8)$.
  - Poisson: reference prior $\pi_0\left(\theta_3\right) \propto 1/\sqrt{\theta_3} \equiv G(1/2, 0)$ or the weakly informative $G(4, 2)$.

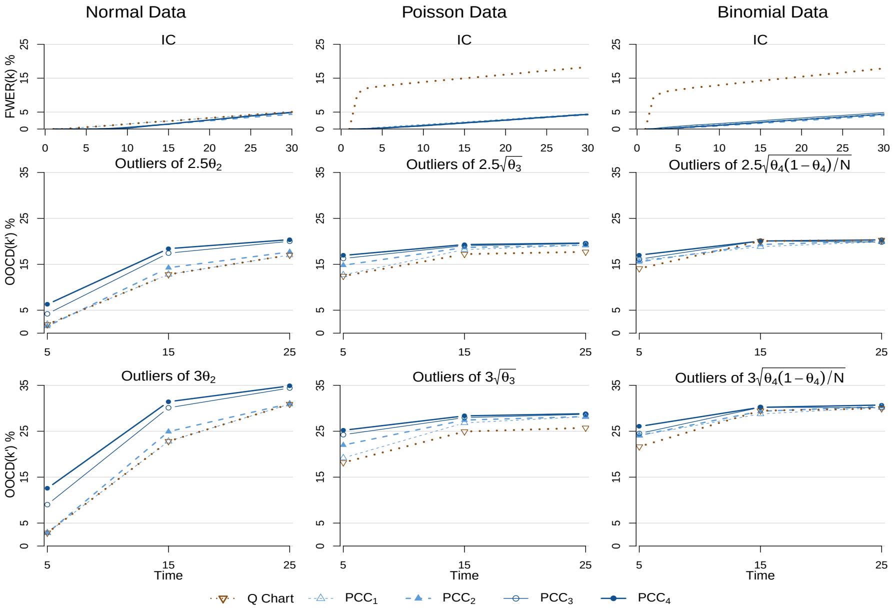
PCC: Competing Methods and Sensitivity Analysis

- We will also examine the sensitivity of the PCC performance for various prior settings.
- For each setup, we will make use of two initial priors (reference/objective and weakly informative) with the absence or the presence of $n_0 = 10$ historical data $\boldsymbol{Y}$. Thus, we will have four versions of PCC (with/without prior knowledge, with/without historical data). The initial priors $\pi_0(\cdot|\boldsymbol{\tau})$, which we will employ are:
  - Normal: reference prior $\pi_0\left(\theta_1, \theta_2^2\right) \propto 1/\theta_2^2 \equiv NIG(0, 0, -1/2, 0)$ or the weakly informative $NIG(0, 2, 1, 0.8)$.
  - Poisson: reference prior $\pi_0\left(\theta_3\right) \propto 1/\sqrt{\theta_3} \equiv G(1/2, 0)$ or the weakly informative $G(4, 2)$.
  - Binomial: reference prior $\pi_0\left(\theta_4\right) \propto 1/\sqrt{\theta_4(1-\theta_4)} \equiv Beta(1/2, 1/2)$ or the weakly informative $Beta(0.5, 4.5)$.

PCC: Initial Priors (Sensitivity Analysis)



PCC: Competing Methods and Sensitivity Analysis



PCC Real Data Application (Medical Lab)

- We will use data that come from the daily Internal Quality Control (IQC) routine of a medical laboratory, monitoring "activated Partial Thromboplastin Time" (aPTT), measured in seconds.

PCC Real Data Application (Medical Lab)

- We will use data that come from the daily Internal Quality Control (IQC) routine of a medical laboratory, monitoring "activated Partial Thromboplastin Time" (aPTT), measured in seconds.
- We gathered 30 daily normal IQC observations (X_i) from a medical lab. Notice that these data are based on control samples and in regular practice will become available sequentially.

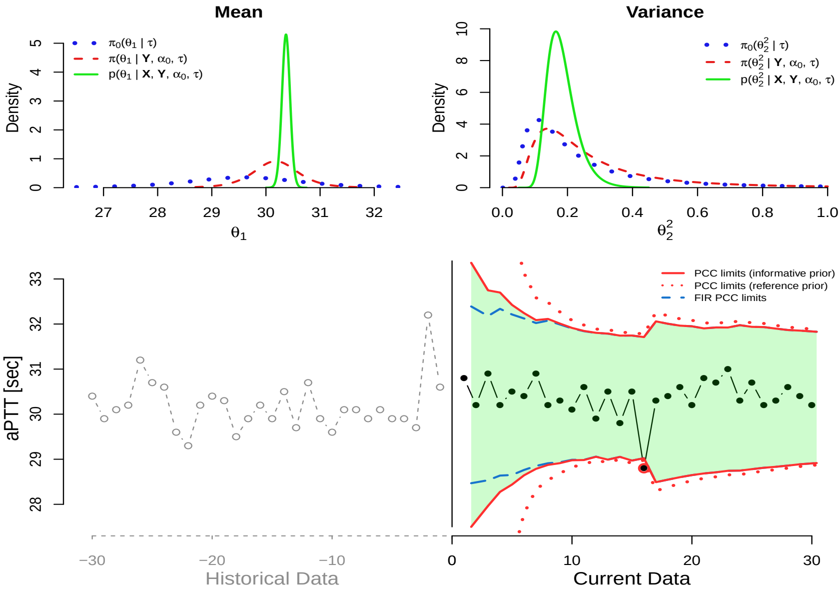
PCC Real Data Application (Medical Lab)

- We will use data that come from the daily Internal Quality Control (IQC) routine of a medical laboratory, monitoring "activated Partial Thromboplastin Time" (aPTT), measured in seconds.
- We gathered 30 daily normal IQC observations (X_i) from a medical lab. Notice that these data are based on control samples and in regular practice will become available sequentially.
- The goal is to accurately detect any transient parameter shift of large size, as this will have an impact on the reported patient results. Thus, it is of major importance to perform on-line monitoring of the process without a phase I exercise.

PCC Real Data Application (Medical Lab)

- We will use data that come from the daily Internal Quality Control (IQC) routine of a medical laboratory, monitoring "activated Partial Thromboplastin Time" (aPTT), measured in seconds.
- We gathered 30 daily normal IQC observations (X_i) from a medical lab. Notice that these data are based on control samples and in regular practice will become available sequentially.
- The goal is to accurately detect any transient parameter shift of large size, as this will have an impact on the reported patient results. Thus, it is of major importance to perform on-line monitoring of the process without a phase I exercise.
- We elicit the prior $\pi_0\left(\theta_1, {\theta_2}^2 | \boldsymbol{\tau}\right) \sim NIG\left(29.6, 1/7, 2, 0.56^2\right)$ and we had $n_0 = 30$ historical data (from a different reagent).

PCC Real Data Application (Medical Lab)



PCC Real Data Application (Industrial Setting)

- PCC illustration for discrete (Poisson) data. The data come from Hansen and Ghare (1987) and were also analyzed by Bayarri and García-Donato (2005).

PCC Real Data Application (Industrial Setting)

- PCC illustration for discrete (Poisson) data. The data come from Hansen and Ghare (1987) and were also analyzed by Bayarri and García-Donato (2005).

- They refer to the number of defects (x_i), per inspected number of units (s_i), encountered in a complex electrical equipment of an assembly line.

PCC Real Data Application (Industrial Setting)

- PCC illustration for discrete (Poisson) data. The data come from Hansen and Ghare (1987) and were also analyzed by Bayarri and García-Donato (2005).

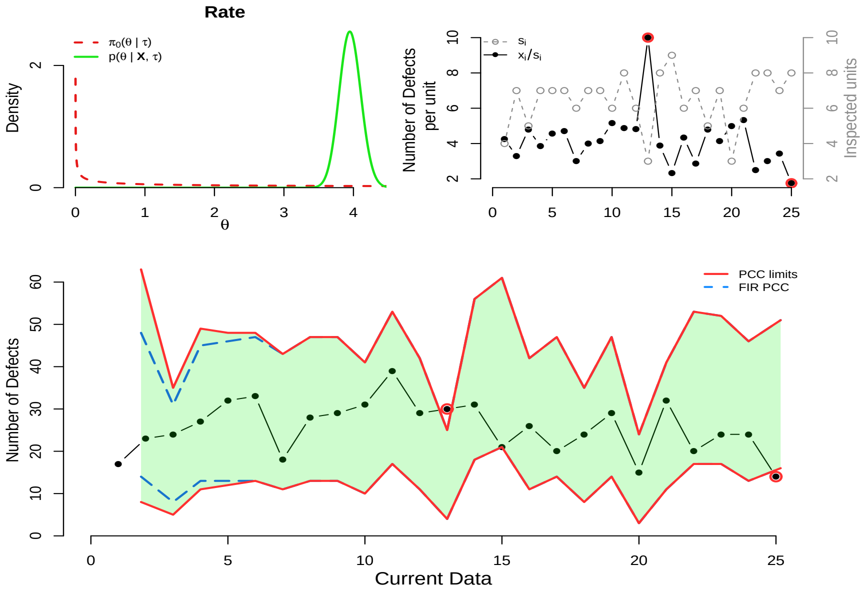
- They refer to the number of defects (x_i), per inspected number of units (s_i), encountered in a complex electrical equipment of an assembly line.

- We have 25 counts arriving sequentially that we will model using the Poisson distribution with unknown rate parameter, i.e. $X_i|\theta \sim P(\theta \cdot s_i)$.

PCC Real Data Application (Industrial Setting)

- PCC illustration for discrete (Poisson) data. The data come from Hansen and Ghare (1987) and were also analyzed by Bayarri and García-Donato (2005).
- They refer to the number of defects (x_i), per inspected number of units (s_i), encountered in a complex electrical equipment of an assembly line.
- We have 25 counts arriving sequentially that we will model using the Poisson distribution with unknown rate parameter, i.e. $X_i|\theta \sim P\left(\theta \cdot s_i\right)$.
- In this data set neither prior information regarding the unknown parameter nor historical data exist. Therefore, we use the reference prior as initial prior for θ, i.e. $\pi_0(\theta|\boldsymbol{\tau}) \propto 1/\sqrt{\theta} \equiv G(1/2, 0)$ and we also set $\alpha_0 = 0$ for the power prior term.

PCC Real Data Application (Industrial Setting)



Conclusions

We introduced the Predictive Control Chart (PCC) mechanism which attempts to solve the online control/monitoring of phase I or short run data, but can be used in phase II data as well. Specifically, PCC:

Conclusions

We introduced the Predictive Control Chart (PCC) mechanism which attempts to solve the online control/monitoring of phase I or short run data, but can be used in phase II data as well. Specifically, PCC:

- Offers axiomatic framework to incorporate prior information.

Conclusions

We introduced the Predictive Control Chart (PCC) mechanism which attempts to solve the online control/monitoring of phase I or short run data, but can be used in phase II data as well. Specifically, PCC:

- Offers axiomatic framework to incorporate prior information.
- It is self-starting (i.e. free of phase I), allowing inference from the second observable.

Conclusions

We introduced the Predictive Control Chart (PCC) mechanism which attempts to solve the online control/monitoring of phase I or short run data, but can be used in phase II data as well. Specifically, PCC:

- Offers axiomatic framework to incorporate prior information.
- It is self-starting (i.e. free of phase I), allowing inference from the second observable.
- Outperforms frequentist's based self starting and non-parametric sequential methods.

Conclusions

We introduced the Predictive Control Chart (PCC) mechanism which attempts to solve the online control/monitoring of phase I or short run data, but can be used in phase II data as well. Specifically, PCC:

- Offers axiomatic framework to incorporate prior information.
- It is self-starting (i.e. free of phase I), allowing inference from the second observable.
- Outperforms frequentist's based self starting and non-parametric sequential methods.
- It is general enough to be used in charting any type of data (discrete or continuous) that belong to the k-PREF.

Conclusions

We introduced the Predictive Control Chart (PCC) mechanism which attempts to solve the online control/monitoring of phase I or short run data, but can be used in phase II data as well. Specifically, PCC:

- Offers axiomatic framework to incorporate prior information.
- It is self-starting (i.e. free of phase I), allowing inference from the second observable.
- Outperforms frequentist's based self starting and non-parametric sequential methods.
- It is general enough to be used in charting any type of data (discrete or continuous) that belong to the k-PREF.
- The present work has been published in the *Journal of Quality Technology* at https://www.tandfonline.com/doi/full/10.1080/00224065.2021.1916413?scroll=top&needAccess=true

Conclusions

We introduced the Predictive Control Chart (PCC) mechanism which attempts to solve the online control/monitoring of phase I or short run data, but can be used in phase II data as well. Specifically, PCC:

- Offers axiomatic framework to incorporate prior information.
- It is self-starting (i.e. free of phase I), allowing inference from the second observable.
- Outperforms frequentist's based self starting and non-parametric sequential methods.
- It is general enough to be used in charting any type of data (discrete or continuous) that belong to the k-PREF.
- The present work has been published in the *Journal of Quality Technology* at https://www.tandfonline.com/doi/full/10.1080/00224065.2021.1916413?scroll=top&needAccess=true
- A library in R has been developed and it is available via Github at https://github.com/BayesianSPCM/BSPCM

Acknowledgements

- We are grateful to **Frederic Sobas** from Hospices Civils de Lyon, who provided the data set used in the Normal PCC, but more importantly for his invaluable feedback from using the suggested PCC (along with other Bayesian SPC/M methods) at the daily Internal Quality Control routine in the medical labs of Hospices Civils de Lyon.

- The was partially funded by the Research Center of the Athens University of Economics and Business (RC/AUEB).

The END

Thank you!

Questions?