Utilisation des données à priori des fournisseurs pour une gestion efficiente des résultats de CIQ: les atouts de la logique Bayésienne vs l'approche conventionnelle

Panagiotis (Panos) Tsiamyrtzis[1,2]

[1]Dept. of Mechanical Engineering, Politecnico di Milano, Italy
panagiotis.tsiamyrtzis@polimi.it

[2]Dept. of Statistics, Athens University of Economics and Business, Greece
pt@aueb.gr

SFTH - Lyon, 01 Oct 2025

Frequentist versus Bayesian Approach in Statistics

Data and Parameters in Statistics

- In parametric statistics a data set is becoming available via a random mechanism.

Data and Parameters in Statistics

- In parametric statistics a data set is becoming available via a random mechanism.
- A distribution $f(X|\theta)$ is used to describe the data generation procedure.

Data and Parameters in Statistics

- In parametric statistics a data set is becoming available via a random mechanism.
- A distribution $f(X|\theta)$ is used to describe the data generation procedure.
- The X refers to the observed data, while θ is the unknown related parameter.

Data and Parameters in Statistics

- In parametric statistics a data set is becoming available via a random mechanism.
- A distribution $f(X|\theta)$ is used to describe the data generation procedure.
- The X refers to the observed data, while θ is the unknown related parameter.
- For example:

Data and Parameters in Statistics

- In parametric statistics a data set is becoming available via a random mechanism.
- A distribution $f(X|\theta)$ is used to describe the data generation procedure.
- The X refers to the observed data, while θ is the unknown related parameter.
- For example:
  - $X|\theta \sim Bernoulli(\theta)$ can be used under certain assumption to model a discrete binary random variable.

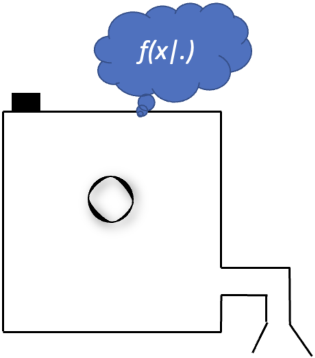
Data and Parameters in Statistics

- In parametric statistics a data set is becoming available via a random mechanism.
- A distribution $f(X|\theta)$ is used to describe the data generation procedure.
- The X refers to the observed data, while θ is the unknown related parameter.
- For example:
    - $X|\theta \sim Bernoulli(\theta)$ can be used under certain assumption to model a discrete binary random variable.
    - $X|\left(\theta_1, \theta_2^2\right) \sim N\left(\theta_1, \theta_2^2\right)$ is a distribution that can model continuous data that obey certain properties.

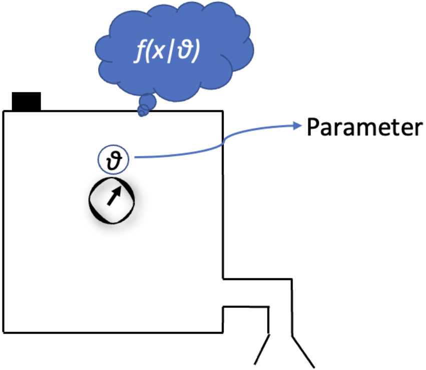
Data and Parameters in Statistics

- In parametric statistics a data set is becoming available via a random mechanism.
- A distribution $f(X|\theta)$ is used to describe the data generation procedure.
- The X refers to the observed data, while θ is the unknown related parameter.
- For example:
  - $X|\theta \sim Bernoulli(\theta)$ can be used under certain assumption to model a discrete binary random variable.
  - $X|\left(\theta_1, \theta_2^2\right) \sim N\left(\theta_1, \theta_2^2\right)$ is a distribution that can model continuous data that obey certain properties.
- Goodness of fit tests and plots should be used to examine how well an assumed theoretical distribution fits the observed data.

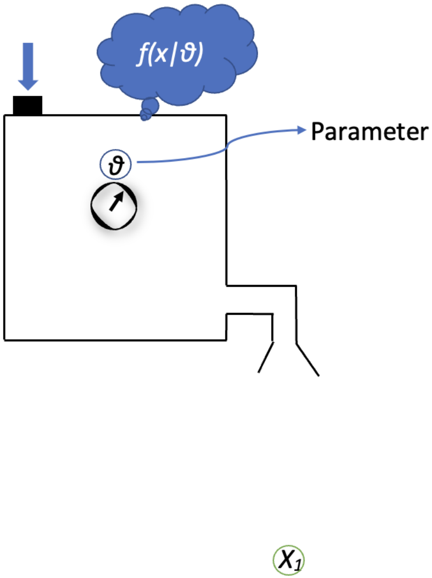
Distribution: random number generation engine



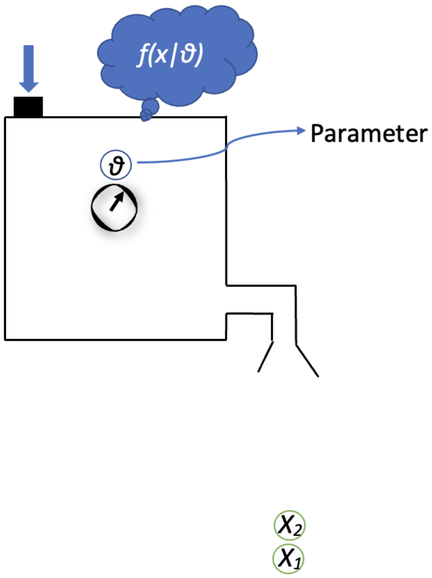
Distribution: random number generation engine



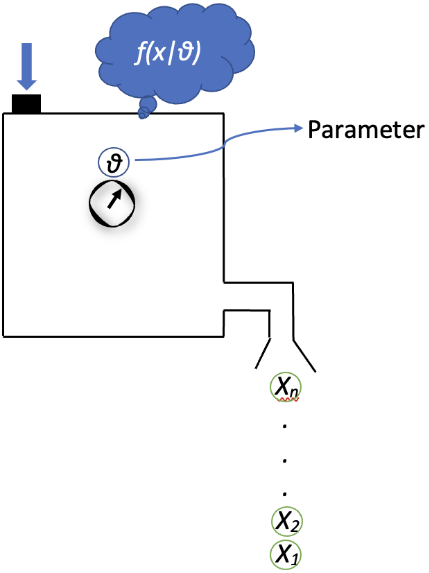
Distribution: random number generation engine



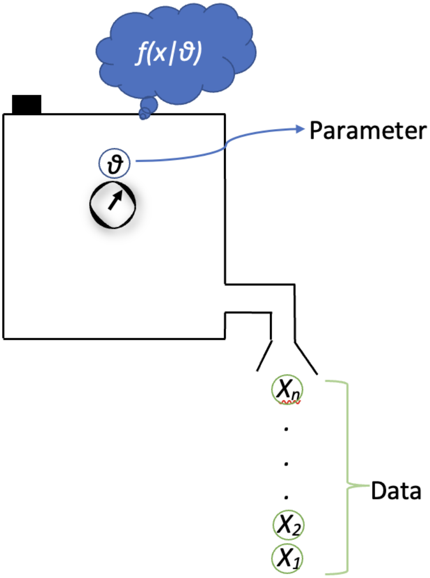
Distribution: random number generation engine



Distribution: random number generation engine



Distribution: random number generation engine



Treating the unknown parameter θ

- Assuming a random sample, the joint distribution of the observed data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is given by:

$$f(\mathbf{x}|\theta) = f(x_1, x_2, \ldots, x_n|\theta) = L(\theta)$$

which is known as likelihood.

Treating the unknown parameter θ

- Assuming a random sample, the joint distribution of the observed data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is given by:

$$f(\mathbf{x}|\theta) = f(x_1, x_2, \ldots, x_n|\theta) = L(\theta)$$

  which is known as likelihood.

- The likelihood is a function of the parameter θ for fixed $\mathbf{x}$ (observed data). It is considered to capture all the information about θ that is available in the data $\mathbf{x}$.

Treating the unknown parameter θ

- Assuming a random sample, the joint distribution of the observed data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is given by:

$$f(\mathbf{x}|\theta) = f(x_1, x_2, \ldots, x_n|\theta) = L(\theta)$$

  which is known as likelihood.

- The likelihood is a function of the parameter θ for fixed $\mathbf{x}$ (observed data). It is considered to capture all the information about θ that is available in the data $\mathbf{x}$.

- There are two (mainstream) schools of thought in statistics regarding the way we deal with the unknown parameter θ:

Treating the unknown parameter θ

- Assuming a random sample, the joint distribution of the observed data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is given by:

$$f(\mathbf{x}|\theta) = f(x_1, x_2, \ldots, x_n|\theta) = L(\theta)$$

  which is known as likelihood.

- The likelihood is a function of the parameter θ for fixed $\mathbf{x}$ (observed data). It is considered to capture all the information about θ that is available in the data $\mathbf{x}$.

- There are two (mainstream) schools of thought in statistics regarding the way we deal with the unknown parameter θ:

  (1) **Frequentist** (Neyman - Pearson - Wald)

Treating the unknown parameter θ

- Assuming a random sample, the joint distribution of the observed data $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is given by:

$$f(\mathbf{x}|\theta) = f(x_1, x_2, \ldots, x_n|\theta) = L(\theta)$$

which is known as likelihood.

- The likelihood is a function of the parameter θ for fixed $\mathbf{x}$ (observed data). It is considered to capture all the information about θ that is available in the data $\mathbf{x}$.

- There are two (mainstream) schools of thought in statistics regarding the way we deal with the unknown parameter θ:

  (1) **Frequentist** (Neyman - Pearson - Wald)

  (2) **Bayesian** (Bayes)

Frequentist School of thought in Statistics

- Within this school we adopt the frequency based interpretation of probability.

Frequentist School of thought in Statistics

- Within this school we adopt the frequency based interpretation of probability.
- The frequentist approach attempts to be "**objective**" in setting the probabilities. This, heavily relies on the assumption that we are capable to repeat an experiment, infinite number of times, under "**identical**" conditions.

Frequentist School of thought in Statistics

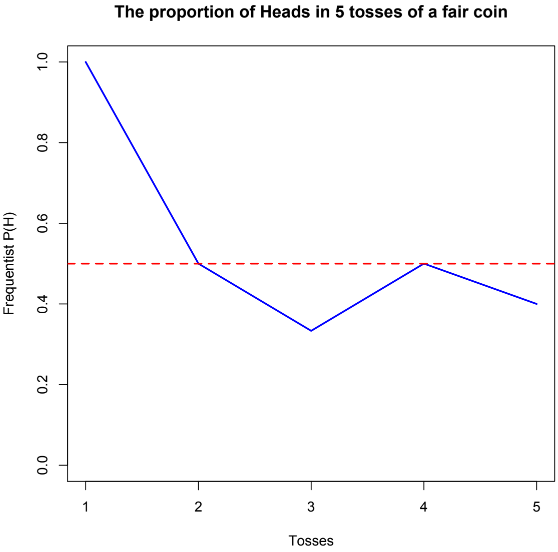
- Within this school we adopt the frequency based interpretation of probability.
- The frequentist approach attempts to be "**objective**" in setting the probabilities. This, heavily relies on the assumption that we are capable to repeat an experiment, infinite number of times, under "**identical**" conditions.
- **Frequentist probability** of an event is defined as the limit of its relative frequency in a large number of trials.

Frequentist School of thought in Statistics

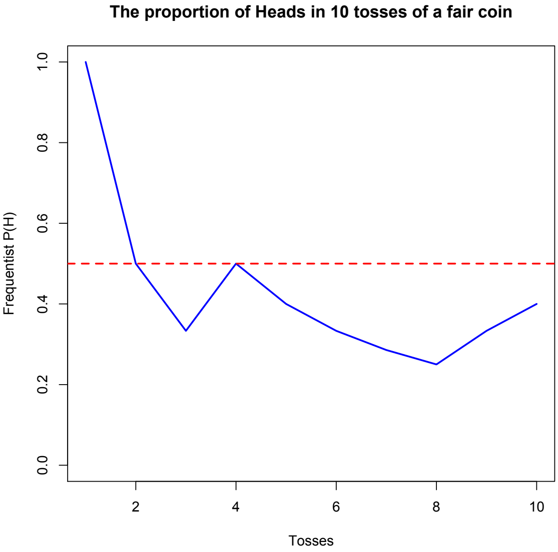
- Within this school we adopt the frequency based interpretation of probability.
- The frequentist approach attempts to be "**objective**" in setting the probabilities. This, heavily relies on the assumption that we are capable to repeat an experiment, infinite number of times, under "**identical**" conditions.
- **Frequentist probability** of an event is defined as the limit of its relative frequency in a large number of trials.
- **Example:** In the case of a fair coin toss, the frequentist probability $P(H)$ is $1/2$, not because there are two equally likely outcomes (classical interpretation of probability) but because in repeated trials the empirical frequency converges to the limit $1/2$ as the number of trials goes to infinity, i.e.

$$P(H) = \frac{\#\text{ of heads}}{\#\text{ of trials}} = \frac{n_H}{n} \xrightarrow{n \to \infty} \frac{1}{2}$$

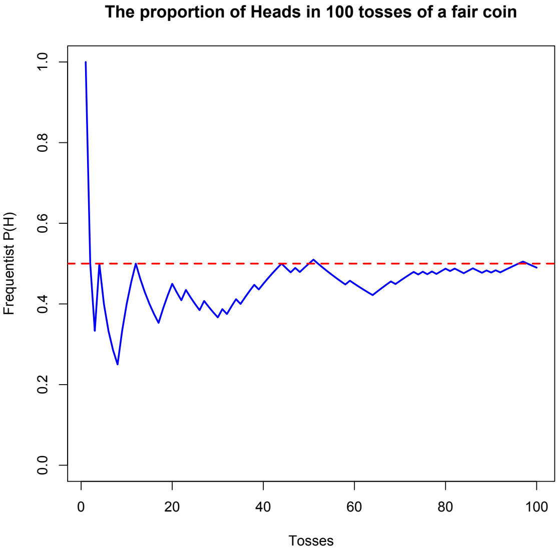
Frequentist School: probability interpretation



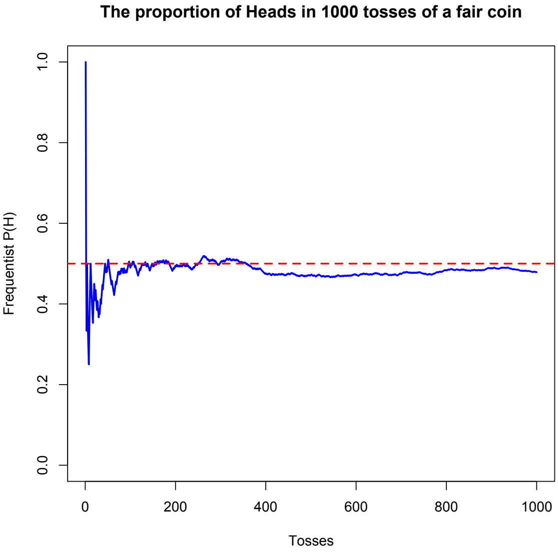
Frequentist School: probability interpretation



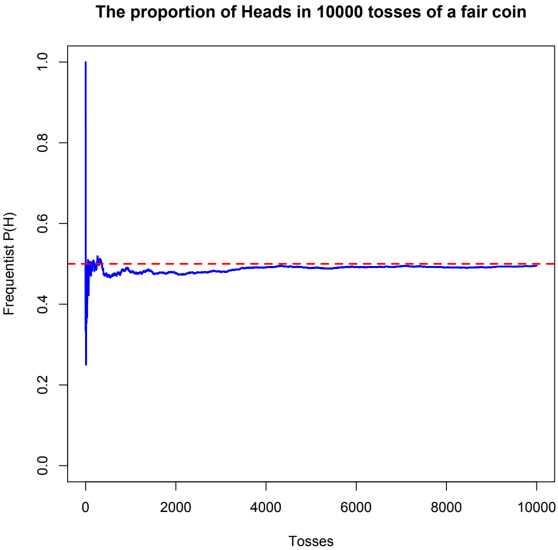
Frequentist School: probability interpretation



Frequentist School: probability interpretation



Frequentist School: probability interpretation



Data and Parameter(s) in Statistics

- The distribution that we chose for our data, comes along with a (univariate or multivariate) set of parameters that fully describe the random mechanism which produces the data. For example:
    - $X|\theta \sim Bernoulli(\theta)$

Data and Parameter(s) in Statistics

- The distribution that we chose for our data, comes along with a (univariate or multivariate) set of parameters that fully describe the random mechanism which produces the data. For example:
  - $X|\theta \sim Bernoulli(\theta)$
  - $X|\left(\theta_1, \theta_2\right) \sim N\left(\theta_1, \theta_2\right)$

Data and Parameter(s) in Statistics

- The distribution that we chose for our data, comes along with a (univariate or multivariate) set of parameters that fully describe the random mechanism which produces the data. For example:
  - $X|\theta \sim Bernoulli(\theta)$
  - $X|\left(\theta_1, \theta_2\right) \sim N\left(\theta_1, \theta_2\right)$
- Usually we are interested in drawing either:

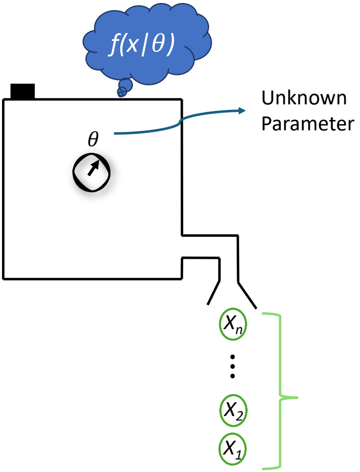
Data and Parameter(s) in Statistics

- The distribution that we chose for our data, comes along with a (univariate or multivariate) set of parameters that fully describe the random mechanism which produces the data. For example:
  - $X|\theta \sim Bernoulli(\theta)$
  - $X|\left(\theta_1, \theta_2\right) \sim N\left(\theta_1, \theta_2\right)$
- Usually we are interested in drawing either:
  - **inference** for the unknown parameter(s) θ, i.e., derive point/interval estimates, hypothesis testing and/or

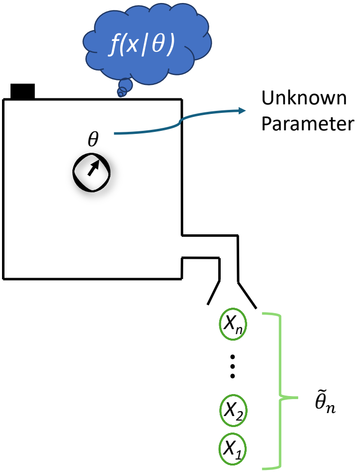
Data and Parameter(s) in Statistics

- The distribution that we chose for our data, comes along with a (univariate or multivariate) set of parameters that fully describe the random mechanism which produces the data. For example:
  - $X|\theta \sim \mathit{Bernoulli}(\theta)$
  - $X|\left(\theta_1, \theta_2\right) \sim N\left(\theta_1, \theta_2\right)$
- Usually we are interested in drawing either:
  - **inference** for the unknown parameter(s) θ, i.e., derive point/interval estimates, hypothesis testing and/or
  - **predictions** for future observable(s).

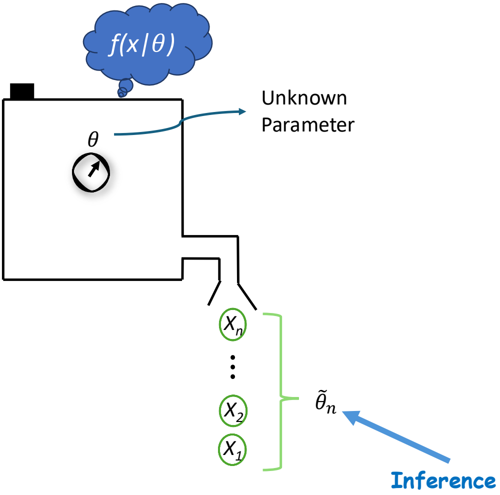
Frequentist School of thought in Statistics



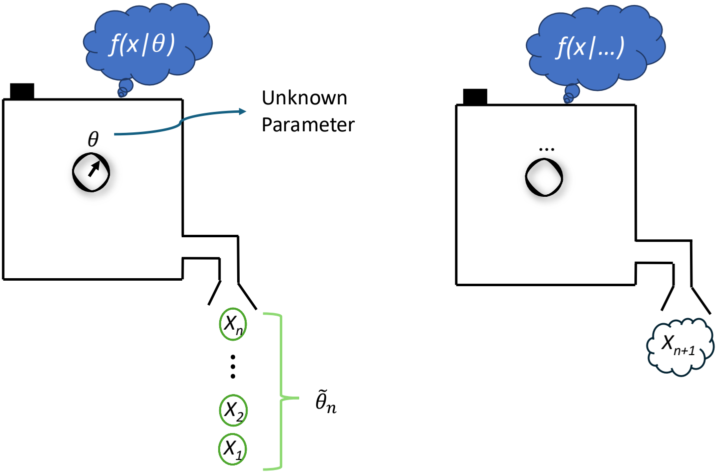
Frequentist School of thought in Statistics



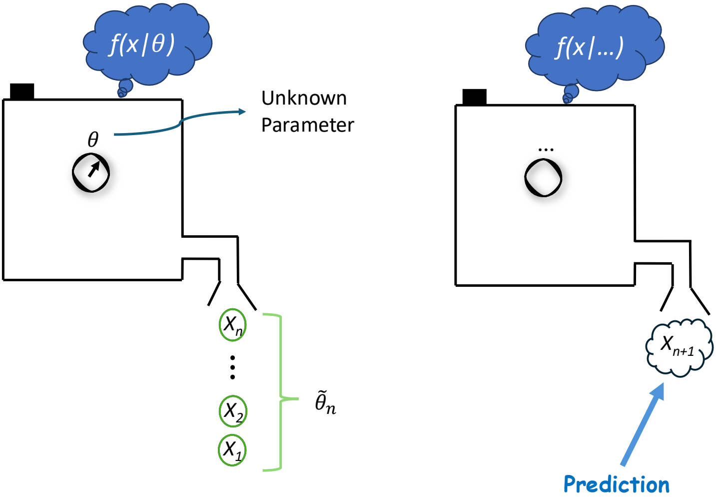
Frequentist School of thought in Statistics



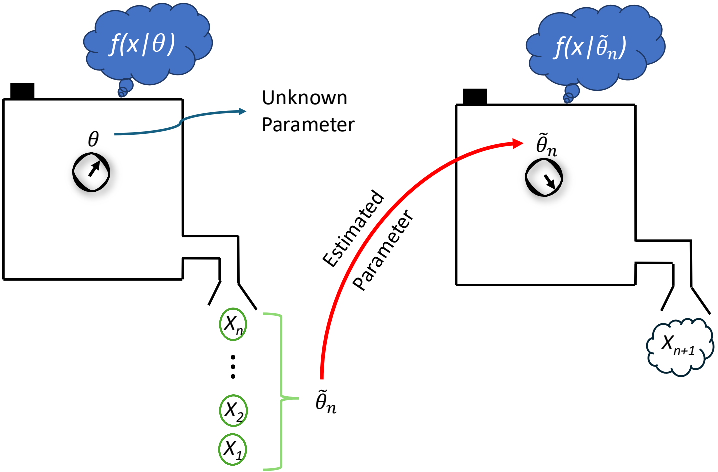
Frequentist School of thought in Statistics



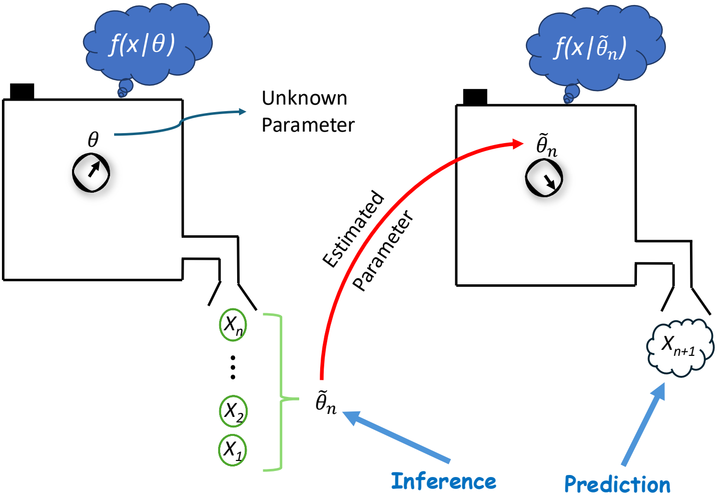
Frequentist School of thought in Statistics



Frequentist School of thought in Statistics



Frequentist School of thought in Statistics



Bayesian School of thought in Statistics

- The data $\mathbf{x}$ become available via the random mechanism $f(\mathbf{x}|\theta)$, where θ is an unknown constant.

Bayesian School of thought in Statistics

- The data $\mathbf{x}$ become available via the random mechanism $f(\mathbf{x}|\theta)$, where θ is an unknown constant.
- When something is unknown to you, i.e. θ, a reasonable thing to do is to use probability theory to quantify your uncertainty.

Bayesian School of thought in Statistics

- The data $\mathbf{x}$ become available via the random mechanism $f(\mathbf{x}|\theta)$, where θ is an unknown constant.
- When something is unknown to you, i.e. θ, a reasonable thing to do is to use probability theory to quantify your uncertainty.
- So in the Bayesian school the parameter θ is considered to be a random variable.

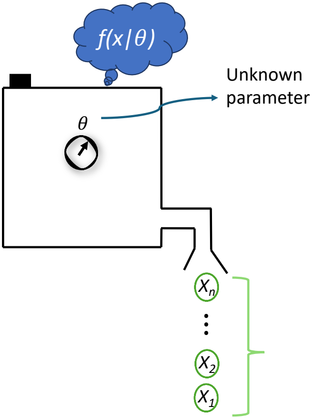
Bayesian School of thought in Statistics

- The data $\mathbf{x}$ become available via the random mechanism $f(\mathbf{x}|\theta)$, where θ is an unknown constant.
- When something is unknown to you, i.e. θ, a reasonable thing to do is to use probability theory to quantify your uncertainty.
- So in the Bayesian school the parameter θ is considered to be a random variable.
- Its distribution, will quantify our (**subjective**) opinion regarding θ (before looking the data) with a prior distribution: $\pi(\theta)$.

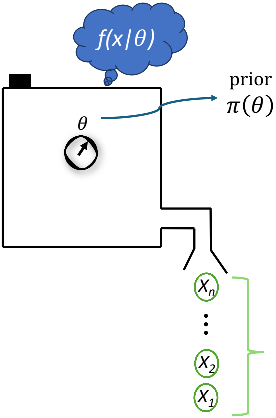
Bayesian School of thought in Statistics

- The data $\mathbf{x}$ become available via the random mechanism $f(\mathbf{x}|\theta)$, where θ is an unknown constant.
- When something is unknown to you, i.e. θ, a reasonable thing to do is to use probability theory to quantify your uncertainty.
- So in the Bayesian school the parameter θ is considered to be a random variable.
- Its distribution, will quantify our (**subjective**) opinion regarding θ (before looking the data) with a prior distribution: $\pi(\theta)$.
- Then Bayes theorem will do the magic **updating** the prior distribution to posterior, in the light of the data.

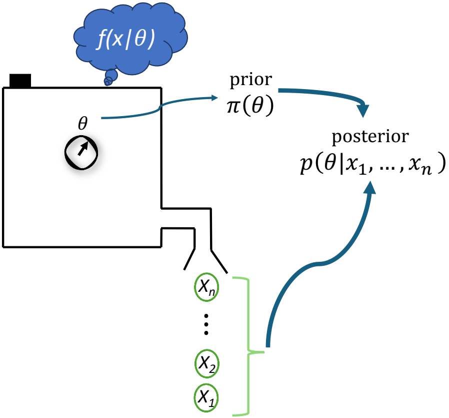
Bayesian School of thought in Statistics



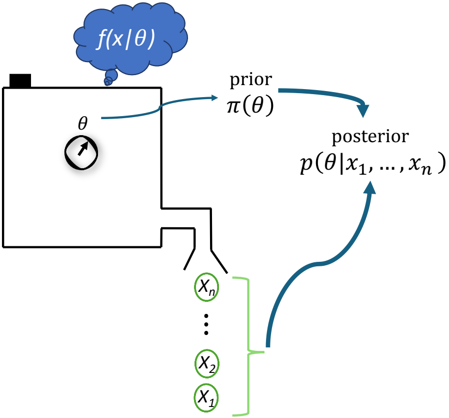
Bayesian School of thought in Statistics

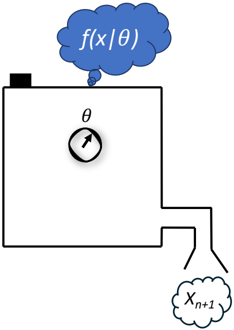


Bayesian School of thought in Statistics

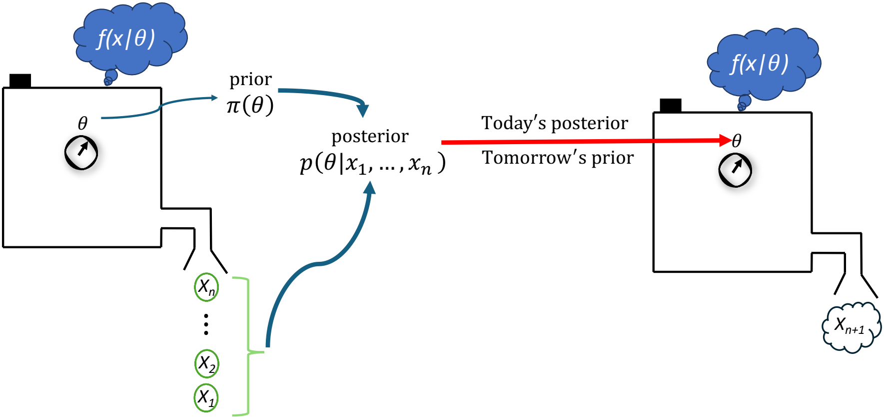


Bayesian School of thought in Statistics

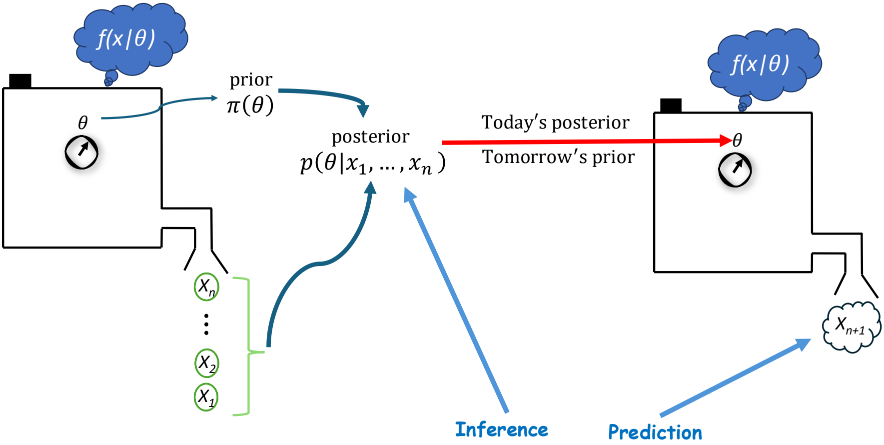




Bayesian School of thought in Statistics



Bayesian School of thought in Statistics



Bayesian School of thought in Statistics

- The Bayes theorem for distributions is given by:

$$p(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)\pi(\theta)}{\int f(\mathbf{x}|\theta)\pi(\theta)d\theta}$$

where:

  - $\mathbf{x} = \{x_1, \ldots, x_n\}$ are the observed data.
  - $f(\mathbf{x}|\theta)$ refers to the likelihood of the data
  - $\pi(\theta)$ is the prior distribution of the parameter θ,
  - $p(\theta|\mathbf{x})$ is the posterior distribution of the parameter θ, given the observed data $\mathbf{x} = \{x_1, \ldots, x_n\}$.

Bayesian School of thought in Statistics

- The Bayes theorem for distributions is given by:

$$p(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)\pi(\theta)}{\int f(\mathbf{x}|\theta)\pi(\theta)d\theta}$$

where:

- $\mathbf{x} = \{x_1, \ldots, x_n\}$ are the observed data.
- $f(\mathbf{x}|\theta)$ refers to the likelihood of the data
- $\pi(\theta)$ is the prior distribution of the parameter θ,
- $p(\theta|\mathbf{x})$ is the posterior distribution of the parameter θ, given the observed data $\mathbf{x} = \{x_1, \ldots, x_n\}$.

- So the Bayes theorem is nothing more that an **updating mechanism**, where the prior is updated to posterior in the light of evidence coming from the available data.

Bayesian School of thought in Statistics

The Bayesian approach consists of the following steps:

Bayesian School of thought in Statistics

The Bayesian approach consists of the following steps:

(a) Define the likelihood: $f(\mathbf{x}|\theta)$.

Bayesian School of thought in Statistics

The Bayesian approach consists of the following steps:

(a) Define the likelihood: $f(\mathbf{x}|\theta)$.

(b) Define the prior distribution: $\pi(\theta)$.

Bayesian School of thought in Statistics

The Bayesian approach consists of the following steps:

(a) Define the likelihood: $f(\mathbf{x}|\theta)$.

(b) Define the prior distribution: $\pi(\theta)$.

(c) Compute the posterior distribution: $p(\theta|\mathbf{x})$.

Bayesian School of thought in Statistics

The Bayesian approach consists of the following steps:

(a) Define the likelihood: $f(\mathbf{x}|\theta)$.

(b) Define the prior distribution: $\pi(\theta)$.

(c) Compute the posterior distribution: $p(\theta|\mathbf{x})$.

(d) Perform decision making: Draw inference regarding θ (point/interval estimates and hypothesis testing).

Bayesian School of thought in Statistics

The Bayesian approach consists of the following steps:

(a) Define the likelihood: $f(\mathbf{x}|\theta)$.

(b) Define the prior distribution: $\pi(\theta)$.

(c) Compute the posterior distribution: $p(\theta|\mathbf{x})$.

(d) Perform decision making: Draw inference regarding θ (point/interval estimates and hypothesis testing).

(e) Derive the predictive distribution $f(x_{n+1}|x_1, x_2, \ldots, x_n)$ of a future observable and make predictions.

Internal & External Bayesian Quality Control

Bayesian Statistical Process Control/Monitoring (SPC/M)

Introduction to SPC/M

- **What is quality?**
  In an abstract way, the quality of a process is **inversely** related to the **variability** of the process.

Introduction to SPC/M

- **What is quality?**
  In an abstract way, the quality of a process is **inversely** related to the **variability** of the process.

- Second law of thermodynamics guarantees that as time passes, the entropy (variability) will increase, i.e. the quality will downgrade.

Introduction to SPC/M

- **What is quality?**
  In an abstract way, the quality of a process is **inversely** related to the **variability** of the process.

- Second law of thermodynamics guarantees that as time passes, the entropy (variability) will increase, i.e. the quality will downgrade.

- Thus, we need methods and tools to routinely monitor the process and **inform us asap** of whether the process's quality is in jeopardy, while maintaining a **low false alarm rate**.

Introduction to SPC/M

- **What is quality?**
  In an abstract way, the quality of a process is **inversely** related to the **variability** of the process.

- Second law of thermodynamics guarantees that as time passes, the entropy (variability) will increase, i.e. the quality will downgrade.

- Thus, we need methods and tools to routinely monitor the process and **inform us asap** of whether the process's quality is in jeopardy, while maintaining a **low false alarm rate**.

- **What is Statistical Process Control & Monitoring (SPC/M)?**
  SPC/M is a method of internal/external quality control, which uses a statistical approach (**control charts**) to monitor and control a process.

Why Statistical Process Control & Monitoring?

- SPC/M aims to detect as soon as possible when a process moves from the *In Control* (**IC**) to the *Out of Control* (**OOC**) state, while we maintain a **low** false alarm rate.

Why Statistical Process Control & Monitoring?

- SPC/M aims to detect as soon as possible when a process moves from the *In Control* (**IC**) to the *Out of Control* (**OOC**) state, while we maintain a **low** false alarm rate.
- The OOC states, can refer to anything that "breaks" the statistical stability of the process and typically this is translated as a parameter shift, which we wish to detect.

Why Statistical Process Control & Monitoring?

- SPC/M aims to detect as soon as possible when a process moves from the *In Control* (**IC**) to the *Out of Control* (**OOC**) state, while we maintain a **low** false alarm rate.
- The OOC states, can refer to anything that "breaks" the statistical stability of the process and typically this is translated as a parameter shift, which we wish to detect.
- We most often focus our attention to detect shifts that are either:

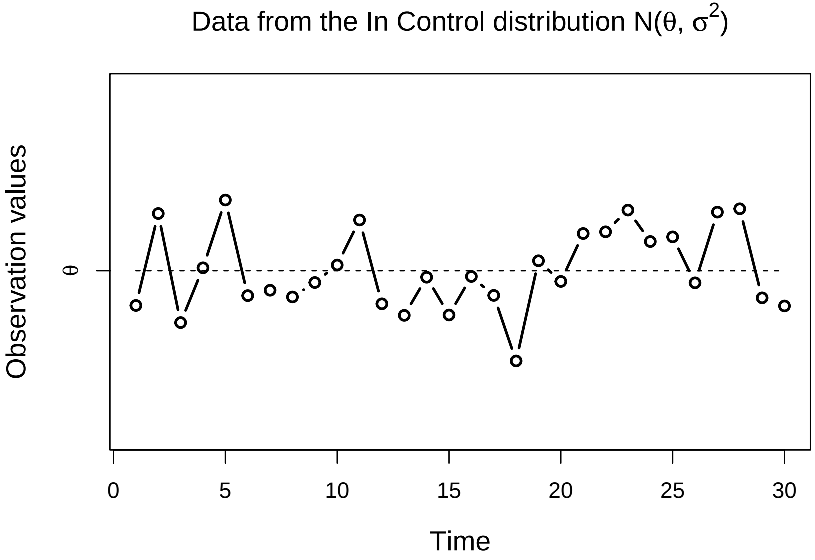
Why Statistical Process Control & Monitoring?

- SPC/M aims to detect as soon as possible when a process moves from the *In Control* (**IC**) to the *Out of Control* (**OOC**) state, while we maintain a **low** false alarm rate.
- The OOC states, can refer to anything that "breaks" the statistical stability of the process and typically this is translated as a parameter shift, which we wish to detect.
- We most often focus our attention to detect shifts that are either:
  - **Transient**, a.k.a. outlier, where we have a *temporal* large shift or

Why Statistical Process Control & Monitoring?

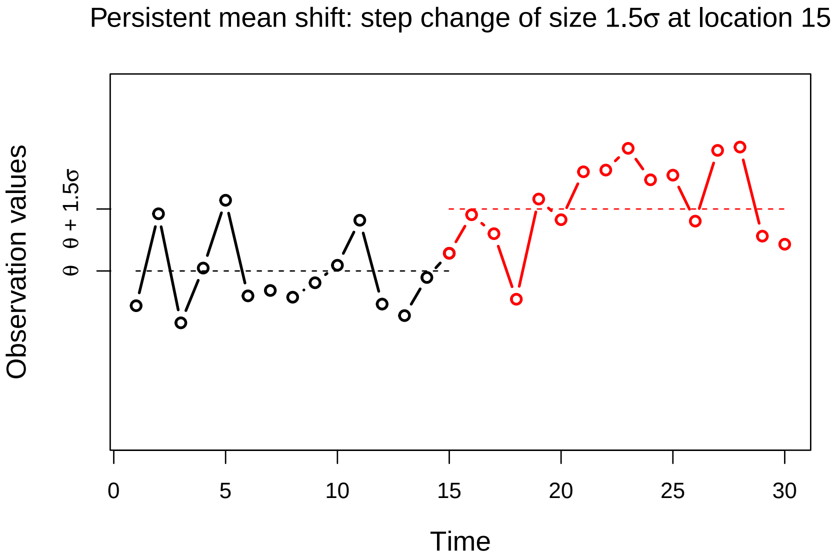
- SPC/M aims to detect as soon as possible when a process moves from the *In Control* (**IC**) to the *Out of Control* (**OOC**) state, while we maintain a **low** false alarm rate.
- The OOC states, can refer to anything that "breaks" the statistical stability of the process and typically this is translated as a parameter shift, which we wish to detect.
- We most often focus our attention to detect shifts that are either:
  - **Transient**, a.k.a. outlier, where we have a *temporal* large shift or
  - **Persistent**, where we have a *permanent* structural change (e.g. change point) shift that is of medium/small size.

Why Statistical Process Control & Monitoring?

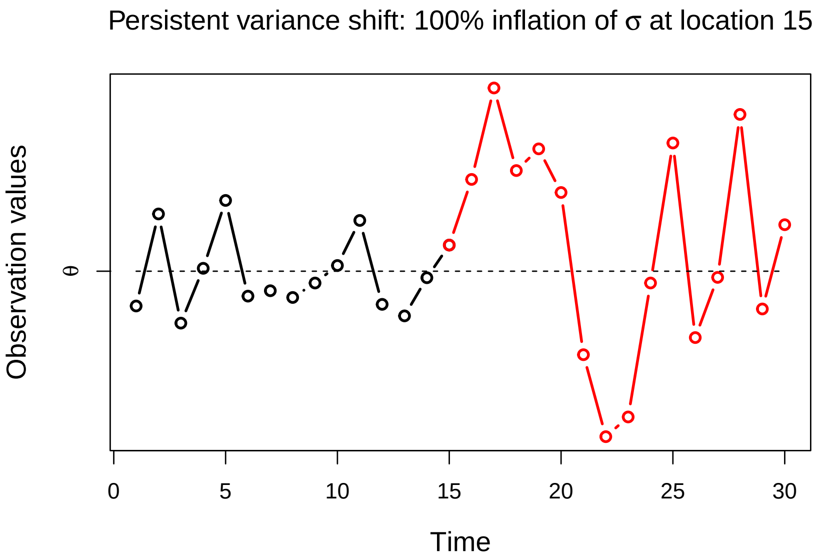


Why Statistical Process Control & Monitoring?

Why Statistical Process Control & Monitoring?



Why Statistical Process Control & Monitoring?



Risks in using the frequentist's approach

- In the frequentist based SPC/M the data come from $f(\boldsymbol{X}|\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ presents the unknown process parameter(s), considered a **constant**.

Risks in using the frequentist's approach

- In the frequentist based SPC/M the data come from $f(\boldsymbol{X}|\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ presents the unknown process parameter(s), considered a **constant**.
- Parametric control chart methods, like 1_{3s}, CUSUM and EWMA, will require the knowledge of the in control parameter(s) value(s) $\boldsymbol{\theta}$.

Risks in using the frequentist's approach

- In the frequentist based SPC/M the data come from $f(\boldsymbol{X}|\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ presents the unknown process parameter(s), considered a **constant**.
- Parametric control chart methods, like 1_{3s}, CUSUM and EWMA, will require the knowledge of the in control parameter(s) value(s) $\boldsymbol{\theta}$.
- In practice, this is achieved employing an **offline** calibration (**phase I**) period, where we derive an estimate of $\boldsymbol{\theta}$ (call it $\tilde{\boldsymbol{\theta}}$), that will be plugged into the likelihood: $f(\boldsymbol{X}|\tilde{\boldsymbol{\theta}})$, to construct the control chart and move to **online** control/monitoring of the process (**phase II**).

Risks in using the frequentist's approach

- In the frequentist based SPC/M the data come from $f(\boldsymbol{X}|\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ presents the unknown process parameter(s), considered a **constant**.
- Parametric control chart methods, like 1_{3s}, CUSUM and EWMA, will require the knowledge of the in control parameter(s) value(s) $\boldsymbol{\theta}$.
- In practice, this is achieved employing an **offline** calibration (**phase I**) period, where we derive an estimate of $\boldsymbol{\theta}$ (call it $\tilde{\boldsymbol{\theta}}$), that will be plugged into the likelihood: $f(\boldsymbol{X}|\tilde{\boldsymbol{\theta}})$, to construct the control chart and move to **online** control/monitoring of the process (**phase II**).

Risks in using the frequentist's approach

- In the frequentist based SPC/M the data come from $f(\boldsymbol{X}|\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ presents the unknown process parameter(s), considered a **constant**.
- Parametric control chart methods, like 1_{3s}, CUSUM and EWMA, will require the knowledge of the in control parameter(s) value(s) $\boldsymbol{\theta}$.
- In practice, this is achieved employing an **offline** calibration (**phase I**) period, where we derive an estimate of $\boldsymbol{\theta}$ (call it $\tilde{\boldsymbol{\theta}}$), that will be plugged into the likelihood: $f(\boldsymbol{X}|\tilde{\boldsymbol{\theta}})$, to construct the control chart and move to **online** control/monitoring of the process (**phase II**).
- This approach is reasonable as long as all the data are IC. OOC points during calibration, will result **contaminated** parameter estimates.

Risks in using the frequentist's approach

- In the frequentist based SPC/M the data come from $f(\boldsymbol{X}|\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ presents the unknown process parameter(s), considered a **constant**.
- Parametric control chart methods, like 1_{3s}, CUSUM and EWMA, will require the knowledge of the in control parameter(s) value(s) $\boldsymbol{\theta}$.
- In practice, this is achieved employing an **offline** calibration (**phase I**) period, where we derive an estimate of $\boldsymbol{\theta}$ (call it $\tilde{\boldsymbol{\theta}}$), that will be plugged into the likelihood: $f(\boldsymbol{X}|\tilde{\boldsymbol{\theta}})$, to construct the control chart and move to **online** control/monitoring of the process (**phase II**).
- This approach is reasonable as long as all the data are IC. OOC points during calibration, will result **contaminated** parameter estimates.
- To anticipate the such risks people tend to use simultaneously several control charts and/or several run rules aiming to increase the power of detecting OOC scenarios. However, **there is no free lunch**! The **more** the charts/rules you combine the **higher** the **false alarm rate**.

Why Bayesian SPC/M?

The frequentist phase I/II based approach has certain deficiencies:

Why Bayesian SPC/M?

The frequentist phase I/II based approach has certain deficiencies:

- Phase I assumes iid data from the in control distribution. What if the parameter experiences transient/persistent shifts during phase I?

Why Bayesian SPC/M?

The frequentist phase I/II based approach has certain deficiencies:

- Phase I assumes iid data from the in control distribution. What if the parameter experiences transient/persistent shifts during phase I?
- Phase I needs to be long to guarantee reliable estimates. What if we have short runs?

Why Bayesian SPC/M?

The frequentist phase I/II based approach has certain deficiencies:

- Phase I assumes iid data from the in control distribution. What if the parameter experiences transient/persistent shifts during phase I?
- Phase I needs to be long to guarantee reliable estimates. What if we have short runs?
- In phase I we can draw decisions only in an offline fashion. What if we need to have online inference (e.g. biomedical applications).

Why Bayesian SPC/M?

The frequentist phase I/II based approach has certain deficiencies:

- Phase I assumes iid data from the in control distribution. What if the parameter experiences transient/persistent shifts during phase I?
- Phase I needs to be long to guarantee reliable estimates. What if we have short runs?
- In phase I we can draw decisions only in an offline fashion. What if we need to have online inference (e.g. biomedical applications).
- Issues encountered in phase I will not only affect testing during that phase. They will greatly influence the phase II chart's performance.

Why Bayesian SPC/M?

The frequentist phase I/II based approach has certain deficiencies:

- Phase I assumes iid data from the in control distribution. What if the parameter experiences transient/persistent shifts during phase I?
- Phase I needs to be long to guarantee reliable estimates. What if we have short runs?
- In phase I we can draw decisions only in an offline fashion. What if we need to have online inference (e.g. biomedical applications).
- Issues encountered in phase I will not only affect testing during that phase. They will greatly influence the phase II chart's performance.
- As more data become available there is no formal mechanism of updating in phase II the (phase I based) parameter estimates.

Why Bayesian SPC/M?

The frequentist phase I/II based approach has certain deficiencies:

- Phase I assumes iid data from the in control distribution. What if the parameter experiences transient/persistent shifts during phase I?
- Phase I needs to be long to guarantee reliable estimates. What if we have short runs?
- In phase I we can draw decisions only in an offline fashion. What if we need to have online inference (e.g. biomedical applications).
- Issues encountered in phase I will not only affect testing during that phase. They will greatly influence the phase II chart's performance.
- As more data become available there is no formal mechanism of updating in phase II the (phase I based) parameter estimates.
- Prior information (typically available) regarding $\boldsymbol{\theta}$ is left unexploited.

Why Bayesian SPC/M?

- Within the Bayesian approach the unknown parameter(s) $\boldsymbol{\theta}$ can be integrated out, deriving the **predictive** distribution.

Why Bayesian SPC/M?

- Within the Bayesian approach the unknown parameter(s) $\boldsymbol{\theta}$ can be integrated out, deriving the **predictive** distribution.
- Use of Bayes theorem will update the (power) prior $\pi(\boldsymbol{\theta})$ to posterior $p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)$ and then for future observable(s) X_{n+1} we get:

Why Bayesian SPC/M?

- Within the Bayesian approach the unknown parameter(s) $\boldsymbol{\theta}$ can be integrated out, deriving the **predictive** distribution.

- Use of Bayes theorem will update the (power) prior $\pi(\boldsymbol{\theta})$ to posterior $p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)$ and then for future observable(s) X_{n+1} we get:

$$f(X_{n+1}|x_1, x_2, \ldots, x_n) = \int f(X_{n+1}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)\, d\boldsymbol{\theta}$$

Why Bayesian SPC/M?

- Within the Bayesian approach the unknown parameter(s) $\boldsymbol{\theta}$ can be integrated out, deriving the **predictive** distribution.

- Use of Bayes theorem will update the (power) prior $\pi(\boldsymbol{\theta})$ to posterior $p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)$ and then for future observable(s) X_{n+1} we get:

$$f(X_{n+1}|x_1, x_2, \ldots, x_n) = \int f(X_{n+1}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)\, d\boldsymbol{\theta}$$

- Based on the predictive distribution we will derive two monitoring schemes:

Why Bayesian SPC/M?

- Within the Bayesian approach the unknown parameter(s) $\boldsymbol{\theta}$ can be integrated out, deriving the **predictive** distribution.

- Use of Bayes theorem will update the (power) prior $\pi(\boldsymbol{\theta})$ to posterior $p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)$ and then for future observable(s) X_{n+1} we get:

$$f(X_{n+1}|x_1, x_2, \ldots, x_n) = \int f(X_{n+1}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)\, d\boldsymbol{\theta}$$

- Based on the predictive distribution we will derive two monitoring schemes:

PCC: Predictive Control Charts, for detecting **transient** shifts of large magnitude (outliers).

Why Bayesian SPC/M?

- Within the Bayesian approach the unknown parameter(s) $\boldsymbol{\theta}$ can be integrated out, deriving the **predictive** distribution.

- Use of Bayes theorem will update the (power) prior $\pi(\boldsymbol{\theta})$ to posterior $p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)$ and then for future observable(s) X_{n+1} we get:

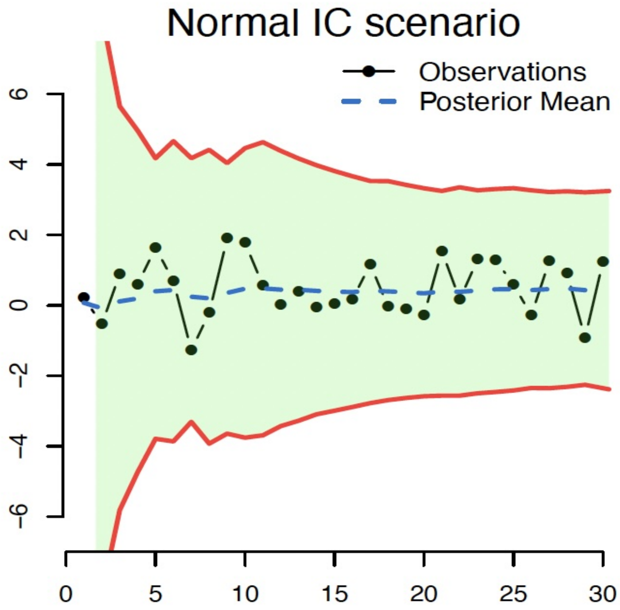
$$f(X_{n+1}|x_1, x_2, \ldots, x_n) = \int f(X_{n+1}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|x_1, x_2, \ldots, x_n)\, d\boldsymbol{\theta}$$

- Based on the predictive distribution we will derive two monitoring schemes:

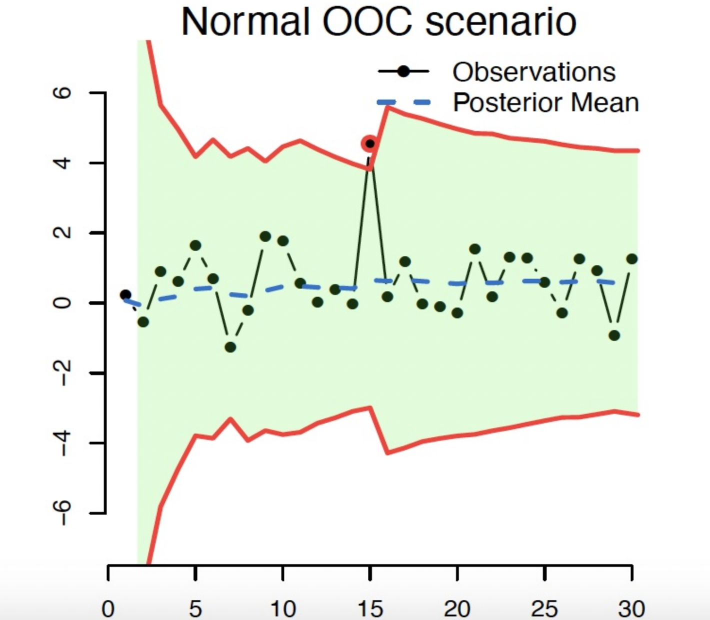
PCC: Predictive Control Charts, for detecting **transient** shifts of large magnitude (outliers).

PRC: Predictive Residual Cusum: for detection of **persistent** shifts of medium/small size (extended to Predictive Ratio Cusum for any distribution in the exponential family).

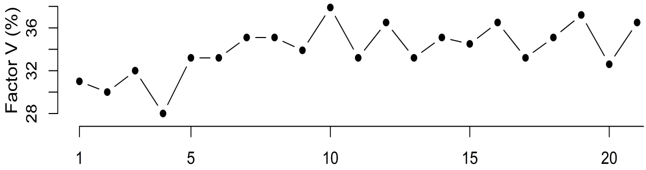
PCC Illustration and Decision Making

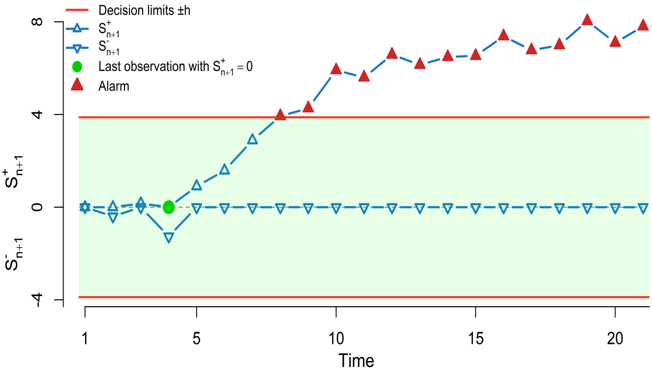


PCC Illustration and Decision Making



PRC Illustration and Decision Making

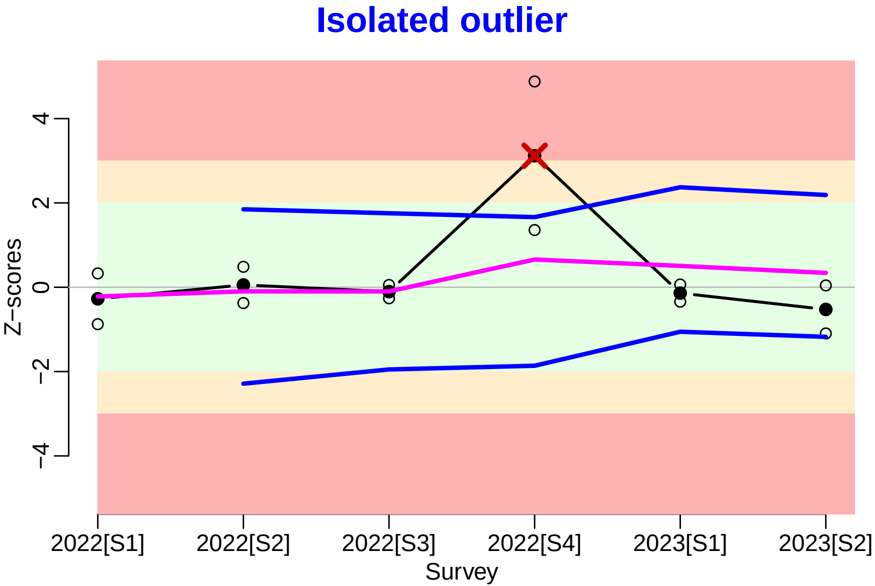




PCC in ECAT's z-score historic evaluation

- ECAT's EQA program evaluates the deviation of the measured result from the assigned value.
- The performance is statistically quantified using Z-scores.
- Univariate and bivariate z-score analysis provides feedback for the current state of the lab.
- In addition, ECAT provides the recent history of the lab's z-scores to help them evaluate their longitudinal performance.

PCC Illustration on ECAT's z-score history



Conclusions

- In Bayesian SPC/M we introduced the **Predictive Control Chart** (**PCC**) and the **Predictive Ratio Cusum** (**PRC**) mechanisms which:
  - can be used in IQC and EQA monitoring processes
  - they utilize available **prior information** and/or **historical data**, boosting the performance.
  - They can identify **outlier** & **change point** problems in the process.
  - They can provide **posterior inference** for the unknown parameter(s) is also available.
  - Both PCC and PRC **outperform** competing alternatives and their are found to be **robust** to various misspecifications.

Conclusions

- In Bayesian SPC/M we introduced the **Predictive Control Chart** (**PCC**) and the **Predictive Ratio Cusum** (**PRC**) mechanisms which:
  - can be used in IQC and EQA monitoring processes
  - they utilize available **prior information** and/or **historical data**, boosting the performance.
  - They can identify **outlier** & **change point** problems in the process.
  - They can provide **posterior inference** for the unknown parameter(s) is also available.
  - Both PCC and PRC **outperform** competing alternatives and their are found to be **robust** to various misspecifications.
- With low volumes of data, standard statistics might be in trouble!

Conclusions

- In Bayesian SPC/M we introduced the **Predictive Control Chart** (**PCC**) and the **Predictive Ratio Cusum** (**PRC**) mechanisms which:
    - can be used in IQC and EQA monitoring processes
    - they utilize available **prior information** and/or **historical data**, boosting the performance.
    - They can identify **outlier** & **change point** problems in the process.
    - They can provide **posterior inference** for the unknown parameter(s) is also available.
    - Both PCC and PRC **outperform** competing alternatives and their are found to be **robust** to various misspecifications.
- With low volumes of data, standard statistics might be in trouble!
- Modern alternatives call for utilizing available prior information, opening widely the door to the **Bayesian** approach!

Acknowledgements

- I am grateful to **Werfen** for their interest in the Bayesian perspective and all their work to carry it to everyday practice in IQC.

Acknowledgements

- I am grateful to **Werfen** for their interest in the Bayesian perspective and all their work to carry it to everyday practice in IQC.
- I would like to thank **Piet Meijer**, @ ECAT, for our collaboration and his vision of introducing state of the art statistical tools in EQA.

Acknowledgements

- I am grateful to **Werfen** for their interest in the Bayesian perspective and all their work to carry it to everyday practice in IQC.
- I would like to thank **Piet Meijer**, @ ECAT, for our collaboration and his vision of introducing state of the art statistical tools in EQA.
- This work would not have been existed without **Frederic Sobas** from Hospices Civils de Lyon, France, who provided an actual problem setting that inspired this work, he supported us with different ways, but most importantly he provided invaluable feedback from using **PCC** & **PRC** at the daily Internal Quality Control routine in the medical labs of Hospices Civils de Lyon.

Bayesian SPC/M: Open Access manuscripts

- Bourazas, K., Kiagias, D., and Tsiamyrtzis, P. (2022). **"Predictive Control Charts (PCC): A Bayesian approach in online monitoring of short runs"**. Journal of Quality Technology, Vol. 54 (4):367–391. https://doi.org/10.1080/00224065.2021.1916413

- Bourazas K., Sobas F. and Tsiamyrtzis, P. (2023). **"Predictive ratio CUSUM (PRC): A Bayesian approach in online change point detection of short runs"**. Journal of Quality Technology, Vol. 55(4):391-403. https://doi.org/10.1080/00224065.2022.2161434

- Bourazas K., Sobas F. and Tsiamyrtzis, P. (2023) **"Design and properties of the predictive ratio cusum (PRC) control charts"**. Journal of Quality Technology, Vol. 55(4):404-421. https://doi.org/10.1080/00224065.2022.2161435

PRC paper won the ASQ's 2024 Brumbaugh Award

BRUMBAUGH AWARD WINNERS

The Brumbaugh Award is presented for the paper making the largest single contribution to the development of industrial application of quality control.

The award has been presented since 1949. See the complete list of award recipients and their winning articles in the table below. While the most recent articles are reserved for journal subscribers, ASQ members have access to past articles. Learn more about ASQ journals.

Could you be the next Brumbaugh Award recipient?

When you publish a technical article with ASQ, it becomes eligible for consideration by the Brumbaugh Award committee. See frequently asked questions about publishing with ASQ.

Award-winning authors and articles

Year	Recipient	Winning Article
2024	Konstantinos Bourazas, Frédéric Sobas, Panagiotis Tsiamyrtzis	Predictive ratio CUSUM (PRC): A Bayesian approach in online change point detection of short runs, Journal of Quality Technology, 55:4, 391-403.

Thomas Bayes

Merci beaucoup!

Merci beaucoup!