

# 1 Introduction

Welcome! This application allows to show the effect of observed data and a priori distribution on posterior distribution. This effect is illustrated with the conjugate normal model. Let us remember that Bayesian perspective quantifies the belief in the occurrence of an event from a concept of subjective probability. For this we establish a probability distribution for the observations of the event  $p(\mathbf{y} | \theta)$ , and a probability distribution for the parameters  $p(\theta)$ , they are considered as random variables. Later, they are used in the Bayes theorem with the purpose of updating the belief and inferring about parameters and future observations through posterior distribution that belongs to a family of known distributions:

$$p(\theta | \mathbf{y}) \propto p(\mathbf{y} | \theta) p(\theta)$$

The application uses the following cases of the normal conjugate model:

## 2 Unknown mean with known variance.

### 2.1 Likelihood distribution

Suppose we have  $\mathbf{y} = (y_1, \dots, y_n)$  a vector of independent observations such that  $\mathbf{y} \stackrel{\text{i.i.d.}}{\sim} N(\theta, \sigma^2)$  with  $\sigma^2$ , therefore, the likelihood distribution is:

$$p(\mathbf{y} | \theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \times \exp\left(\frac{-1}{2\sigma^2}(y_i - \theta)^2\right), \quad (1)$$

### 2.2 Prior distribution

$\theta \sim N(\mu_0, \tau_0^2)$  whose hyperparameters  $\theta$  and  $\tau_0^2$  are known.

$$p(\theta) = \frac{1}{\sqrt{2\pi\tau_0^2}} \times \exp\left(\frac{-1}{2\tau_0^2}(\theta - \mu_0)^2\right), \quad (2)$$

### 2.3 Posterior distribution

$$\theta | \mathbf{y} \sim N(\theta | \mu_n, \tau_n^2)$$

$$p(\theta | \mathbf{y}) \propto \exp\left(\frac{-1}{2\tau_n^2}(\theta - \mu_n)^2\right) \quad (3)$$

where  $\frac{1}{\tau_n^2} = \frac{1}{\sigma^2} + \frac{1}{\tau_0^2}$  and  $\mu_n$  is the new mean of  $\theta$ , composed of the weighted average between the a priori mean and the sample mean with weights sample mean with weights proportional to their respective precisions,  $\mu_n = \frac{\sum_{i=1}^n y_i \tau_n^2 + \mu_0 \sigma^2}{\tau_n^2 + \sigma^2}$

### 2.4 known values and parameters

Each of the values required for the calculation of the distributions is described below:

- Tamaño de muestra (n)
- Varianza ( $\sigma^2$ ): It's the variance of the likelihood distribution

Both values are needed to simulate the data set.

- **Media muestral** ( $\bar{y}_n$ ): the a priori average is estimated based on this value, as follows  $\mu_0 = \bar{y}_n \pm d\sigma$
- **Standard deviation number** ( $d$ ): It indicates how many standard deviations we want the a priori mean of the sample mean. Thus, it can be set in a range of  $\pm 3$  standard deviations with steps of 0.5.
- **Prior standard deviation**( $\tau_0$ )

### 3 Known mean with unknown variance.

#### 3.1 Likelihood distribution

The likelihood distribution for a vector of observations  $\mathbf{y} = (y_1, \dots, y_n)$  is:

$$p(\mathbf{y} | \theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \times \exp\left(\frac{-1}{2\sigma^2}(y_i - \mu)^2\right) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \times \exp\left(\frac{-1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2\right) \quad (4)$$

where  $\mathbf{y} \stackrel{\text{i.i.d.}}{\sim} N(\theta, \sigma^2)$  and  $\nu = \frac{1}{n} \sum_{i=1}^n (y_i - \mu)^2$

#### 3.2 Prior distribution

$\sigma^2 \sim \text{Gamma-inverse}\left(\frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2}\right)$

$$p(\sigma^2) = \frac{\left(\frac{\nu_0\sigma_0^2}{2}\right)^{\frac{\nu_0}{2}}}{\Gamma\left(\frac{\nu_0}{2}\right)} (\sigma^2)^{-\frac{\nu_0}{2+1}} \times \exp\left(\frac{-\nu_0\sigma_0^2}{\sigma^2}\right) \quad (5)$$

#### 3.3 Posterior distribution

$\sigma^2 | \mathbf{y} \sim \text{Gamma-inverse}\left(\frac{\nu_0+n}{2}, \frac{n\nu+\nu_0\sigma^2}{2}\right)$

$$p(\sigma^2 | \mathbf{y}) \propto (\sigma^2)^{-\frac{n}{2}} e^{-\frac{n}{2\sigma^2}\nu} (\sigma^2)^{-\frac{\nu_0}{2+1}} e^{-\frac{\nu_0\sigma_0^2}{\sigma^2}} \quad (6)$$

#### 3.4 known values and parameters

Each of the values required for the calculation of the distributions is described below:

- Sample size (n)
- **Media** ( $\mu$ ): It's the population mean.

Both values are needed to simulate the data set.

- $\nu$ : It corresponds to the squared sums of the differences between the observed values and the known mean. It takes values from 0.1 to 5, with step of 0.1 and if this value is very small the prior and posterior distribution will be centered on the mean.
- **First hyperparameter** ( $\alpha$ ):  $\frac{\nu_0}{2}$ . It is the scaling parameter of the Inverse Gamma.
- **Second hyperparameter** ( $\beta$ ):  $\frac{\nu_0\sigma_0^2}{2}$ . It is the shape parameter of the Inverse Gamma

### 4 Mean and unknown variance with a priori distribution of the mean depending on the variance.

#### 4.1 Likelihood distribution

This case refers to a multiparametric model. Let  $\mathbf{y} = (y_1, \dots, y_n)$  a vector of independent observations that distributes  $\mathbf{y} \stackrel{\text{i.i.d.}}{\sim} N(\theta, \sigma^2)$  with unknown  $\theta$  and  $\sigma^2$

$$p(\mathbf{y} | \theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \times \exp\left(\frac{-1}{2\sigma^2} \sum_{i=1}^n (n-1)S^2 + n(\bar{y} - \theta)^2\right), \quad (7)$$

## 4.2 Prior distribution

The prior distribution for  $\theta$  is assumed to depend on  $\sigma^2$ ,  $p(\theta | \sigma^2)$ , while the a priori distribution for  $\sigma^2$  doesn't depend on  $\theta$  and it can be written as  $p(\sigma^2)$ . Therefore, this case has two distributions a priori:

$$\theta | \sigma^2 \sim N(\mu_0, \frac{\sigma^2}{\kappa_0})$$

$\sigma^2 \sim \text{Gamma-inverse}(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2})$  where  $\kappa$  represents the a priori belief we have about the parameter  $\theta$ .

## 4.3 Posterior distribution

The posterior marginal distribution of  $\theta$  is:

$$p(\sigma^2 | \mathbf{y}) \sigma^2 | \mathbf{y} \sim \text{Gamma-inverse}(\frac{\nu_n}{2}, \frac{\nu_n \sigma_n^2}{2}) \text{ where } \nu_n = \nu_0 + n, \sigma_n^2 = \frac{\nu_0 \sigma_0^2 + (n-1)S^2 + \frac{n\kappa_0(\bar{y}_n - \mu_0)^2}{n + \kappa_0}}{n + \nu_0}$$

$$\text{The marginal posterior distribution of } \theta \text{ is } \theta | \mathbf{y} \sim t_{n+\nu_0}(\mu_0, \frac{\sigma_n^2}{\kappa_0 + n}) \text{ where } \mu_n = \frac{\mu_0 \kappa_0 + n\bar{y}_n}{n + \kappa_0}$$

## 4.4 known values and parameters

Each of the values required for the calculation of the distributions is described below:

- Sample size (n)

Both values are needed to simulate the data set.

- $\mu_0$ : represents the mean of the prior distribution of conditional on variance  $\sigma^2$  and is a free parameter.
- $\kappa_0$ : It is the a priori belief that we have about the parameter  $\sigma$ , this parameter indicates how much uncertainty we have about the parameter  $\theta$ , if there is a lot of uncertainty  $\kappa_0$  takes small values, but if there is enough knowledge about  $\theta$ ,  $\kappa_0$  takes large values. It should be noted that  $\kappa_0$  takes only positive values.
- $\nu_0$ : It is the value that allows obtaining the parameters of the prior distribution of  $\sigma^2$ , that is, they make part of the parameters  $\alpha$  and  $\beta$  of the Gamma Inverse. This value depends on the knowledge about the parameter  $\sigma^2$  and takes positive values.
- $\sigma_0^2$ : It value that allows obtaining the parameter  $\beta$  of the inverse gamma distribution, a priori for  $\sigma^2$ . This value depends on the knowledge about the parameter  $\sigma^2$  and must be a positive value.