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($\mathbf{y} \mid \theta$), and a probability distribution for the parameters p(θ), 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($\theta \mid \mathbf{y}$) $\propto$ p($\mathbf{y} \mid \theta$) p(θ)

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

2 Unknown mean with known variance.

2.1 Likelihood distribution

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

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

2.2 Prior distribution

$\theta \sim \mathrm{N}(\mu_0, \tau_0^2)$ whose hyperparameters θ and τ_0^2 are known.

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

2.3 Posterior distribution

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

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

where $\frac{1}{\tau_n^2} = \frac{1}{\sigma^2} + \frac{1}{\tau_0^2}$ and μ_n is the new mean of θ, 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, $\mu_n = \frac{\sum_{i=1}^{n} y_i \tau_n^2 + \mu_n \sigma^2}{\tau_n^2 + \sigma^2}$

2.4 known values and parameters

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

- **Sample size** (n)
- **Variance (σ^2):** It's the population variance.

Both values are needed to simulate the data set.

- **Sample mean ($\bar{y_n}$)**: the priori mean 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 ± 3 standard deviations with steps of 0.5.
- **Prior standard deviation(τ_0)**

3 Known mean with unknown variance.

3.1 Likelihood distribution

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

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

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

3.2 Prior distribution

$\sigma^2 \sim Gamma - inverse(\frac{v_0}{2}, \frac{v_0\sigma_0^2}{2})$

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

3.3 Posterior distribution

$\sigma^2 \mid \mathbf{y} \sim Gamma - inverse(\frac{\nu_0+n}{2}, \frac{n\nu+v_0\sigma^2}{2})$

$$p(\sigma^2 \mid \mathbf{y}) \propto (\sigma^2)^{-\frac{n}{2}} \times \exp\left(\frac{-n}{2\sigma^2}v\right)(\sigma^2)^{-\frac{v_0}{2+1}} \times \exp\left(\frac{-\nu_0\sigma_0^2}{2\sigma^2}\right) \tag{6}$$

3.4 known values and parameters

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

- **Sample size** (n)
- **Media (μ)**: It's the population mean.
- **Sample variance (S^2)**: Variability we want the likelihood distribution to have.

Both values are needed to simulate the data set.

- $\boldsymbol{\nu}$: 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.
- **First hyperparameter (α)**: $\frac{\nu_0}{2}$. It is the scaling parameter of the prior distribution.
- **Second hyperparameter (β)**: $\frac{\nu_0\sigma_0^2}{2}$. It is the shape parameter of the prior distribution.

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 $\mathbf{y} = (y_1, \ldots, y_n)$ a vector of independent observations that distributes $\mathbf{y} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(\theta, \sigma^2)$ with unknown θ and σ^2

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

where $S^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}$

4.2 Prior distribution

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

- $\theta \mid \sigma^2 \sim \mathrm{N}(\mu_0, \frac{\sigma^2}{\kappa_0})$
- $\sigma^2 \sim$ Gamma- inverse$(\frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2})$

4.3 Posterior distribution

The posterior marginal distribution of θ is:

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

And the marginal posterior distribution of θ is:

$\theta \mid \mathbf{y} \sim t_{n+\nu_0}(\mu_n, \frac{\sigma_n^2}{\kappa_0+n})$, 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)
- **Sample mean ($\bar{y_n}$)**:
- **Sample variance (S^2)**: Variability we want the likelihood distribution to have.

Both values are needed to simulate the data set.

- $\boldsymbol{\mu_0}$: represents the mean of the a priori distribution conditional on the variance σ^2 and is a free parameter.
- $\boldsymbol{\kappa_0}$: It is the a priori belief that we have about the parameter σ, this parameter indicates how much uncertainty we have about the parameter θ, if there is a lot of uncertainty κ_0 takes small values, but if there is enough knowledge about θ, κ_0 takes large values. It should be noted that κ_0 takes only positive values.
- $\boldsymbol{\alpha_0}$: It's the shape parameter for the prior distribution of sigma distribution. $\alpha_0 = \frac{\nu_0}{2}$
- $\boldsymbol{\beta_0}$: It's the scale parameter for the prior distribution of sigma. $\beta_0 = \frac{\nu_0\sigma_0^2}{2}$