Laboratory Session : May 6, 2025
Exercises due on : May 25, 2025

Exercise 1

- A well established and diffused method for detecting a disease in blood fails to detect the presence of disease in 15% of the patients that actually have the disease. A young UniPD startUp has developed an innovative technique of screening. During the qualification phase, a random sample of $n = 75$ patients known to have the disease is screened using the new method.

(a) Write the probability mass function of getting a false negative y , using the new technique.

(b) Let $n = 75$ be the patient sample, the new method fails to detect the disease in $y = 6$ cases. What is the frequentist estimator for the failure probability of the new method ?

(c) In Bayesian setting, evaluate the posterior probability, assuming a beta distribution with mean value 0.15 and standard deviation 0.14. Plot the posterior distribution for y, and mark on the plot the mean value and variance.

(d) Perform a hypothesis test assuming that if the probability of failing to the detect the desease in ill patients is greater or equal than 15%, the new test is no better that the traditional method. Test the sample at a 5% level of significance in the Bayesian setting.

(e) Perform the same hypothesis test in the classical frequentist setting.

Exercise 2

- A researcher has collected the following $n = 20$ observations that are supposed to be drawn from an unimodal Beta distribution Beta($\alpha = 2, \beta$), where $\alpha = 2$ is fixed and β is the unknown parameter.

$$0.269,\ 0.344,\ 0.802,\ 0.418,\ 0.433,\ 0.835,\ 0.52,\ 0.139,\ 0.243,\ 0.294,\ 0.723,\ 0.493,\ 0.504,\ 0.428,$$

$$0.27,\ 0.53, 0.057,\ 0.585,\ 0.288,\ 0.171$$

- Let the prior be the following piece-wise function:

$$p(\beta) = \begin{cases} 0.2\beta & \text{for } 0.1 \leq \beta < 2 \\ \frac{1}{0.8\sqrt{2\pi}} \exp\left(-\frac{(\beta-2.8)^2}{2(0.8)^2}\right) & \text{for } 2 \leq \beta < 4 \\ 0.1 & \text{for } 4 \leq \beta \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

(a) find the posterior distribution, the posterior mean and standard deviation

(b) find the 95% credibility interval for β

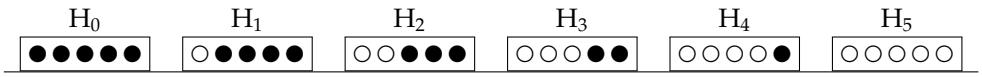
(c) plot the posterior distribution, indicating on the same plot: the mean value, the standard deviation, and the 95% credibility interval

(d) plot (in another chunck but on the same canvas) the prior, the likelihood and the posterior distribution

(optional) repeat the exercise using an uninformative prior

Exercise 3 - Six Boxes Toy Model : inference

- The six boxes toy model is described in reference [1].
- Labeling the boxes as follows:



- write a program in R that:

1) selects a random box
2) makes random sampling from the box
3) prints on the standard output the probability of selecting each box
4) plots the probability for each box as a function of the number of trial

References

[1] G. D'Agostini, *Probability, propensity and probabilities of propensities (and of probabilities)*, `https://arxiv.org/pdf/1612.05292.pdf`
G. D'Agostini, *More lessons form the six box toy experiment*, `https://arxiv.org/pdf/1701.01143.pdf`