Laboratory Session : May 7, 2024
Exercises due on : May 20, 2024

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 method 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) what is the probability distribution of y, the number of times the new method fails to detect the disease ?

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

(c) setup a bayesian computation of 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 test of hypothesis 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 way.

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

Exercise 2

- a researcher has collected $n = 16$ observations that are supposed to come from a `Normal` distribution with known variance $\sigma^2 = 4$:

4.09	4.68	1.87	2.62	5.58	8.68	4.07	4.78
4.79	4.49	5.85	5.09	2.40	6.27	6.30	4.47

- assuming the prior is a step funtion:

$$g(\mu) = \begin{cases} \mu & \text{for } 0 < \mu \leq 3\,, \\ 3 & \text{for } 3 < \mu \leq 5\,, \\ 8 - \mu & \text{for } 5 < \mu \leq 8\,, \\ 0 & \text{for } \mu > 8\,. \end{cases}$$

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

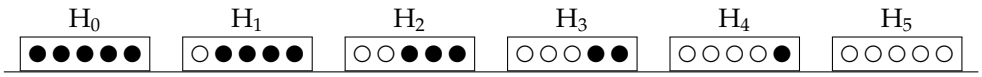
(b) find the 95% credibility interval for μ

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

(d) plot, on the same graph, the prior, the likelihood and the posterior distribution

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`