CS7DS3 Assignment 1

February 26, 2025

To be submitted by **1 p.m.** on **Monday 10th March, 2025**. Submit solutions through blackboard. Legible scans are fine for some questions but please be aware that the Turnitin submission requires a minimum amount of printed text. All plots should be produced using standard software.

Please remember to print your **name** and **student number** on your submission.

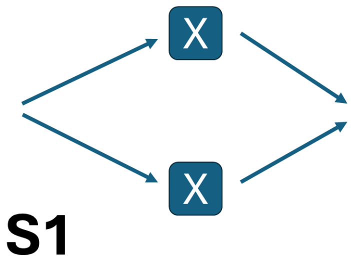
Each part of each question is worth one mark. For example, q.1 is worth one mark, q.2 is worth 3 marks, etc.

Please show your workings. Where relevant, results and definitions, etc. from class can be used, but please reference appropriately (e.g., "as stated in Bayesian Inference 1 slide 6…") Where code has been used, an outline description, along with relevant results, is sufficient.

If you have any questions about the assignment, email me: arwhite@tcd.ie.

Problem set - 15 marks total

A simple electronic system, S_1, consists of two electrical components x_1 and x_2 and is visualised below. For this system, the components are said to be in parallel; that is, the system will fail only if both individual components fail. Assume that the components in the system are independent, and fail individually with a common probability $\theta = \mathbb{P}(X \text{ fails})$.



1) Let $\theta_1 = \mathbb{P}(S_1 \text{ fails})$. Express θ_1 in terms of θ.

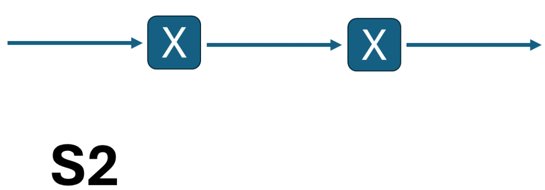
2) Suppose that 1,000 S_1 systems were to be tested. Let A be the total number of observed failures, and assume (for this question only) that $\theta = 0.08$. Compute:

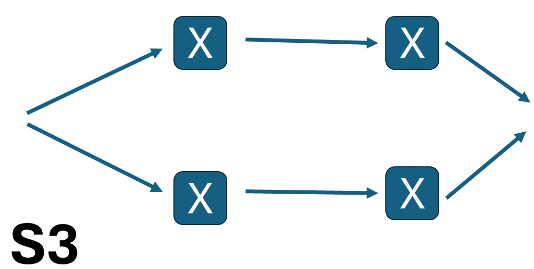
   i) $\mathbb{E}[A]$ and $\mathbb{V}\text{ar}[A]$, the expected number of failures and the variance for number of failures;

   ii) $\mathbb{P}(5 \leq A \leq 10)$;

   iii) the minimum number of failures k^* such that $\mathbb{P}(A \leq k^*) > 0.95$.

3) Two additional systems, S_2 and S_3 are also made using the same kind of components. These systems are shown below. The S_2 system is said to be in series; that is, this system will fail if either (or both) individual components fail. The S_3 system consists of a combination of in series and in parallel subcomponents.





   In terms of θ, what is:

   i) $\mathbb{P}(S_2 \text{ fails}) = \theta_2$?

   ii) $\mathbb{P}(S_3 \text{ fails}) = \theta_3$?

4) An experiment is conducted to test the reliability of individual components. For this experiment, $n = 25$ components are tested, and $k = 3$ failures observed. (That is, data $x = x_1, \ldots, x_n$ are collected, with $n = 25$, and a total $\sum_{i=1}^{n} x_i = k = 3$ failures are observed.) An expert was also consulted regarding the reliability of these components. She expects that the probability of an individual component failure to be somewhere between 5% and 10%. She also says that, in her experience, it is highly unusual to observe failure rates for components of about 20% or higher.

   On the basis of this advice, a Bayesian analysis was proposed to model component failure, and beta distribution prior of $\mathcal{B}e(3, 30)$ was specified.

   i) Describe the key properties of the proposed prior distribution. How well does it match the expert's description, in your opinion?

   ii) Using the expert prior, as well as the data from the outlined experiment, construct a posterior distribution for $p(\theta|x)$ and describe its key features.

   iii) Compare the posterior distribution $p(\theta|x)$ to its prior distribution $p(\theta)$. How informative is this prior, in relation to the data? Explain your reasoning.

   iv) Using Monte Carlo methods or otherwise, use $p(\theta|x)$ to construct posterior distributions for θ_1, θ_2, and θ_3. Briefly comment on the forms of these distributions.

5) A company wishes to purchase 1,000 electronic systems of type S_3, for a total price of €1,000. However, the purchase comes with some conditions. If less than 50 of the purchased systems fail, the company will do nothing, i.e., pay the full price. If 50 or more, but less than 100, systems fail, the company will request reimbursement for those systems only, i..e, pay a price of €1,000 less the number of failed systems. However, should more than 100 systems fail, the company will request a full refund, i.e., they will not pay anything.

   i) What is the expected price for the systems, once the conditions are accounted for?

   ii) Briefly describe how sensitive your results for the expected price are on the expert prior distribution that you used. If you were to instead use a "non-informative" $\mathcal{B}e(1, 1)$ prior for p, how much would this change your expected price estimate?

6) A separate study is conducted, in which the reliability of S_1 systems are tested directly, rather than individual components. A sample of n such systems are tested, with k observed to fail. That is, data $y = y_1, \ldots, y_n$ are collected, and a total $\sum_{i=1}^{n} y_i = k$ failures are observed.

   i) Express the likelihood for the data y in terms of a) θ_1 and b) θ.

ii) Show that $\hat{\theta}$, the maximum likelihood estimate for θ, has the form

$$\hat{\theta} = \sqrt{\frac{k}{n}}.$$

iii) Do you think a Bayesian approach for estimating θ would be difficult in this case? Briefly explain your answer.