STAT 447 Assignment 2

Caden Hewlett

2024-01-17

Question 1 : Define a Bayesian Model

In order to perform inference on the unknown quantities, we must specify how they relate to the data; i.e., we need a probabilistic model. Assume that every Delta 7925H rocket has the same probability p of failing. For simplicity, let us assume that p is allowed to take values on an evenly space grid

$$p \in \left\{ \frac{k}{K} : k \in \{0, \dots, K\} \right\}$$

for some fixed $k \in \mathbb{N}$. Furthermore, we have access to a collection of numbers ρ_k such that:

$$\forall k \in \{0, \dots, K\} : \ \mathbb{P}\left(p = \frac{k}{K}\right) = \rho_k$$

Let Y_i denote a binary variable with $Y_i = 1$ encoding a success, and $Y_i = 0$ a failure. We assume that, conditionally on p, the Y_i's are independent of each other.

We will use the following prior:

$$\rho_k \propto \frac{k}{K}\left(1 - \frac{k}{K}\right)$$

From now on, use $K = 20$

Part 1

What are the unknown quantities in this scenario? And what is the data?

Solution: In this case, we can think of two unknown quantities and one piece of known data. One unknown metric is Y_i, whether or not the rocket is going to fail. The other unknown quantity is the success probability of the rockets - this was denoted p. This is much like the "coin flip" scenario from the previous homework, where Y_i was "whether or not we flip a heads," and the distribution of coin biases (i.e. the probability we pick a coin with bias k/K) was ρ_k.

In this case, the data is given in the setup section, and is simply that as of Jan 2024, Delta 7925H rockets have been launched 3 times, with 0 failed launches. Further, we have ρ_k, the probability that the *true* probability p is k/K for a given discrete k.

Part 2

Here, letting $k = 20$, we can think of the random Variable X as the probability that the rocket has failure rate k, for $k \in [0, 20]$ i.e.

$$X \sim \text{unif}(\{0, 1, 2, \dots, 20\})$$

Similarly, the conditional distribution of Y_i given X can be written as:

$$Y_i \sim \text{bern}\left(\frac{X}{20}\left(1 - \frac{X}{20}\right)\right)$$

Together, these form the joint distribution $p_{X,Y}(x, y)$ as we saw in lecture, due to the chain rule.