
8.1.0 Introduction

In real life, we work with data that are affected by randomness, and we need to extract information and draw conclusions from the data. The randomness might come from a variety of sources. Here are two examples of such situations:

1. Suppose that we would like to predict the outcome of an election. Since we cannot poll the entire population, we will choose a random sample from the population and ask them who they plan to vote for. In this experiment, the randomness comes from the sampling. Note also that if our poll is conducted one month before the election, another source of randomness is that people might change their opinions during the one month period.
2. In a wireless communication system, a message is transmitted from a transmitter to a receiver. However, the receiver receives a corrupted version (a noisy version) of the transmitted signal. The receiver needs to extract the original message from the received noisy version. Here, the randomness comes from the noise.

Examples like these are abundant. Dealing with such situations is the subject of the field of statistical inference.

Statistical inference is a collection of methods that deal with drawing conclusions from data that are prone to random variation.

Clearly, we use our knowledge of probability theory when we work on statistical inference problems. However, the big addition here is that we need to work with real **data**. The probability problems that we have seen in this book so far were clearly defined and the probability models were given to us. For example, you might have seen a problem like this:

Let X be a normal random variable with mean $\mu = 100$ and variance $\sigma^2 = 15$.
Find the probability that $X > 110$.

In real life, we might not know the distribution of X , so we need to collect data, and from the data we should conclude whether X has a normal distribution or not. Now, suppose that we can use the central limit theorem to argue that X is normally distributed. Even in that case, we need to collect data to be able estimate μ and σ .

Here is a general setup for a statistical inference problem: There is an unknown quantity that we would like to estimate. We get some data. From the data, we estimate the desired quantity. There are two major approaches to this problem:

1. **Frequentist (classical) Inference:** In this approach, the unknown quantity θ is assumed to be a fixed quantity. That is, θ is a deterministic (non-random) quantity that is to be estimated by the observed data. For example, in the polling problem stated above we might consider θ as the percentage of people who will vote for a certain candidate, call him/her Candidate A. After asking n randomly chosen voters, we might estimate θ by

$$\hat{\Theta} = \frac{Y}{n},$$

where Y is the number of people (among the randomly chosen voters) who say they will vote for Candidate A. Although θ is assumed to be a non-random quantity, our estimate of θ , which we show by $\hat{\Theta}$ is a random variable, because it depends on our random sample.

2. **Bayesian Inference:** In the Bayesian approach the unknown quantity Θ is assumed to be a random variable, and we assume that we have some initial guess about the distribution of Θ . After observing the data, we update the distribution of Θ using Bayes' Rule.

As an example, consider the communication system in which the information is transmitted in the form of bits, i.e., 0's and 1's. Let's assume that, in each transmission, the transmitter sends a 1 with probability p , or it sends a 0 with probability $1 - p$. Thus, if Θ is the transmitted bit, then $\Theta \sim \text{Bernoulli}(p)$. At the receiver, X , which is a noisy version of Θ , is received. The receiver has to recover Θ from X . Here, to estimate Θ , we use our prior knowledge that $\Theta \sim \text{Bernoulli}(p)$.

In summary, you may say that frequentist (classical) inference deals with estimating non-random quantities, while Bayesian inference deals with estimating random variables. We will discuss frequentist and Bayesian approaches more in detail in this and the next chapter. Nevertheless, it is important to note that both approaches are very useful and widely used in practice. In this chapter, we will focus on frequentist methods, while in the next chapter we will discuss Bayesian methods.
