Project: Beta Distribution

Abstract

In this course, we are going to study several fundamental distributions. These distributions are models of the random variables with some deterministic hyperparameter. A natural question is what if the hyperparameter is also a random variable. In this project, we are going to study the Beta distribution, which is a distribution to model the hyperparameter of a random variable.

The Beta distribution is a family of continuous probability distributions defined on the interval $[0, 1]$ that are parameterized by two positive shape parameters, denoted by α and β. These parameters determine the shape of the distribution. The Beta distribution has been widely used in various fields due to its flexibility to take on different shapes and its ability to model random variables limited to the interval of $[0, 1]$.

Here are two applications of the Beta distribution:

1. **Modeling Success Rates or Proportions:** Since the Beta distribution is bounded between 0 and 1, it is often used to model proportions or success rates. For example, in A/B testing, where you want to compare the success rates of two different treatments, the success rates can be modeled as Beta distributed.

2. **Bayesian Statistics:** The Beta distribution is often used as a prior distribution for the parameter p of a Bernoulli or a Binomial distribution in Bayesian statistics. This is because the Beta distribution is a conjugate prior of these distributions. When a Beta distribution is used as a prior, the posterior distribution is also a Beta distribution, which simplifies the computation of the posterior.

Though the Beta distribution is a flexible distribution that can model a wide variety of shapes on the interval $[0, 1]$, but it does have some limitations:

1. The Beta distribution is defined only on the interval $[0, 1]$. This means it's not appropriate for variables that can take values outside this interval. If your data includes values at exactly 0 or 1, they will be outside the support of the Beta distribution.

2. Although the Beta distribution can take many shapes, it may not be able to accurately represent all types of data. For instance, it cannot model multimodal distributions (distributions with multiple peaks) on the interval $[0, 1]$.

3. When used in a Bayesian context, inference with the Beta distribution can become mathematically complex if it is not being used as a conjugate prior. This is true of any distribution, but it's worth noting since the Beta distribution is often used as a prior in Bayesian statistics.

4. The parameters of a Beta distribution do not have a direct "real-world" interpretation, unlike the mean of a Normal distribution or the rate of an Exponential distribution. This can make it more difficult to specify a prior if using the Beta distribution in a Bayesian context.

In this project, there are three concepts required to be mentioned:

1. Derive the Beta distribution from the scratch.

2. Show how the parameters of the Beta distribution affect the shape of the distribution.

3. Show why choose the parameter of the Beta distribution in Bayesian context is important. Comparing the positive and negative examples.