4.8: Beta Distributions

In this section, we introduce beta distributions, which are very useful in a branch of statistics known as Bayesian Statistics.

Beta Distributions

Definition 4.8.1

A random variable X has a ***beta distribution*** with parameters $\alpha, \beta > 0$, write $X \sim \text{beta}(\alpha, \beta)$, if X has pdf given by

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & \text{for } 0 \leq x \leq 1, \\ 0 & \text{otherwise,} \end{cases} \tag{4.8.1}$$

Note that the gamma function, $\Gamma(\alpha)$, is defined in Definition 4.5.2.

In the formula for the pdf of the beta distribution given in Equation 4.8.1, note that the term with the gamma functions, i.e., $\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ is the *scaling constant* so that the pdf is valid, i.e., integrates to 1. This is similar to the role the gamma function plays for the gamma distribution introduced in Section 4.5. Ignoring the scaling constant for the beta distribution, we can focus on what is referred to as the *kernel* of the distribution, which is given by

$$x^{\alpha-1}(1-x)^{\beta-1}, \quad \text{for } x \in [0,1]. \tag{4.8.2}$$

The parameters, α and β, are both ***shape parameters*** for the beta distribution, varying their values changes the shape of the pdf.

As is the case for the normal, gamma, and chi-squared distributions, there is no closed form equation for the cdf of the beta distribution and computer software must be used to calculate beta probabilities. Here is a link to a beta calculator online.

Beta distributions are useful for modeling random variables that only take values on the unit interval $[0,1]$. In fact, if both parameters are equal to one, i.e., $\alpha=\beta=1$, the corresponding beta distribution is equal to the uniform$[0,1]$ distribution. In statistics, beta distributions are used to model proportions of random samples taken from a population that have a certain characteristic of interest. For example, the proportion of surface area in a randomly selected urban neighborhood that is green space, i.e., parks or garden area.

We state the following important properties of beta distributions without proof.

Properties of Beta Distributions

If $X \sim \text{beta}(\alpha, \beta)$, then:

1. the mean of X is $\text{E}[X] = \dfrac{\alpha}{\alpha+\beta}$,
2. the variance of X is $\text{Var}(X) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.