Parametric Distributions

PSE 3.1

In some cases, a researcher may know (or be willing to assume that they know) the distribution of a random variable. More commonly, a researcher might be willing to assume that they know the distribution of a random variable up to a finite number of **parameters**, where the parameters allow for some flexibility in the shape of the distribution. In this section, we will briefly cover some of the most common/useful parametric distributions (though there are a number of other distributions mentioned in this chapter in the textbook that are worth mentioning).

Bernoulli Distribution

PSE 3.2

A random variable that follows a **Bernoulli distribution** takes either the value 0 or 1 with some probability p. An example is whether or not a particular person is employed. In particular, we can write

$$\begin{aligned}\mathrm{P}(X = 1) &= p \\ \mathrm{P}(X = 0) &= 1 - p\end{aligned}$$

which fully summarizes the distribution of X. In practice, in some cases p might be known (e.g., for flipping a coin, $p = 0.5$), but in most interesting cases you would need to somehow estimate p (we'll return to this sort of issue later). It is also useful to have a single, full expression for the pmf of a Bernoulli random variable, which is given by

$$\mathrm{P}(X = x) = p^x (1 - p)^{(1-x)}$$

In particular, just try plugging in $x = 1$ and $x = 0$ here and you will get $\mathrm{P}(X = 1) = p$ and $\mathrm{P}(X = 0) = (1 - p)$ as above.

Two interesting properties of Bernoulli random variables are:

1. $\mathbb{E}[X] = p$, in words: the expected value of X equals p
2. $\mathrm{var}(X) = p(1 - p)$ in words: the variance of X equals $p(1 - p)$.

Here is a proof of the first property, and I'll leave proving the second property as an excerise.

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x \in \{0,1\}} x\mathrm{P}(X = x) \\ &= 0(1 - p) + 1p \\ &= p\end{aligned}$$

Normal Distribution

PSE 3.12

Arguably the most important parametric distribution is the normal distribution. We will soon see that random variables following a standard normal distribution show up naturally in statistics due to the central limit theorem.

A random variable X that follows a normal distribution is continuous and the normal distribution is indexed by two parameters μ (its means) and σ^2 (its variance). It is common to write $X \sim N(\mu, \sigma^2)$ for a random variable that follow a normal distribution. The pdf of X is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

A special case (but important case) is when $\mu = 0$ and $\sigma^2 = 1$. In this case, X is said to follow a **standard normal distribution**.

Another useful property of normal distributions is stated in the following propsition

Proposition: Suppose that $X \sim N(\mu, \sigma^2)$ and we define $Y = a + bX$ for some $a, b \in \mathbb{R}$. Then Y is also normally distributed as $N(a + b\mu, b^2\sigma^2)$.

A useful consequence of the previous proposition is that, when $X \sim N(\mu, \sigma^2)$, it can be "standardized" by considering the transformed random variable $Z = (X - \mu)/\sigma$. This sort of transformation will be useful to us in the statistics portion of the course.

There is also some specialized notation that is worth mentioning for standard normal random variables. The cdf of a standard normal distribution is often denoted by Φ, and the pdf is often denoted by ϕ.