

## Unit 5, Lecture 1

*Numerical Methods and Statistics***Companion Reading**

Bulmer Chapter 6

**1 Distributions**

Some random variables/sample spaces are so common, we have a name and equations that describe them. For example, consider the simplest probability mass function/distribution:

**1.1 Bernoulli Distribution**

The Bernoulli distribution has a sample space of size 2 (e.g., 0, 1) with probability:

$$P(X = 1) = p, P(X = 0) = 1 - p \quad (1)$$

The expected value is  $p$  and the variance is  $p(1 - p)$ . An example of a process that follows a Bernoulli distribution is flipping a coin once. Typically 1 is denoted as success and 0 as failure.

There are three important things to recall about a distribution: its **sample space type** (continuous/discrete), its **support** or **sample space**, and what it describes. You should remember these things for many common distributions and then look up the details such as their variance in your textbook or online. Support (the probability is greater than 0) is

$$x \in P(X = x) > 0 \quad (2)$$

Support is important when the sample space contains zero probability elements. For example, if in a Bernoulli  $p = 0$ , that means there is no probability mass for  $x = 1$  and thus  $x = 1$  is no longer in the support.

**1.2 Geometric Distribution**

A geometric distribution is discrete and has a sample space of  $[1, \infty)$ . It is the number of trials required until success. Specifically, we have a Bernoulli process where the probability of success is a constant  $p$  and  $n$  is the number of failures until a success is achieved:

$$P(X = n) = (1 - p)^{n-1}p \quad (3)$$

The expected value is  $1/p$  and the variance is:

$$\frac{1 - p}{p^2} \quad (4)$$

An example is the number of times until a heads is seen while flipping a coin. Notice that the Geometric is unbounded in sample space. This makes it simple to distinguish between the other discrete probability distributions. For example, a Bernoulli distribution could be if you like a Facebook post or not. There are two outcomes. A Geometric distribution could be the number of Facebook posts you view before you like one. Depending on how cynical you are, there is no upper bound on the number you view.

### 1.3 Binomial Distribution

A binomial distribution is discrete and has a sample space of  $[0, N]$ . It is the number of successes in  $N$  trials. It is an important distribution for considering permutations. Its equation is:

$$P(X = n) = \binom{N}{n} p^n (1-p)^{N-n} \quad (5)$$

$$\binom{N}{n} = \frac{N!}{(N-n)!n!} \quad (6)$$

and the expected value is  $Np$ . The variance is  $Np(1-p)$ . An example is the number of boys in a family of eight children.

An easy way to distinguish a Bernoulli from a Binomial is to think about if the sample space has more than 2 elements. For example, I may retweet a tweet or not. That's a Bernoulli. If I examine 5 tweets and I may retweet any of them, the number of retweets I make is a Binomial, since it can vary from 0 to 5 (6 possibilities).

### 1.4 Poisson Distribution

A Poisson distribution is discrete and has a sample space of  $[0, \infty)$ . It can be thought of an approximation to the Binomial distribution in the case of very low chances of success ( $p \ll 0.5$ ), where  $\mu = Np$ . Its equation is

$$P(X = x) = \frac{e^{-\mu} \mu^x}{x!} \quad (7)$$

and the expected value is  $\mu$ . The variance is also  $\mu$ . An example is the number of deaths by horse kick in the 14 corps of the Prussian Army. It can also be viewed as the number of events in a fixed time interval. For example, if we are watching a live stream of an event, the number of comments could be modeled by a Poisson distribution.

### 1.5 Exponential Distribution

The exponential distribution is a continuous distribution with a support of  $[0, \infty)$ . It generally describes time between events, especially events whose occurrence follow the Poisson distribution. Its distribution is

$$p(x) = \lambda e^{-\lambda x} \quad (8)$$

and the expected value is  $1/\lambda$ . You may also think of it in terms of a residence time  $\tau = 1/\lambda$ . Note that the equation is for its probability distribution function (pdf) and not probability *mass* function. This is indicated by the lower case  $p$ . The variance is  $\lambda^{-2}$  or  $\tau^2$ . Note that we can more easily imagine the process in terms of its cumulative probability (see drawing) and not instantaneous pdf.

Following our previous example of a live stream, the time between comments will follow an exponential distribution if the comments follow a Poisson distribution. Even though the Poisson is discrete, the time between comments is continuous since time is a continuous value.

### 1.6 Normal Distribution

The normal distribution is a continuous distribution with a support of  $(-\infty, \infty)$ . We will see why later in the semester, but a normal distribution is almost always a good guess for the description of a process if it is continuous. Its equation is:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (9)$$