2.5 Common Distributions in the wild

One notable advantage in studying random variables, is that many seemingly different situations can be reduced down to the same family of distributions. In this chapter we will study a selection of classic distributions which characterise many different kind of phenomena seen in the world. Throughout this section, all random variables will only take values in $\mathbb{N}_0$.

2.5.1 Bernoulli Random Variables

We actually saw this type of random variable in Example 2.4.3. A Bernoulli random variable can only take two values, either zero or one. Suppose $0 \leq p \leq 1$ is a probability, we say that X is Bernoulli distributed with probability p, if we have that $\mathbb{P}[X = 1] = p$ and $\mathbb{P}[X = 0] = 1 - p$. As some notation we write $X \sim \mathrm{Be}(p)$ to denote that X is Bernoulli distributed with probability p.

Problem 2.5.1. Suppose $X \sim \mathrm{Be}(p)$, show that $\mathbb{E}[X] = p$. Furthermore show that $\mathbb{E}[X^2] = p$, thus show that $\mathrm{Var}[X] = p(1 - p)$.

Example 2.5.1. Suppose that $X_1 \sim \mathrm{Be}(1/3)$ and $X_2 \sim \mathrm{Be}(3/4)$, what is the probability that $X_1 > X_2$?

We note that X_1 and X_2 are both Bernoulli distributed, so they may only take either the value of zero or one. So if $X_1 > X_2$ then we must have that $X_1 = 1$ and $X_2 = 0$. Written symbolically we have:

$$\mathbb{P}(X_1 > X_2) = \mathbb{P}(X_1 = 1 \cap X_2 = 0).$$

Now both as both of these events are independent, (i.e knowing the value of X_1 does not affect the value of X_2) we can apply Lemma 1.7.2 to find the following:

$$\mathbb{P}(X_1 = 1 \cap X_2 = 0) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 0) = \frac{1}{3} \cdot \left(1 - \frac{3}{4}\right) = \frac{1}{12}.$$

2.5.2 The Binomial Distribution

The Binomial distribution can be seen as an extension of the Bernoulli distribution. Before we define the Binomial distribution we briefly introduce some notation:

A Crash Course in Combinatorics

Definition 2.5.1. Suppose n is a natural number. Then we denote the *factorial* of n as, $n!$. We define it as follows:

$$n! := n \times (n - 1) \times (n - 2) \times \ldots \times 2 \times 1.$$

Essentially to find the factorial of n we just take a product of all numbers from one up to n. For example $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$. We also remark that $0! = 1$.

Definition 2.5.2. Suppose n and k are natural numbers, where $k \leq n$, we denote the *Binomial coefficient*, read as "n choose k", as follows:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Briefly the binomial coefficient describes the following scenario: Suppose you have a workforce made up of 40 members, and you want to select a committee consisting of 5 members from this group of 40. Then the binomial coefficient:

$$\binom{40}{5} = \frac{40!}{5! \times 35!} = 658008,$$

tells us the number of different ways we can select five members of the committee from a group of 40. The notion of choosing combinations is ubiquitous across many of mathematics, computer science and in real-world scenarios, however in-depth study is beyond the scope of this course. Your calculator will be able to compute both factorials and binomial coefficients, as they crop-up in the definition of the Binomial Distribution.

2.5.3 The Binomial Distribution

Definition 2.5.3. Let p be a probability and n a positive integer. We say that X is distributed binomially with parameters n and p if the following holds for each $i \in \mathbb{N}_0$:

$$\mathbb{P}(X = i) = \binom{n}{i} p^i (1-p)^{n-i}.$$

If X has a binomial distribution with parameters n and p, then we write that $X \sim \mathrm{Bin}(n, p)$.

The definition of the binomial distribution is definitely not as a friendly as that of the Bernoulli distribution. With the binomial distribution it is more important to understand the situation where it applies. A binomial distribution describes experiments which are made up of repeated sequences of identical and independent success or failure trials. For example suppose you have a nest containing eight different eggs. Suppose that each egg will hatch successfully with a probability of $3/4$ independently of all other eggs. Then the random Y which counts the numbers of eggs that hatch successfully has a binomial distribution, precisely $Y \sim \mathrm{Bin}(8, 3/4)$. The first parameter n is the number of trials that take place, while the second parameter p, is the success probability of the trials. We briefly remark the variance and the standard deviation of a binomial random variable:

Lemma 2.5.1. Suppose $X \sim \mathrm{Bin}(n, p)$, then we have that:

(i) $\mathbb{E}[X] = np$.

(ii) $\mathrm{Var}[X] = np(1-p)$.

Example 2.5.2. It is known that in trials for a certain drug, headaches occur as a side effect in 5% of the population. A sample of 30 participants test the drug and are observed for headaches. Let H be the number of participants in the sample which exhibit a headache as a side effect.

(i) What is the distribution of H?

(ii) What is the expected number of participants that will exhibit a headache?

(iii) What is the probability that exactly four participants will exhibit a headache?

(iv) What is the probability that at most 3 participants will exhibit a headache?

(v) what is the probability that at least 4 participants will exhibit a headache?

In this example we can think of each person being an independent trial, where the person either has a headache or they do not. The probability they will have a headache is 0.05, independently of any other person within the sample. There are 30 people in the sample, hence there are 30 trials. Therefore we have that $H \sim \text{Bin}(30, 0.05)$.

We apply Lemma 2.5.1 (i), with $n = 30$ and $p = 0.05$. As a result, we find that the expected number of headache sufferers is, $\mathbb{E}[H] = 30 \times 0.05 = 1.5$.

We are now looking to find $\mathbb{P}(H = 4)$. Therefore by using the definition of the binomial distribution (Definition 2.5.3), we have that:

$$\mathbb{P}(H = 4) = \binom{30}{4} \times (0.05)^4 \times (0.95)^{30-4} = 27405 \times (0.05)^4 \times (0.95)^{26} = 0.04514.$$

For the next part we are looking for $\mathbb{P}(X \leq 3)$. We note if $H \leq 3$, then $H = 0$, or $H = 1$, or $H = 2$, or $H = 3$. All of these events are disjoint therefore:

$$\mathbb{P}(H \leq 3) = \mathbb{P}(H = 0) + \mathbb{P}(H = 1) + \mathbb{P}(H = 2) + \mathbb{P}(H = 3).$$

For each of these terms we just apply the formula from Definition 2.5.3, I do encourage you to check you are happy where these numbers have come from:

$$\mathbb{P}(H \leq 3) = 0.2146 + 0.3389 + 0.2586 + 0.1270 = 0.939.$$

For the final part we are now trying to compute $\mathbb{P}(H \geq 4)$. If we try to proceed as we did in the previous part we would run into trouble as we would need to compute 26 different terms. However we note that the events $\{H \leq 3\}$ and $\{H \geq 4\}$ are complement events. Either H is less than or equal to 3, or it is at least four, exactly one of these must be true. Therefore we may apply Lemma 1.4.3 to find that:

$$\mathbb{P}(H \leq 3) + \mathbb{P}(H \geq 4) = 1.$$

Therefore it follows that $\mathbb{P}(H \geq 4) = 1 - 0.939 = 0.061$.

A few things to remark about the previous example. Whenever tackling these sorts of questions always clearly state the distribution the random variable that you are working with. Another thing to note is when computing these cumulative events (i.e $\mathbb{P}(X \geq 10)$) sometimes it is easier to consider the complement event and then apply Lemma 1.4.3. Finally this example demonstrates some idea of the shape of the binomial distribution. Most of the probability is concentrated around the expectation, and the further we get from the expectation the less likely we are to observe the outcome. As you might expect $\mathbb{P}(H = 30)$ is the least likely outcome, with a probability of $(0.05)^{30} \approx 9.31 \times 10^{-40}$.