

# Chapter 6

## Random Variables

### 6.1 Random Variables

In this chapter, we will put together the ideas we have developed for probability and descriptive statistics, to build tools that will help us understand statistics such as the average of a random sample.

The story begins with an *outcome space*, that is, the set of all possible outcomes of an experiment that involves chance. Standard notation for this space is  $\Omega$ , the upper case Greek letter Omega. **For mathematical simplicity, we will assume that the outcome space  $\Omega$  is finite.** Each outcome  $\omega$  (that's lower case omega) is assigned a probability, and the total probability of all the outcomes is 1.

**Example 1.** You flip a fair coin three times, and record what face the coin landed on. What is a reasonable outcome space  $\Omega$  for this experiment, and what is the probability distribution on that space ?

You can think of  $\Omega$  as consisting of 8 equally likely outcomes, since the coin is not biased. Here's a table representing  $\Omega$  and the probabilities of all the outcomes.

$\omega$	TTT	TTH	THT	HTT	THH	HTH	HHT	HHH
$P(\omega)$	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

A *random variable*  $X$  is a real-valued function defined on  $\Omega$ . That is, the domain of  $X$  is  $\Omega$  and the range of  $X$  is the real line.

In the example above,  $X$  could be the number of times the letter H appears in an outcome. You can think of  $X$  as the number of heads in three tosses of a coin.

$\omega$	TTT	TTH	THT	HTT	THH	HTH	HHT	HHH
$P(\omega)$	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8
$X(\omega)$	0	1	1	1	2	2	2	3

The probability function on  $\Omega$  determines probabilities for  $X$ . For example, the chance that  $X$  is 1 is defined as follows:

$$P(X = 1) = P(\{\omega : X(\omega) = 1\}) = P(\text{TTH, THT, HTT}) = 1/8 + 1/8 + 1/8 = 3/8$$

## 6.2 Probability distribution

The *probability distribution* of  $X$  is a distribution on the range of  $X$ . It specifies all the possible values of  $X$  along with all their probabilities.

For example, for  $X$  defined in the example above, the probability distribution is given by

$x$	0	1	2	3
$P(X = x)$	1/8	3/8	3/8	1/8

All the probabilities in a distribution must add up to 1.

Note that the probabilities on the range of  $X$  are determined by the probabilities on the domain. In our example, had all 8 elements in the domain not been equally likely, the possible values of  $X$  would still have been the same but the probabilities might have been different.

**Example 2: Uniform distribution** A word is picked at random from the set data, science, connector, class. Let  $V$  be the number of distinct vowels in the word. What is the distribution of  $V$ ?

**Solution.**  $\Omega = \{\text{science, computer}\}$ . The probability distribution on  $\Omega$  assigns probability  $1/2$  to each element. Now  $V(\text{science}) = 2$  and  $V(\text{computer}) = 3$ .

So the distribution of  $V$  puts probability  $1/2$  on each of the values 2 and 3. We say that  $V$  has the *uniform* distribution on  $\{2, 3\}$ .

**Example 3: Binomial distribution.** Suppose a coin is tossed  $n$  times. Let  $X$  be the number of heads. Find the distribution of  $X$ .

**Solution.** Each possible outcome of  $n$  tosses is a sequence of  $n$  H's and T's. You have seen in an earlier exercise that there are  $2^n$  equally likely such sequences.

When you are finding a distribution, an excellent idea is to start with the possible values rather than the probabilities. The possible number of heads in  $n$  tosses is 0 through  $n$ .

For any fixed  $k$  in the range 0 through  $n$ , the number of sequences that contain exactly  $k$  H's is  $\binom{n}{k}$ . So the distribution of  $X$  is

$$P(X = k) = \frac{\binom{n}{k}}{2^n}, \quad k = 0, 1, \dots, n$$

This is called the *binomial distribution with parameters  $n$  and  $1/2$* . You should convince yourself that the formula gives a sensible answer in the two edge cases  $k = 0$  and  $k = n$ .

If the coin is unfair, it is still true that there are  $\binom{n}{k}$  sequences with exactly  $k$  H's, but all  $2^n$  sequences are no longer equally likely. To find the probability of each sequence, let  $p$  be the chance that the coin lands heads on a single toss, and make the natural assumption that the outcome of any set of tosses does not affect chances for any other set.

Let's work out an example under these assumptions before finding the general formula. Suppose  $n = 3$  as in Example 1, and let  $k = 2$ . Then the sequences corresponding to 2 heads in 3 tosses are HHT, HTH, and THH. Each of these has probability  $p^2(1-p)$ , because there are two factors of  $p$  and one of  $(1-p)$  in different order. So the chance of 2 heads in 3 tosses is  $P(\text{HHT or HTH or THH}) = 3p^2(1-p)$ .

Notice that the 3 in the answer above is the number of sequences that have 2 H's, that is,  $3 = \binom{3}{2}$ .

Now we are ready to extend the argument to the general case. Suppose the coin lands heads on a single toss with probability  $p$ . If  $X$  is the number of heads on  $n$  tosses, then the distribution of  $X$  is

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

This is called the *binomial distribution with parameters  $n$  and  $p$* . Notice that the fair coin is a special case:

$$P(X = k) = \frac{\binom{n}{k}}{2^n} = \binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k}$$

The name *binomial* comes from the fact that the probabilities in the binomial distribution are the terms in the *binomial expansion*  $(a + b)^n$  for  $a = p$  and  $b = 1 - p$ . This also shows why the sum of the terms is equal to 1.

The distribution can be used to find probabilities of events, as in the following examples.

The chance of getting between 45 and 55 heads (inclusive) in 100 tosses of a fair coin is

$$\sum_{k=45}^{55} \frac{\binom{100}{k}}{2^{100}}$$

The chance of getting fewer than 10 sixes in 12 rolls of a die is

$$\sum_{k=0}^9 \binom{12}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{12-k} = 1 - \sum_{k=10}^{12} \binom{12}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{12-k}$$

The second form is simpler to calculate, as the sum only has three terms.

**Example 4.** A 5-card poker hand is dealt from a standard deck. What is the distribution of the number of queens in the hand?

**Solution.** Let's give the variable a name:  $Q$  for "queens". We will start by listing the possible values of  $Q$ . We are dealing 5 cards but there are only 4 aces in the deck. So the possible values of  $Q$  are 0, 1, 2, 3, 4.

The total number of possible hands is the total number of subsets of 5 that can be chosen from among 52 cards. That's  $\binom{52}{5}$ , and they are all equally likely.

Now fix  $k$  in the range 0 through 4. To find  $P(Q = k)$ , we will need the number of hands that contain exactly  $k$  queens. These  $k$  queens can be chosen in  $\binom{4}{k}$  ways. For each of these ways, there are  $\binom{48}{5-k}$  ways of choosing the remaining  $5 - k$  cards in the hand. So the number of 5-card hands that contain exactly  $k$  queens is  $\binom{4}{k} \binom{48}{5-k}$ .

So the distribution of  $Q$  is given by

$$P(Q = k) = \frac{\binom{4}{k} \binom{48}{5-k}}{\binom{52}{5}}, \quad k = 0, 1, 2, 3, 4$$

The chance of three or more queens in a 5-card hand is

$$P(Q \geq 3) = \frac{\binom{4}{3} \binom{48}{2}}{\binom{52}{5}} + \frac{\binom{4}{4} \binom{48}{1}}{\binom{52}{5}}$$

### 6.3 Functions of Random Variables

We will frequently be interested in quantities that can be calculated based on a random variable. In other words, we will be interested in functions of random variables. A function of a random variable is also a random variable, because a function of a function on  $\Omega$  is also a function on  $\Omega$ .

**Example 5.** Let  $X$  have the binomial  $(4, 1/2)$  distribution. Let  $Y$  be the absolute deviation of  $X$  from 2, that is,  $Y = |X - 2|$ . What is the distribution of  $Y$ ?

**Solution.** The table below gives the distribution of  $X$  along with the possible value  $y$  of  $Y$  computed from each possible value  $x$  of  $X$ :

$y$	2	1	0	1	2
$x$	0	1	2	3	4
$P(X = x)$	1/16	4/16	6/16	4/16	1/16

Collect terms get the distribution of  $Y$ :

$y$	0	1	2
$P(Y = y)$	6/16	8/16	2/16

**Example 6: Net gain on red at roulette.** In Nevada roulette, the bet on "red" pays 1 to 1 and you have 18 in 38 chances of winning. This means that if you bet a dollar on "red" and the winning color is not red, you lose your dollar; if the winning color is red, then your net winnings are a dollar. Suppose the roulette wheel is spun 10 times, and you bet a dollar on red each time. What is the chance that you make money?

**Solution.** Let  $X$  be the number of times you win. Then  $X$  has the binomial distribution with parameters  $n = 10$  and  $p = 18/38$ . But the question is about your net gain being positive overall. So let your net gain be  $G$ , and let's try and work out what  $G$  has to be for you to make money.

For any possible value  $x$  of  $X$ , the amount of money you make is the total of \$1 for each of the  $x$  bets that you win and -\$1 for each of the  $10 - x$  bets that you lose. This corresponds to making  $g$  dollars overall, where

$$g = 1 \cdot x + (-1) \cdot (10 - x)$$

So

$$g = 2x - 10$$

So your random net gain  $G$  is the following function of the random number of bets  $X$  that you win:

$$G = 2X - 10$$

The question asks for  $P(G > 0)$ .

$$P(G > 0) = P(2X - 10 > 0) = P(X > 5) = \sum_{x=6}^{10} \binom{10}{x} \left(\frac{18}{38}\right)^x \left(\frac{20}{38}\right)^{10-x}$$