Some notes on random variables: probability mass functions, expected value, and the binomial distribution

1 Random variables, probability mass functions, and expected value

We begin with some definitions from probability.

Definition 1.1. An *experiment* is any procedure that can be repeated indefinitely, and that has a well-defined set of possible outcomes. The set of all such outcomes is called the *sample space* of the experiment.

An *event* is a subset of the sample space. That is, an event is a set of possible outcomes. If A is an event, then we denote by $P(A)$ the probability that A occurs.

A *random variable* X is a way of assigning numerical values to outcomes or events of an experiment. A *discrete* random variable is one that takes on only finitely many or countably many values.

All random variables for us will be discrete.

Example 1.2. An example of an experiment would be the process of flipping six fair coins. The sample space for such an experiment could be described as the set S of all six-letter strings where each letter is an H (for heads) or a T (for tails). Note that $|S| = 2^6 = 64$.

An example of an event A in this sample space might be the event where exactly two of the coins land heads. You might recall from an activity we did in class that

$$A = \{\text{HHTTTT}, \text{HTHTTT}, \text{HTTHTT}, \text{HTTTHT}, \text{HTTTTH}, \text{THHTTT}, \text{THTHTT}, \text{THTTHT}, \text{THTTTH}, \text{TTHHTT}, \text{TTHTHT}, \text{TTHTTH}, \text{TTTHHT}, \text{TTTHTH}, \text{TTTTHH}\},$$

so that $|A| = 15$.

As we mentioned in class, if all outcomes of an experiment are equally likely (as they are in this case, since the coins are fair), then for any event A,

$$P(A) = \frac{|A|}{|S|}.$$

So in this case, for example, we have

$$P(A) = \frac{15}{64} \approx 23.43\%.$$

An example of a random variable X for this experiment might be

$$X = \text{the number of heads that show up.}$$

Example 1.3. Another example of an experiment might be the process of rolling two fair dice. The set S of outcomes here would be the set of ordered pairs of numbers that come up (where each number is a 1 through a 6). For this experiment, we might define a random variable X by

$$X = \text{the number of 5's that come up.}$$

Of course, other random variables are possible; for example:

$$\begin{aligned} Y &= \text{the sum of the numbers on the two dice,} \\ Q &= \text{the smallest number that comes up,} \\ Z &= \text{the largest number that comes up,} \\ \Lambda &= \text{the average of the numbers that come up,} \end{aligned}$$

and so on.

We need some more definitions.

Definition 1.4. Given a random variable X and a possible value x of X, we write $P(X = x)$ for the probability that X takes the value x.

The *probability mass function* for X is just a description of what $P(X = x)$ equals for all possible values x of X.

Example 1.5. Consider the experiment of rolling two fair dice, as in Example 1.3 above. Let X and Z be as above: X is the number of 5's that come up, and Z is the largest number that comes up.

(a) What is the probability mass function for X?

(b) What is the probability mass function for Z?

Solution. (a) X can take the values $0, 1$, or 2. So finding the probability mass function for X means computing $P(X = x)$ for $x = 0, 1, 2$.

As in Example 1.2, we can compute the probability mass function as follows: for each possible value x, count the number of outcomes for which $X = x$, and divide this by the total number of possible outcomes. As mentioned above, this strategy works when all outcomes are equally likely, as they are in this case. (Any pair of numbers coming up is as likely as any other pair.)

To do the counting, let's write down the sample space S for this experiment. We have

$$\begin{aligned} S = \{&11, 12, 13, 14, 15, 16, 21, 22, 23, 24, 25, 26, 31, 32, 33, 34, 35, 36, \\ &41, 42, 43, 44, 45, 46, 51, 52, 53, 54, 55, 56, 61, 62, 63, 64, 65, 66.\} \end{aligned}$$

Note that $|S| = 6^2 = 36$.

We see that 25 of these outcomes have zero 5's, 10 of them (15,25,35,45,51,52,53,54,56,65) have exactly one 5, and only one (55) has two 5's.

So we compute:

$$P(X = 0) = \frac{25}{36} \approx 69.44\%, \qquad P(X = 1) = \frac{10}{36} \approx 27.78\%, \qquad P(X = 2) = \frac{1}{36} \approx 2.78\%.$$

Note that the probabilities add up to $\frac{36}{36} = 100\% = 1$. (Sometimes, because of roundoff error, you might get a number just a bit different from 100%, when you add up your decimal perecentages.)

(b) Z can take the values 1 through 6. So finding the probability mass function for Z means computing $P(Z = z)$ for $z = 1, 2, 3, 4, 5, 6$.

As we did for X, we count the appropriate outcomes from the sample space S above, and divide by the size of S. For example, there's only one outcome where the largest number showing is 1: namely, the outcome 11. So

$$P(Z = 1) = \frac{1}{36} \approx 2.78\%.$$

Similarly, there are three outcomes for which the largest number showing is 2: 12, 21, and 22. So

$$P(Z = 2) = \frac{3}{36} \approx 8.33\%.$$

There are five outcomes (13, 23, 31, 32, and 33) for which the largest number showing is 3, and so on. We compute:

$$P(Z = 3) = \frac{5}{36} \approx 13.89\%, \qquad P(Z = 4) = \frac{7}{36} \approx 19.44\%,$$

$$P(Z = 5) = \frac{9}{36} = 25.00\%, \qquad P(Z = 6) = \frac{11}{36} \approx 30.56\%.$$

You can check that, again, the probabilities add up to $\frac{36}{36} = 100\% = 1$.

We conclude this section with one more definition.

Definition 1.6. The *expected value* $E(X)$ of a random variable X is defined by

$$E(X) = \sum_x x \cdot P(X = x),$$

where the sum is taken over all possible values x of X.

Remark: the definition of expected value says: Take each possible value x of X, "weight" it by the proportion of the time that this value occurs (that is, by $P(X = x)$), and add all these weighted values up. So expected value is sort of an average of what you expect X to be.

Example 1.7. What are the expected values of the random variables X and Z of Example 1.5 above?

Solution. The expected value $E(X)$ of X is given by

$$\begin{aligned} E(X) &= \sum_{x=0}^{2} x \cdot P(X = x) = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) \\ &= 0 \cdot \frac{25}{26} + 1 \cdot \frac{10}{36} + 2 \cdot 136 = \frac{12}{36} \approx 0.3333. \end{aligned}$$

That is, if we roll two fair dice and record the number of 5's that come up, then on average we would expect to see 0.3333 5's. Note that this makes perfect sense: we're rolling two dice, and since there are 6 numbers on each die, we'd expect on average to see $1/6$ of a 5 on each die, so we'd expect on average to see $2/6 = 1/3 \approx 0.3333$ of a 5 on the two dice together.

Next, the expected value $E(Z)$ of Z is given by

$$E(Z) = \sum_{z=1}^{6} z \cdot P(Z = z) = 1 \cdot \frac{1}{36} + 2 \cdot \frac{3}{36} + 3 \cdot \frac{5}{36} + 4 \cdot \frac{7}{36} + 5 \cdot \frac{9}{36} + 6 \cdot \frac{11}{36}$$
$$\approx 4.4722.$$

On average, then, we'd expect the largest of the two numbers appearing on the dice to be 4.4722 (which is not something you could come up with intuitively).

Section 1 Exercises.

Exercise 1.1. Find the probability mass function and the expected value for the random variables Y and Q of Example 1.3 above. (Use the method of Example 1.5.)

Exercise 1.2. Flip four fair coins, and record each outcome as a string of H's and T's, as in Example 1.2 above, except that here we have only four coins, not six.

(a) Write down the sample space S for this experiment. That is, write down, explicitly, all possible outcomes.

(b) Let X be the number of heads showing on your flip of the four coins. Find the probability mass function for X, as well as the expected value $E(X)$. Use the method of Example 1.5 above. That is: for each value x of X, *count* how many outcomes have x heads, and divide by the total number of possible outcomes to find $P(X = x)$.
Why does your answer for $E(X)$ make intuitive sense?

2 The binomial distribution

A *binomial* experiment is one with only two possible outcomes. We generally call one of these outcomes a *success* and the other a *failure*. We typically denote the probability of success by p.

We have:

Theorem 2.1. Let X denote the number of successes in a single trial of a binomial experiment that has P(success)$= p$. Then $E(X) = p$.

Proof. If X is as described, then X takes the value 1 with probability p, and takes the value 0 with probability $1 - p$, and takes no other values, so

$$E(X) = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) = 0 \cdot (1 - p) + 1 \cdot p = p.$$

□

Example 2.2. If X denotes the number of shots made by an 80% free-throw shooter in a single attempt, find $E(X)$.

Solution. By Theorem 2.1, we have $E(X) = p = 0.8$.

Again, this answer makes sense: If the shooter hits 80% of their free throws, then on average we'd expect that, on a single shot, they'd make 0.8 free throws.

Often we will want to understand what things look like if we *repeat* a binomial experiment n times. One question we might want to ask in such a context is: what is the probability that exactly k of these n trials of our binomial experiment will result in success? Here $0 \leq k \leq n$.

To answer this question, we will first need a basic fact from probability theory. This fact says: suppose A and B are events, and AB denotes the event where both A *and* B occur. (Recall that events are subsets of the sample space S. In terms of set theory, AB is the same as $A \cap B$.) Then, *if* A and B are *independent* events, meaning, essentially, that whether one occurs does not affect whether the other does, then

$$P(AB) = P(A)P(B).$$

This result should not surprise you: it says, *roughly*, that "and" means "multiply," when it comes to probabilities (just as it does when it comes to counting).

For example, flip two fair coins; let A be the event that the first coin lands heads, and B the event that the second coin does. Since $P(A)$ and $P(B)$ are both equal to $1/2$, and since the outcomes of the two flips are independent, we have

$$P(AB) = P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = 25\%.$$

This makes intuitive sense, since, in the two flips, there are four possibilities: both are heads; the first is heads and the second tails; the first is tails and the second heads; or both are tails. And all four of these possibilities are equally likely, so the probability of any one of these things happening is $1/4$.

Back to the question, then: we repeat a binomial experiment, with probability of success p, n times, and we want to know the probability that exactly k of these trials result in a success. To determine this probability we first note that, since $P(\text{success}) = p$, we also have $P(\text{failure}) = 1-p$. Next: there are $\binom{n}{k}$ different ways to slot the k successes in among the n trials. For each such way of choosing where to put our k successes, what is the probability that there actually is a success in each of these places, and a failure in each other place? Well, the probability of success in any given slot is p, so the probability of success in k different slots is

$$\underbrace{p \cdot p \cdot p \cdots p}_{k \text{ times}} = p^k,$$

by the fact noted above. (Here we're assuming, as we always will with a binomial experiment, that successive trials are independent of each other.) Similarly, the probability of failure in the remaining $n-k$ slots is $(1-p)^{n-k}$. Putting all of this together, we have the following.

Theorem 2.3. If a binomial experiment, with $P(\text{success}) = p$, is repeated n times, and all trials are independent, then the probability of exactly k successes among those n trials is given by

$$P(k \text{ successes in } n \text{ trials}) = \binom{n}{k} p^k (1-p)^{n-k} \qquad (0 \leq k \leq n).$$

Remark: when we write P(k successes in n trials), we will always mean the probability of *exactly* k successes (rather than *at least* k, or *at most* k, or anything of that nature) among the n trials.

Example 2.4. An 80% free-throw player attempts four free throws.

(a) Find the probability mass function for the number X of free throws made out of the four.

(b) Find the probability that the player makes at least two free throws.

(c) Find the expected number of free throws made.

Solution. (a) By Theorem 2.3, we have:

$$\begin{aligned} P(X=0) &= \binom{4}{0}(0.8)^0(0.2)^{4-0} = 0.0016, \\ P(X=1) &= \binom{4}{1}(0.8)^1(0.2)^{4-1} = 0.0256, \\ P(X=2) &= \binom{4}{2}(0.8)^2(0.2)^{4-2} = 0.1536, \\ P(X=3) &= \binom{4}{3}(0.8)^3(0.2)^{4-3} = 0.4096, \\ P(X=4) &= \binom{4}{4}(0.8)^4(0.2)^{4-4} = 0.4096. \end{aligned}$$

Note that the probabilities add up to 1.

(b) We compute that

$$P(X \text{ is at least two}) = P(X=2)+P(X=3)+P(X=4) = 0.1536+0.4096+0.4096 = 0.9728.$$

(c) We have

$$\begin{aligned} E(X) &= 0\cdot P(X=0) + 1\cdot P(X=1) + \cdots + 4\cdot P(X=4) \\ &= 0\cdot 0.0016 + 1\cdot 0.0256 + 2\cdot 0.1536 + 3\cdot 0.4096 + 4\cdot 0.4096 = 3.2. \end{aligned}$$

Remark: in part (b) of the above example, we used the fact that the probability of X equalling 2 or 3 or 4 is the sum of the probabilities $P(X=2)$, $P(X=3)$, and $P(X=4)$. It's not hard to show that, in general, probability mass functions add this way.

It's not surprising that, in the above example, $E(X)=3.2$: if the player hits 80% of their shots then, out of 4 shots, we would expect them to hit 80% of those four, and $0.8\times 4 = 3.2$.

Section 2 Exercises.

Exercise 2.1. A 40% three-point field goal shooter attempts five three-point shots in a game.

(a) Find the probability mass function for the number X of three-point shots made out of the five. Confirm that your probabilities add up to one.

(b) Find the probability that the player makes at least two three-point shots, out of the five taken. Hint: it might be easier to first compute the probability that they make fewer than two.

(c) Find the expected number of three-point shots made. Use the method of Example 2.4 above (that is, use the definition of expected value, and the probabilities that you computed in part (a) of this exercise).

(d) Why does your answer to part (c) of this exercise make intultive sense?

Exercise 2.2. Consider the following game. You pay \$10, and pick a number from 1 through 6. A fair die is rolled three times. You are awarded \$0 if your number does not come up at all in the three rolls; you are awarded \$20 if your number comes up once, \$30 if it comes up twice, and \$40 if it comes up all three times.

Let X be the number of times your chosen number comes up. Then X is the number of successes in three trials of a binomial experiment, with $P(\text{success}) = 1/6$.

(a) Find the probability mass function for X.

(b) Should you play the game? Hint: your expected payoff, in dollars, is

$$-10 \cdot P(X = 0) + 10 \cdot P(X = 1) + 20 \cdot P(X = 2) + 30 \cdot P(X = 3).$$

Exercise 2.3. Find the probability mass function and the expected value for the random variable X of Example 1.2 (the example where you flip six fair coins). Use Theorem 2.3 above (that is, *do not* actually count outcomes, as you did in Exercise 1.2 above).