

Chapter 5

Key Ideas

Random Variables – Discrete and Continuous, Expected Value
Probability Distributions – Properties, Mean, Variance and Standard Deviation
Unusual Results and the Rare Event Rule
Binomial Distribution – Properties, Finding Probabilities

Section 5-1: Overview

Chapters 1-3 involved summarization of data sets. Chapter 4 involved probability calculations. In Chapter 5, we will combine the two techniques and begin trying to *predict* what will happen in the sample based on probability. This can be done by talking about things called random variables.

Section 5-2: Random Variables

Definitions:

- A *random variable* is a variable that has a single numerical value, determined by chance, for each outcome of a procedure.
- A random variable is *discrete* if it takes on a countable number of values (i.e. there are gaps between values).
- A random variable is *continuous* if there are an infinite number of values the random variable can take, and they are densely packed together (i.e. there are no gaps between values).
- A *probability distribution* is a description of the chance a random variable has of taking on particular values. It is often displayed in a graph, table, or formula.
- A *probability histogram* is a display of a probability distribution of a discrete random variable. It is a histogram where each bar's height represents the probability that the random variable takes on a particular value.

Notation:

- Random Variables are usually denoted X or Y.
- The probability a random variable X takes on a value x is $P(x) = P(X=x)$
- A Probability Distribution, then, is a specification of $P(x)$ for each x that is an outcome of a procedure. As mentioned before, this could be done in a graph, table, or formula.

Properties of a Probability Distribution

There are 2 properties a probability distribution must satisfy. First of all, since $P(x)$ describes a probability, each $P(x)$ must be between 0 and 1. Also, since X *must* take one value after the procedure, the sum of the $P(x)$ values should be 1. So we have:

$$1. \sum_x P(X = x) = 1$$

$$2. 0 \leq P(x) \leq 1$$

Example:

Suppose we record four consecutive baby births in a hospital. Let X = The difference in the number of girls and boys born. X is discrete, since it can only have the values 0, 2, or 4. (Why can't it be 1 or 3?) Furthermore, if we write out the sample space for this procedure, we can find the probability that X equals 0, 2, or 4:

$S = \{mmmm, mmmf, mmfm, mfmm, fmmm, mmff, mfmf, mffm, fmf, fimm, fmmf, mfff, fmff, fmf, fffm, ffff\}$

There are 16 total cases, and each one is equally likely.

$P(X = 0) = 6/16 = 0.375$ (these are the cases mmff, mfmf, mffm, fmf, fimm, fmmf)

$P(X = 2) = 8/16 = 0.5$ (these are the cases mmmf, mmfm, mfmm, fmmm, mfff, fmff, fmf, fffm)

$P(X = 4) = 2/16 = 0.125$ (these are the cases mmmm and ffff)

Is this a probability distribution?

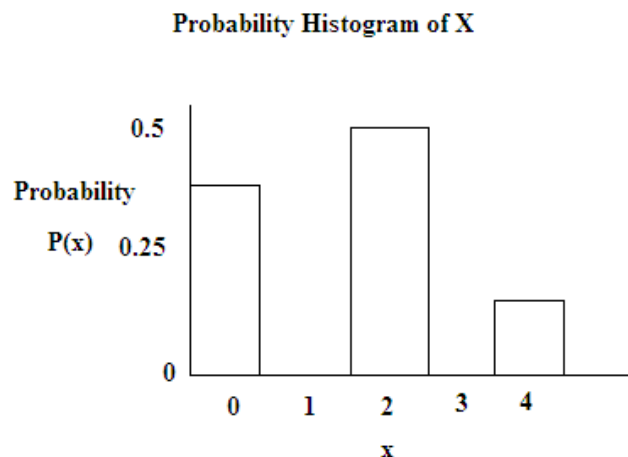
$$1. \sum_x P(X = x) = 0.375 + 0.5 + 0.125 = 1$$

$$2. 0 \leq 0.125 < 0.375 < 0.5 \leq 1$$

So it *is* a probability distribution.

We can display it in a table or as a graph:

X	0	2	4
$P(X=x)$	0.375	0.5	0.125



Another Example:

Is the following a probability distribution?

x	1	4	9	16	25	36	49	64	81	100	121
P(X=x)	0.20	0.15	0.14	0.12	0.10	0.09	0.07	0.05	0.04	0.02	0.01

Clearly, each probability is between 0 and 1. Now we need to see if they sum to 1.

$$\sum_x P(X = x) = 0.20 + 0.15 + 0.14 + 0.12 + 0.10 + 0.09 + 0.07 + 0.05 + 0.04 + 0.02 + 0.01 = 0.99$$

Since they do not sum to 1, it is *not* a probability distribution.

Mean and Variance of a Probability Distribution

The *mean* of a probability distribution (also called the *expected value* of the random variable X) is the value we would expect as the long-term average after repeating the procedure over and over again. There are two common forms of notation for the mean of a probability distribution: μ (denotes “mean”) and $E[X]$ (denotes “expected value of X”). We will use the $E[X]$ notation, primarily.

The *variance* of a probability distribution is a measure of the expected variation in the random variable. The variance, like the mean, also has two notations: σ^2 (denotes “variance”) and $\text{Var}[X]$ (denotes “variance of X”). We will usually use the $\text{Var}[X]$ notation.

To compute these values, we have the following formulas:

$$\mu = E[X] = \sum_x x \cdot P(x)$$

$$\sigma^2 = \text{Var}[X] = \sum_x (x - \mu)^2 \cdot P(x)$$

Shortcut Formula: $\sigma^2 = \text{Var}[X] = E[X^2] - E[X]^2$, where $E[X^2] = \sum_x x^2 \cdot P(x)$

The *standard deviation* of a probability distribution is the square root of the variance, $\sigma = \sqrt{\text{Var}[X]}$

Examples:

For each distribution, find the mean, variance, and standard deviation.

1. X = # of heads in 2 coin flips

x	0	1	2
P(x)	0.25	0.5	0.25

Mean: $\mu = E[X] = \sum_x x \cdot P(x) = 0(0.25) + 1(0.5) + 2(0.25) = 1$

This should make sense... over the long-term, we should expect 1 head in every 2 coin flips.

Variance: $\sigma^2 = \text{Var}[X] = \sum_x (x - \mu)^2 \cdot P(x) = (0-1)^2(0.25) + (1-1)^2(0.5) + (2-1)^2(0.25) = 0.5$

OR

$E[X^2] = \sum_x x^2 \cdot P(x) = 0^2(0.25) + 1^2(0.5) + 2^2(0.25) = 1.5$

Variance: $\sigma^2 = \text{Var}[X] = E[X^2] - E[X]^2 = 1.5 - 1^2 = 0.5$

Standard Deviation: $\sigma = \sqrt{\text{Var}[X]} = \sqrt{0.5} = 0.7071$

2. $X =$ Roll on a 6-sided die

x	1	2	3	4	5	6
$P(X=x)$	1/6	1/6	1/6	1/6	1/6	1/6

$$\text{Mean: } \mu = E[X] = \sum_x x \cdot P(x) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \frac{21}{6} = 3.5$$

This should make sense... over the long-term, we should have an average of about 3.5 rolled on the die.

$$\text{Variance: } \sigma^2 = \text{Var}[X] = \sum_x (x - \mu)^2 \cdot P(x) = (1 - 3.5)^2 \left(\frac{1}{6}\right) + (2 - 3.5)^2 \left(\frac{1}{6}\right) + \dots + (6 - 3.5)^2 \left(\frac{1}{6}\right) = 2.9167$$

OR

$$E[X^2] = \sum_x x^2 \cdot P(x) = 1^2 \left(\frac{1}{6}\right) + 2^2 \left(\frac{1}{6}\right) + 3^2 \left(\frac{1}{6}\right) + 4^2 \left(\frac{1}{6}\right) + 5^2 \left(\frac{1}{6}\right) + 6^2 \left(\frac{1}{6}\right) = \frac{91}{6} = 15.1667$$

$$\text{Variance: } \sigma^2 = \text{Var}[X] = E[X^2] - E[X]^2 = 15.1667 - 3.5^2 = 15.1667 - 12.25 = 2.9167$$

$$\text{Standard Deviation: } \sigma = \sqrt{\text{Var}[X]} = \sqrt{2.9167} = 1.7078$$

3. Suppose you want to play the Pick 3 lottery. You choose any number from 000 to 999 (1000 numbers total) and pay \$1. If your number is selected, you win \$750. Otherwise you win nothing. Let $X =$ Profit made playing one game of the lottery.

To find the probability distribution, notice that you can either win or lose. $P(\text{Win}) = 1/1000 = 0.001$, and the profit made in that case is $\$750 - \$1 = \$749$. $P(\text{Lose}) = 999/1000 = 0.999$, and profit made in that case is $-\$1$. So we have:

x	-1	749
$P(x)$	0.999	0.001

$$\text{Mean: } \mu = E[X] = \sum_x x \cdot P(x) = -1(0.999) + 749(0.001) = -0.25$$

So over the long run, you would expect to lose about 25 cents per play.

$$\begin{aligned} \text{Variance: } \sigma^2 = \text{Var}[X] &= \sum_x (x - \mu)^2 \cdot P(x) = (-1 - (-0.25))^2 (0.999) + (749 - (-0.25))^2 (0.001) \\ &= (-0.75)^2 (0.999) + (749.25)^2 (0.001) = 561.9375 \end{aligned}$$

OR

$$E[X^2] = \sum_x x^2 \cdot P(x) = (-1)^2 (0.999) + 749^2 (0.001) = 562$$

$$\text{Variance: } \sigma^2 = \text{Var}[X] = E[X^2] - E[X]^2 = 562 - (-0.25)^2 = 562 - 0.0625 = 561.9375$$

$$\text{Standard Deviation: } \sigma = \sqrt{\text{Var}[X]} = \sqrt{561.9375} = 23.7052$$

4. Try this one on your own... answers are provided.

x	0	2	4	6
$P(x)$	0.1	0.3	0.2	0.4

Mean: 3.8

Variance: 4.36

Standard Deviation: 2.09

Unusual Results

Recall that before, we said an “unusual observation” had a value falling more than 2 standard deviations away from the mean. Now, we define an *unusual value* of X to be a value falling more than 2 standard deviations away from the mean of the probability distribution.

To determine if a set of results is unusual, suppose the number of successes (times an event occurred) in n trials is x .

If $P(X \geq x)$ is smaller than 0.05, we say there was an *unusually high number of successes*.

If $P(X \leq x)$ is smaller than 0.05, we say there was an *unusually low number of successes*.

In other words, if the chances of getting a value that high or that low are very small, then the results are unusually high or low, respectively.

Example:

Suppose you flip a coin 8 times, and get 6 heads. Is this an unusually high number of heads?

The probability of getting at least 6 heads in 8 flips can be computed to be $P(X \geq 6) = 0.144$. Therefore, we conclude that 6 heads on 8 coin flips is not unusual.

Now suppose you flip a coin 8 times again, and get 1 head. Is this an unusually low number of heads?

The probability of getting at most 1 head in 8 flips can be computed to be $P(X \leq 1) = 0.035$. Since this is smaller than 0.05, we conclude that this is an unusually small number of heads.

Section 5-3: Binomial Probability Distributions

The previous section described the general concept of a discrete probability distribution. However, one type of distribution appears frequently in applications and is worth studying by itself. It is called a *binomial distribution*. A binomial probability distribution has a number of characteristics. Any application that shares these characteristics can be described using the distribution, which is why it is so powerful.

Characteristics of the Binomial Distribution

1. The procedure has n repeated *independent* trials.
2. Each trial has two outcomes: success or failure.
3. Each trial has the same probability of success p .

Examples:

X = Number of heads in 5 flips of a coin ($n = 5, p = 0.5$)

X = Number of defective units in a sample of 100 units, where $P(\text{defective}) = 0.01$ ($n = 100, p = 0.01$)

(Note: a success doesn't necessarily have to be a good thing to happen)

X = Number of U.S. citizens supporting a particular candidate for president in a sample of 1000, where 70% of americans support the candidate ($n = 1000, p = 0.70$)

The Binomial Distribution Formula

To develop a formula to find $P(X = x)$, let's consider an example.

Ted is at an amusement park and decides to play one of the carnival games. In this game, he is trying to throw a beanbag into a hole in the wall. Suppose he has a 30% chance of getting the beanbag into the hole each game. Let X be the number of times Ted wins (gets the beanbag in the hole) out of 4 games. Hopefully, it is clear that X will follow a binomial distribution, with $n = 4$ and $p = 0.30$. Let's try to figure out the probability that he wins 1 out of the 4 games (i.e. $P(X = 1)$).

To do this, we need him to win one game (the probability of winning is 0.30) and lose 3 games (the probability of losing is 0.70). So by the multiplication rule for independent events, the chance of this happening is $0.30 \cdot 0.70 \cdot 0.70 \cdot 0.70 = 0.1029$. But there is one other thing to consider – namely that this can happen in 4 different ways:

- 1st Way: Win Lose Lose Lose
- 2nd Way: Lose Win Lose Lose
- 3rd Way: Lose Lose Win Lose
- 4th Way: Lose Lose Lose Win

Each one has probability 0.1029, so the overall probability is $P(X = 1) = 4 \cdot 0.1029 = 0.4116$.

In general, for n games, x wins, and chance of winning p , we could generalize the formula for $P(X = x)$.

We need Ted to win x times and lose $n - x$ times, which is $p^x(1 - p)^{n-x}$. Then we have to multiply that by the number of ways you can win x times out of n games. It turns out this is given by the formula:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}, \text{ where } n! = n(n-1)(n-2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$$

For example, in the case discussed above, we had $n = 4$ and $x = 1$, so $\binom{4}{1} = \frac{4!}{1!3!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{1 \cdot 3 \cdot 2 \cdot 1} = 4$

Thus, for a general binomial distribution, we obtain:

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

Why is This a Probability Distribution?

In case you are wondering, this actually is a probability distribution. To show that this is the case, we can use a theorem from high

school algebra: $(a + b)^n = \sum_{x=1}^n \binom{n}{x} a^x b^{n-x}$. This theorem is called the Binomial Formula, which is where the binomial distribution

gets its name. Notice that the items in the sum look just like the probabilities in the binomial distribution. Therefore, we have:

$$1. \sum_x P(x) = \sum_{x=1}^n \binom{n}{x} p^x (1-p)^{n-x} = (p + (1-p))^n = 1^n = 1$$

2. Clearly, $0 \leq P(x)$ because all the terms are positive. Combining this fact with knowledge that the sum of the $P(x)$ terms is 1 tells us that $P(x) \leq 1$ as well.

Therefore, we see that the binomial distribution is a probability distribution.

How to Compute Binomial Probabilities

Table A-1 in Appendix A gives the binomial probabilities for various values of n , p , and x . Using this table, it is easy to find desired probabilities that X takes on certain values. (Note that probabilities smaller than 0.0005 appear as 0+ in the table)

Example:

In a manufacturing process, the probability of a unit being defective is 0.05. Suppose we sample 10 of these units. Find the probability that...

- Exactly 3 of the units are defective.
- Less than 2 units are defective.
- At least 1 unit is defective.

Solutions:

First note that in this situation, $n = 10$ and $p = 0.05$

a. From the table, under $n = 10$, $p = 0.05$, and $x = 3$, we find that $P(X = 3) = \underline{0.010}$

b. Less than 2 units defective means X is less than 2. So, looking at the table for $x = 0$ and $x = 1$, we get:

$$P(X < 2) = P(X = 0) + P(X = 1) = 0.599 + 0.315 = \underline{0.914}$$

c. At least 1 unit is defective means that X is greater than or equal to 1. There are two possibilities for computing this.

The hard way is to add up the probabilities for 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. However, half of these probabilities are listed as 0+, so we cannot know exactly what value they have.

The easier way is to use the complement rule. $P(X \geq 1) = 1 - P(X = 0) = 1 - 0.599 = \underline{0.401}$. Also, this result will be more accurate.

Mean and Variance of the Binomial Distribution

If $P(\text{Success}) = 0.30$, and we have 10 trials, how many successes would you expect? Probably 3. This intuition is true in the general case with n trials and $P(\text{Success}) = p$.

If X follows a Binomial Distribution with n trials and probability of success p , then:

$$E[X] = np$$

$$\text{Var}[X] = np(1 - p)$$