

# CHAPTER 6

## Random Variables

6.3a

### Binomial and Geometric Random Variables

The Practice of Statistics, 5th Edition  
Starnes, Tabor, Yates, Moore



# Binomial and Geometric Random Variables

---

## Learning Objectives

After this section, you should be able to:

- ✓ **DETERMINE** whether the conditions for using a binomial random variable are met.
- ✓ **COMPUTE** and **INTERPRET** probabilities involving binomial distributions.
- ✓ **CALCULATE** the mean and standard deviation of a binomial random variable. **INTERPRET** these values in context.
- ✓ **FIND** probabilities involving geometric random variables.

# Binomial Settings

---

When the same chance process is repeated several times, we are often interested in whether a particular outcome does or doesn't happen on each repetition. Some random variables count the number of times the outcome of interest occurs in a fixed number of repetitions. They are called **binomial random variables**.

A **binomial setting** arises when we perform several independent trials of the same chance process and record the number of times that a particular outcome occurs. The four conditions for a binomial setting are:

- B** Binary? The possible outcomes of each trial can be classified as “success” or “failure.”
- I** Independent? Trials must be independent; that is, knowing the result of one trial must not tell us anything about the result of any other trial.
- N** Number? The number of trials  $n$  of the chance process must be fixed in advance.
- S** Success? There is the Same probability  $p$  of success on each trial.

# Binomial Random Variables

---

Consider tossing a coin  $n$  times. Each toss gives either heads or tails. Knowing the outcome of one toss does not change the probability of an outcome on any other toss.

If we define heads as a success, then  $p$  is the probability of a head and is 0.5 on any toss.

The number of heads in  $n$  tosses is a **binomial random variable  $X$** . The probability distribution of  $X$  is called a **binomial distribution**.

The count  $X$  of successes in a binomial setting is a **binomial random variable**. The probability distribution of  $X$  is a **binomial distribution** with parameters  $n$  and  $p$ , where  $n$  is the number of trials of the chance process and  $p$  is the probability of a success on any one trial. The possible values of  $X$  are the whole numbers from 0 to  $n$ .

## *Dice, hoops, and cars*

---

Determine whether the random variables below have a binomial distribution. Justify your answer.

- (a) Roll a fair die 10 times and let  $X$  = the number of sixes.
- (b) Shoot a basketball 20 times from various distances on the court. Let  $Y$  = number of shots made.
- (c) Observe the next 100 cars that go by and let  $C$  = color.

***Roll a fair die 10 times and let  $X$  = the number of sixes***

*Binary?* Yes; success = six, failure = not a six.

*Independent?* Yes; knowing the outcomes of past rolls tells you nothing about the outcomes of future rolls.

*Number?* Yes; there are  $n = 10$  trials.

*Success?* Yes; the probability of success is always  $p = 1/6$ . This is a binomial setting.

The number of sixes  $X$  is a binomial random variable with  $n = 10$  and  $p = 1/6$ .       $B(10, 1/6)$

# *Dice, hoops, and cars*

---

**Shoot a basketball 20 times from various distances on the court.**

**Let  $Y$  = number of shots made.**

*Binary?* Yes; success = make the shot, failure = miss the shot.

*Independent?* Yes; evidence suggests that it is reasonable to assume that knowing the outcome of a shot tells us nothing about the outcome of other shots.

*Number?* Yes; there are  $n = 20$  trials.

*Success?* No; the probability of success changes because the shots are taken from various distances. Because the probability of success is not constant,  $Y$  is not a binomial random variable.

**Observe the next 100 cars that go by and let  $C$  = color.**

*Binary?* No. There are more than two possible colors.

*Independent?* Yes, knowing the color of one car tells you nothing about the color of other cars.

*Number?* Yes; there are  $n = 100$  trials.

*Success?* A success hasn't been defined, so we cannot determine if the probability of success is always the same. Because there are more than two possible outcomes,  $C$  is not a binomial random variable.

# Cards and dice

---

Determine whether the random variables below have a binomial distribution. Justify your answer.

- (a) Shuffle a deck of cards. Turn over the top card. Put the card back in the deck, and shuffle again. Repeat this process 10 times. Let  $X$  = the number of aces you observe.
- (b) Roll a fair die 100 times. Sometime during the 100 rolls, one corner of the die chips off. Let  $W$  = the number of 5s you roll.

## cards

*Binary?* Yes; success = get an ace, failure = don't get an ace.

*Independent?* Yes; Because you are replacing the card in the deck and shuffling each time, the result of one trial does not tell you anything about the outcome of any other trial.

*Number?* Yes; there are  $n = 10$  trials.

*Success?* Yes; the probability of success is always  $p = 4/52$ .

This is a binomial setting. The number of aces  $X$  is a binomial random variable with  $n = 10$  and  $p = 4/52$ .  $B(10, 4/52)$

# *dice*

---

**Roll a fair dice 100 times. Sometime during the 100 rolls, one corner of the die chips off. Let  $W$  = the number of 5s you roll.**

*Binary?* Yes; success = roll a 5, failure = don't roll a 5.

*Independent?* Yes; Because you are rolling a die, the outcome of any one trial does not tell you anything about the outcome of any other trial.

*Number?* Yes; there are  $n = 100$  trials.

*Success?* No; the probability of success changes when the corner of the die is chipped off.

If the corner didn't chip off?

$B(100, 1/6)$

# Binomial Probabilities

In a binomial setting, we can define a random variable (say,  $X$ ) as the number of successes in  $n$  independent trials. We are interested in finding the probability distribution of  $X$ .

Each child of a particular pair of parents has probability 0.25 of having type O blood. Genetics says that children receive genes from each of their parents independently. If these parents have 5 children, the count  $X$  of children with type O blood is a binomial random variable with  $n = 5$  trials and probability  $p = 0.25$  of a success on each trial. In this setting, a child with type O blood is a “success” (S) and a child with another blood type is a “failure” (F). What's  $P(X = 2)$ ?

$$P(\text{SSFFF}) = (0.25)(0.25)(0.75)(0.75)(0.75) = (0.25)^2(0.75)^3 = 0.02637$$

However, there are a number of different arrangements in which 2 out of the 5 children have type O blood:

<b>SSFFF</b>	<b>SFSFF</b>	<b>SFFSF</b>	<b>SFFFS</b>	<b>FSSFF</b>
<b>FSFSF</b>	<b>FSFFS</b>	<b>FFSSF</b>	<b>FFSFS</b>	<b>FFFSS</b>

Verify that in each arrangement,  $P(X = 2) = (0.25)^2(0.75)^3 = 0.02637$

$$\text{Therefore, } P(X = 2) = 10(0.25)^2(0.75)^3 = 0.2637$$

# Binomial Coefficient

Note, in the previous example, *any one arrangement* of 2 S' s and 3 F' s had the same probability. This is true because no matter what arrangement, we' d multiply together 0.25 twice and 0.75 three times.

We can generalize this for any setting in which we are interested in  $k$  successes in  $n$  trials. That is,

$$\begin{aligned}P(X = k) &= P(\text{exactly } k \text{ successes in } n \text{ trials}) \\ &= \text{number of arrangements} \times p^k (1 - p)^{n-k}\end{aligned}$$

The number of ways of arranging  $k$  successes among  $n$  observations is given by the **binomial coefficient**

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

for  $k = 0, 1, 2, \dots, n$  where

$$n! = n(n - 1)(n - 2) \cdot \dots \cdot (3)(2)(1)$$

and  $0! = 1$ .

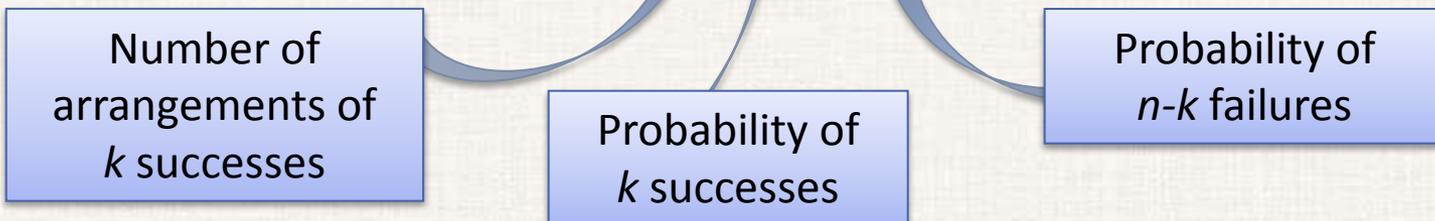
# Binomial Probability Formula

The binomial coefficient counts the number of different ways in which  $k$  successes can be arranged among  $n$  trials. The binomial probability  $P(X = k)$  is this count multiplied by the probability of any one specific arrangement of the  $k$  successes.

## Binomial Probability

If  $X$  has the binomial distribution with  $n$  trials and probability  $p$  of success on each trial, the possible values of  $X$  are  $0, 1, 2, \dots, n$ . If  $k$  is any one of these values,

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



# How to Find Binomial Probabilities

---

## How to Find Binomial Probabilities

**Step 1: State the distribution and the values of interest.** Specify a binomial distribution with the number of trials  $n$ , success probability  $p$ , and the values of the variable clearly identified.

**Step 2: Perform calculations—show your work!**

Do one of the following:

- (i) Use the binomial probability formula to find the desired probability; or
- (ii) Use `binompdf` or `binomcdf` command and label each of the inputs.

**Step 3: Answer the question.**

# Example: How to Find Binomial Probabilities

Each child of a particular pair of parents has probability 0.25 of having blood type O. Suppose the parents have 5 children

(a) Find the probability that exactly 3 of the children have type O blood.

Let  $X$  = the number of children with type O blood. We know  $X$  has a binomial distribution with  $n = 5$  and  $p = 0.25$ . ***binompdf(5, 0.25, 3)***

$$P(X = 3) = \binom{5}{3} (0.25)^3 (0.75)^2 = 10(0.25)^3 (0.75)^2 = 0.08789$$

(b) Should the parents be surprised if more than 3 of their children have type O blood? To answer this, we need to find  $P(X > 3)$ .

$$P(X > 3) = P(X = 4) + P(X = 5)$$

***1 - binomcdf(5, 0.25, 3)***

$$\begin{aligned} &= \binom{5}{4} (0.25)^4 (0.75)^1 + \binom{5}{5} (0.25)^5 (0.75)^0 \\ &= 5(0.25)^4 (0.75)^1 + 1(0.25)^5 (0.75)^0 \\ &= 0.01465 + 0.00098 = 0.01563 \end{aligned}$$

Since there is only a 1.5% chance that more than 3 children out of 5 would have Type O blood, the parents should be surprised!

# Rolling doubles

---

When rolling two dice, the probability of rolling doubles is  $1/6$ . Suppose that a game player rolls the dice 4 times, hoping to roll doubles.

- (a) Find the probability that the player gets doubles twice in 4 attempts.
- (b) Should the players be surprised if he gets doubles more than twice in 4 attempts?

Let  $X$  = number of doubles.  $X$  has binomial distribution with  $n=4$  and  $p=\frac{1}{6}$

$$P(X = 2) = \binom{4}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 = \text{binompdf}(\text{trials: } 4, \text{ p: } 1/6, \text{ x value: } 2) = .116$$

There is an 11.6% chance of getting exactly two doubles in 4 attempts.

$$\begin{aligned} P(X > 2) &= P(X = 3) + P(X = 4) = \binom{4}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right) + \binom{4}{4} \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^0 \\ &= 1 - P(X \leq 2) = 1 - \text{binomcdf}(\text{trials: } 4, \text{ p: } 1/6, \text{ x value: } 2) = 0.016 \end{aligned}$$

Because there is only a 1.6% chance of getting more than two doubles in 4 rolls, the player should be surprised if this happens.

## *The last Kiss*

---

Do people have a preference for the last thing they taste? Researchers at the University of Michigan designed a study to find out. The researchers gave 22 students five different Hershey's Kisses (milk chocolate, dark chocolate, crème, caramel, and almond) in random order and asked the student to rate each one. Participants were not told how many Kisses they would be tasting. However, when the 5<sup>th</sup> and final Kiss was presented, participants were told that it would be their last one. Of the 22 students, 14 of them gave the final Kiss the highest rating.

Assume that the participants in the study don't have a special preference for the last thing they taste. That is, assume that the probability a person would prefer the last Kiss tasted is  $p = 0.20$ .

- (a) What is the probability that exactly 5 of the 22 participants would prefer the last Kiss they tried?
- (b) What is the probability that 14 or more of the 22 participants would prefer the last Kiss they tried?

# The last Kiss...results

---

Remember ALL three steps.

1. State the distribution and value of interest.
2. Perform calculations showing work.
3. Answer the question.

(a) Let  $X$  = the number of participants who prefer the last Kiss they taste.  $X$  has binomial distribution with  $n = 22$  and  $p = 0.2$ .

$$P(X = 5) = \binom{22}{5} (0.2)^5 (0.8)^{17} = \text{binompdf}(\text{trials: } 22, p: 0.2, x \text{ value: } 5) = 0.1898$$

There is about a 19% chance that exactly 5 participants would choose the last Kiss, assuming they have no special preference for the last thing they taste.

$$\begin{aligned} (b) P(X \geq 14) &= 1 - P(X \leq 13) = 1 - \text{binomcdf}(\text{trials: } 22, p: 0.2, x \text{ value } 13) \\ &= 1.01...E-5 = 0.00001 \end{aligned}$$

There is about a 0.001% chance that 14 or more participants would choose the last Kiss, assuming that they had no special preference for the last thing they taste. Because the probability is so small, there is convincing evidence that the participants have a preference for the last thing they taste. It is almost impossible to get 14 or more just by chance.

# Binomial and Geometric Random Variables

## Section Summary

In this section, we learned how to...

- ✓ **DETERMINE** whether the conditions for using a binomial random variable are met.
- ✓ **COMPUTE** and **INTERPRET** probabilities involving binomial distributions.
- ✓ **CALCULATE** the mean and standard deviation of a binomial random variable. **INTERPRET** these values in context.
- ✓ **FIND** probabilities involving geometric random variables.
- ✓ Read p. 386-396 ccc 63, 65, 66, 69, 71, 73, 75, 77 (first 3 from 6.2)