

6.5

The Normal Approximation to the Binomial Distribution

Background

The normal distribution has many applications and is used extensively in statistical inference. In addition, many distributions, or populations, are approximately normal. The Central Limit Theorem, which is introduced in Section 7.2, provides some theoretical justification for this empirical evidence. In these cases, the normal distribution may be used to compute approximate probabilities.

Recall that the binomial distribution was defined in Section 5.4. Although it seems strange, under certain circumstances a (continuous) normal distribution can be used to approximate a (discrete) binomial distribution.

Suppose X is a binomial random variable with $n = 10$ and $p = 0.5$. The probability histogram for this distribution is shown in **Figure 6.70**. In this case, the area of each rectangle represents the probability associated with a specific value of the random variable. For example, in **Figure 6.71** the total area of the rectangular shaded regions is $P(4 \leq X \leq 7)$.

For each rectangle, the width is 1 and the height is equal to the probability associated with a specific value. Therefore, the area of each rectangle (width \times height) also represents probability.

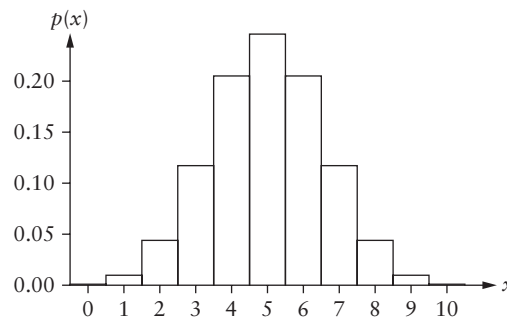


Figure 6.70 Probability histogram associated with a binomial random variable defined by $n = 10$ and $p = 0.5$.

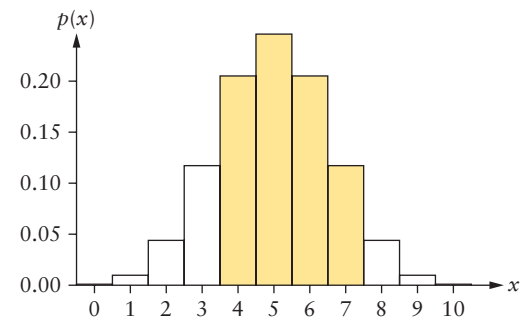


Figure 6.71 The area of the shaded rectangles is $P(4 \leq X \leq 7)$.

The probability histogram corresponding to the binomial distribution in Figure 6.70 is approximately bell-shaped. Consider a normal random variable with a mean and variance from the corresponding binomial distribution. That is, $\mu = np = (10)(0.5) = 5$ and $\sigma^2 = np(1-p) = (10)(0.5)(0.5) = 2.5$. The graph of the probability density function for this normal random variable and the probability histogram for the original binomial random variable are shown in **Figure 6.72**.

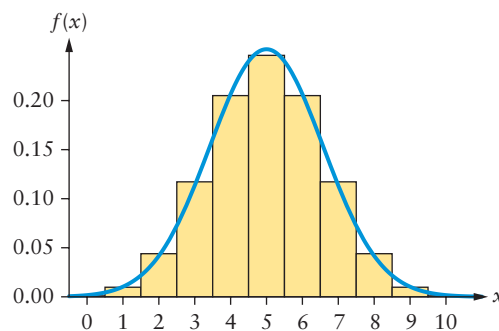


Figure 6.72 Probability histogram and probability density function.

The graph of the probability density function appears to pass through the middle top of each rectangle and is a smooth, continuous approximation to the probability histogram.

The probability of an event involving a normal random variable is associated with the area under the graph of the density function. In this example, the total area under the probability density function seems to be fairly close to the total area of the probability histogram rectangles. Therefore, it seems reasonable to use the normal random variable to approximate probabilities associated with the binomial random variable. There is still one minor issue, but it is easily addressed. Here is an example to illustrate the approximation process.

Approximation and the Continuity Correction

EXAMPLE 6.12 Exact Versus Approximate Probability

Suppose X is a binomial random variable with 10 trials and probability of success 0.5. That is, $X \sim B(10, 0.5)$.

- (a) Find the exact probability, $P(4 \leq X \leq 7)$.
 (b) Use the normal approximation to find $P(4 \leq X \leq 7)$.

Solution

- (a) Use the techniques introduced in Section 5.4.

$$\begin{aligned} P(4 \leq X \leq 7) &= P(X \leq 7) - P(X \leq 3) && \text{Use cumulative probability.} \\ &= 0.9453 - 0.1719 && \text{Use Appendix Table 1.} \\ &= 0.7734 && \text{Simplify.} \end{aligned}$$

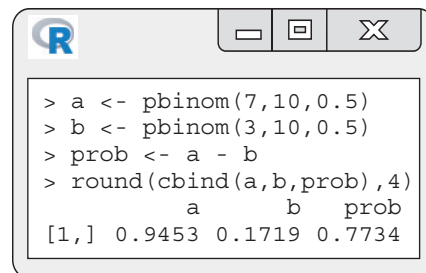
- (b) For $X \sim B(10, 0.5)$, $\mu = 5$ and $\sigma^2 = 2.5$.

Figure 6.72 suggests that $X \overset{\sim}{\sim} N(5, 2.5)$. Using this approximation,

$$\begin{aligned} P(4 \leq X \leq 7) &\approx P(4 \leq X \leq 7) \\ &\quad \uparrow \qquad \qquad \uparrow \\ &\quad \text{Binomial RV} \quad \text{Normal RV} \\ &= P\left(\frac{4-5}{\sqrt{2.5}} \leq \frac{X-5}{\sqrt{2.5}} \leq \frac{7-5}{\sqrt{2.5}}\right) && \text{Standardize.} \\ &= P(-0.63 \leq Z \leq 1.26) && \text{Use Equation 6.8; simplify.} \\ &= P(Z \leq 1.26) - P(Z \leq -0.63) && \text{Use cumulative probability.} \\ &= 0.8962 - 0.2643 && \text{Use Appendix Table 3.} \\ &= 0.6319 \end{aligned}$$

Figure 6.73 and Figure 6.74 show technology solutions.

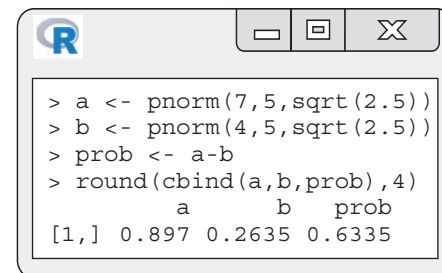
The symbol $\overset{\sim}{\sim}$ means is approximately distributed as.



```

R
> a <- pbinom(7,10,0.5)
> b <- pbinom(3,10,0.5)
> prob <- a - b
> round(cbind(a,b,prob),4)
      a      b  prob
[1,] 0.9453 0.1719 0.7734
  
```

Figure 6.73 Probability calculations associated with the binomial random variable.



```

R
> a <- pnorm(7,5,sqrt(2.5))
> b <- pnorm(4,5,sqrt(2.5))
> prob <- a-b
> round(cbind(a,b,prob),4)
      a      b  prob
[1,] 0.897 0.2635 0.6335
  
```

Figure 6.74 Probability calculations associated with the normal random variable.

This approximation isn't very close to the exact probability, and Figure 6.72 suggests a reason for the difference. By finding the area under the normal probability density function from 4 to 7, we have left out half of the area of the rectangle above 4 and half of the area of the rectangle above 7.

The width of each rectangle is 1, so the edges of a rectangle above x are $x - 0.5$ and $x + 0.5$. For example, the rectangle above 4 has edges 3.5 and 4.5.

To capture the entire area of the rectangles above 4 and 7, we find the area under the normal probability density function from 3.5 to 7.5. This is called the *continuity correction*. We correct a continuous probability approximation associated with a discrete random variable (binomial) by adding or subtracting 0.5 for each endpoint, so as to capture all the area of the corresponding rectangles. **Figure 6.75** and **Figure 6.76** illustrate the continuity correction in this example.

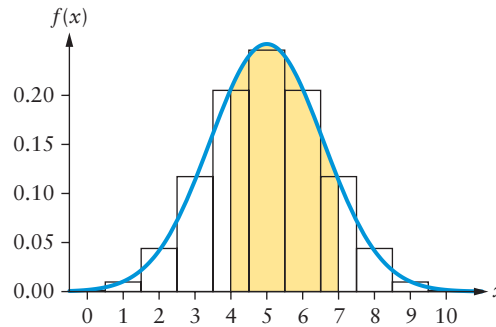


Figure 6.75 This approximation excludes half of the area of the rectangles above 4 and 7.

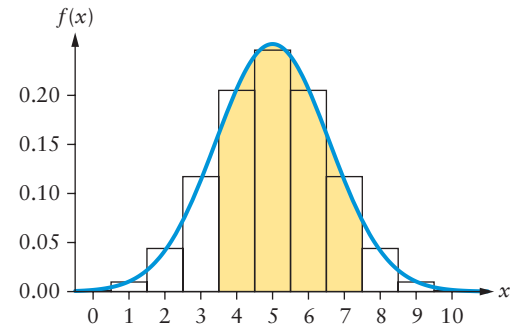


Figure 6.76 This approximation includes more of the area/probability associated with 4 and 7.

Using the continuity correction,

$$P(4 \leq X \leq 7) \approx P(3.5 \leq X \leq 7.5)$$

\uparrow \uparrow
 Binomial RV Normal RV

$$\begin{aligned}
 &= P\left(\frac{3.5-5}{\sqrt{2.5}} \leq \frac{X-5}{\sqrt{2.5}} \leq \frac{7.5-5}{\sqrt{2.5}}\right) \\
 &= P(-0.95 \leq Z \leq 1.58) \\
 &= P(Z \leq 1.58) - P(Z \leq -0.95) \\
 &= 0.9429 - 0.1711 \\
 &= 0.7718
 \end{aligned}$$

Standardize.

Use Equation 6.8; simplify.

Use cumulative probability.

Use Appendix Table 3.

This probability found using the continuity correction, 0.7718, is much closer to the true probability than the value computed in part (b).

Figure 6.77 shows a technology solution using the continuity correction.

```

R
> a <- pnorm(7.5,5,sqrt(2.5))
> b <- pnorm(3.5,5,sqrt(2.5))
> prob <- a-b
> round(cbind(a,b,prob),4)
      a      b      prob
[1,] 0.9431 0.1714 0.7717
  
```

Figure 6.77 The normal approximation using the continuity correction.

Note: A round-off error is introduced in the preceding calculations, in the standardization process. The technology results are more accurate and better illustrate the impact of the continuity correction.

TRY IT NOW Go to Exercises 6.151 and 6.153

Summary and Application

The previous example suggests the following approximation procedure.

The Normal Approximation to the Binomial Distribution

Suppose X is a binomial random variable with n trials and probability of success p , $X \sim B(n, p)$. If n is large and both $np \geq 10$ and $n(1-p) \geq 10$, then the random variable X is approximately normal with mean $\mu = np$ and variance $\sigma^2 = np(1-p)$, $X \sim N(np, np(1-p))$.



A CLOSER LOOK

1. There is no threshold value for n . However, a large sample isn't enough for approximate normality. The two products np and $n(1-p)$ must both be greater than or equal to 10. This is called the nonskewness criterion. If both inequalities are satisfied, then it is reasonable to assume that the distribution of the approximate normal distribution is centered far enough away from 0 or 1, and with the values $\mu \pm 3\sigma$ inside the interval $[0, 1]$.
2. When using the continuity correction, think carefully about whether to include or exclude the area of each rectangle associated with the binomial probability histogram. Consider creating a figure to illustrate the probability question. This will help you decide whether to add or subtract 0.5.
3. For n very large (e.g., $n \geq 5000$), probability calculations involving the exact binomial distribution can be very long and inaccurate, even when using technology. In this case, the normal approximation to the binomial distribution provides a reliable, quick alternative.

EXAMPLE 6.13 Telephone Spam

We've all had those annoying cell phone calls from Heather, Daisy, or Oscar, trying to sell life insurance or an additional car warranty. First Orion, a call protection agency, recently issued a report that suggested 45% of all cell phone calls in 2019 will be spam.¹ Suppose 500 cell phone calls are selected at random. Use the normal approximation to the binomial distribution to answer the following questions.

- (a) Find the probability that at least 240 of the cell phone calls will be spam.
- (b) Find the probability that at exactly 245 of the cell phone calls will be spam.
- (c) Suppose 200 of the cell phone calls are spam. Is there any evidence to suggest that the proportion of cell phone calls that are spam is less than 0.45? Justify your answer.

Solution

- (a) Let X be the number of cell phone calls that are spam. X is a binomial random variable with $n = 500$ and $p = 0.45$, $X \sim B(500, 0.45)$. Use Equation 5.8 to find the mean and variance of the binomial random variable X .

$$\mu = np = (500)(0.45) = 225$$

$$\sigma^2 = np(1-p) = (500)(0.45)(0.55) = 123.75$$

For $n = 500$ and $p = 0.45$, check the nonskewness criterion.

$$np = (500)(0.45) = 225 \geq 10 \quad \text{and} \quad n(1-p) = (500)(0.55) = 275 \geq 10$$

Both inequalities are satisfied. The distribution of X is approximately normal with $\mu = 225$ and $\sigma^2 = 123.75$: $X \sim N(225, 123.75)$.

The probability that at least 240 cell phone calls will be spam is

$$P(X \geq 240) \approx P(X \geq 239.5)$$

↑ ↑

Binomial RV Normal RV

$$= P\left(\frac{X - 225}{\sqrt{123.75}} \geq \frac{239.5 - 225}{\sqrt{123.75}}\right)$$

$$= 1 - P(Z \leq 1.30)$$

$$= 1 - 0.9032$$

$$= 0.0968$$

Use the normal approximation with a continuity correction.

Standardize.

Simplify; use cumulative probability.

Use Appendix Table 3.

Simplify.

The probability that at least 240 cell phone calls will be spam is approximately 0.0968. **Figure 6.78** shows part of the graph of the binomial probability histogram and approximate normal probability density function. We need to include the area of the rectangle corresponding to 240. Using the continuity correction with the normal approximation, start at $240 - 0.5 = 239.5$.

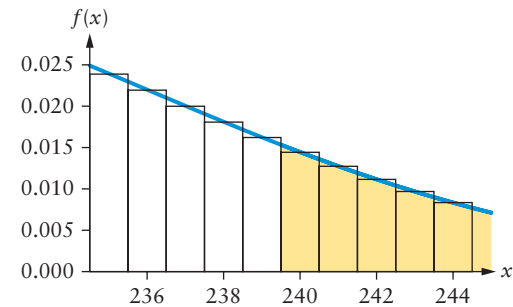


Figure 6.78 Using the normal approximation, start at 239.5.

Figure 6.79 shows a technology solution.

```

R
> prob <- pnorm(239.5, 225, sqrt(123.75), lower.tail=FALSE)
> round(prob, 4)
[1] 0.0962

```

Figure 6.79 Probability calculation using the normal approximation.

(b) The probability that exactly 245 cell phone calls are spam is

$$P(X = 245) \approx P(244.5 \leq X \leq 245.5)$$

↑ ↑

Binomial RV Normal RV

$$= P\left(\frac{244.5 - 225}{\sqrt{123.75}} \leq \frac{X - 225}{\sqrt{123.75}} \leq \frac{245.5 - 225}{\sqrt{123.75}}\right)$$

$$= P(1.75 \leq Z \leq 1.84)$$

$$= P(Z \leq 1.84) - P(Z \leq 1.75)$$

$$= 0.9671 - 0.9599$$

$$= 0.0072$$

Use the normal approximation with a continuity correction.

Standardize.

Use Equation 6.8; simplify.

Use cumulative probability.

Use Appendix Table 3.

The probability that exactly 245 cell phone calls are spam is approximately 0.0072. **Figure 6.80** shows part of the graph of the binomial probability histogram and approximate normal probability density function. To approximate the area of the single rectangle corresponding to 245, use the endpoints of the rectangle.

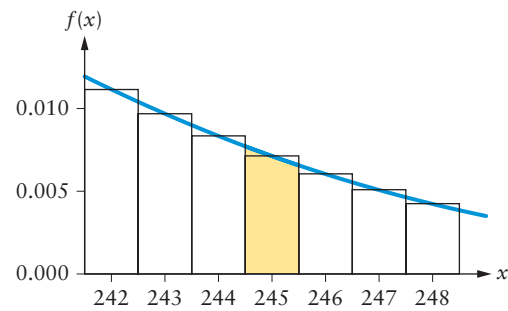


Figure 6.80 Use the continuity correction to approximate the area of a single rectangle.

Figure 6.81 shows a technology solution.

```

R
> mu <- 225
> sigma <- sqrt(123.75)
> a <- pnorm(245.5, mu, sigma)
> b <- pnorm(244.5, mu, sigma)
> prob <- a - b
> round(cbind(a, b, prob), 4)
      a      b  prob
[1,] 0.9673 0.9602 0.0071

```

Figure 6.81 Probability calculations using the normal approximation.

- (c) The claim made in the report is that $p = 0.45$. This implies that the random variable X has a binomial distribution with $n = 500$ and $p = 0.45$. Using the normal approximation to the binomial distribution, X has approximately a normal distribution with $\mu = (500)(0.45) = 225$ and $\sigma^2 = (500)(0.45)(0.55) = 123.75$.

Claim: $p = 0.45 \Rightarrow \mu = 225 \Rightarrow X \sim N(225, 123.65)$

The experimental outcome is that 200 cell phone calls were spam.

Experiment: $x = 200$

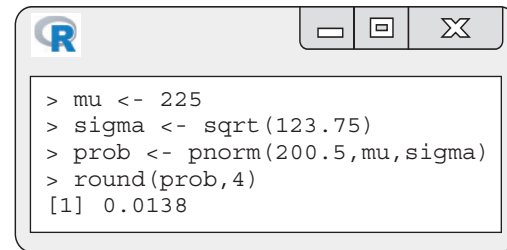
Because $x = 200$ is to the left of the mean, and because we are searching for evidence that the proportion of cell phone calls is less than 0.45, we consider a left-tail probability.

Likelihood:

$$\begin{aligned}
 P(X \leq 200) &\approx P(X \leq 200.5) && \text{Use the normal approximation} \\
 &\uparrow \qquad \qquad \uparrow && \text{with a continuity correction.} \\
 \text{Binomial RV} & \quad \text{Normal RV} \\
 &= P\left(\frac{X - 225}{\sqrt{123.75}} \leq \frac{200.5 - 225}{\sqrt{123.75}}\right) && \text{Standardize.} \\
 &= P(Z \leq -2.20) && \text{Use Equation 6.8; simplify.} \\
 &= 0.0139 && \text{Use Appendix Table 3.}
 \end{aligned}$$

Conclusion: Because the tail probability is so small (less than 0.05), $x = 200$ is considered to be a very unusual observation (if $p = 0.45$). Therefore, there is evidence to suggest that the claim is wrong, that $p < 0.45$ —that is, the proportion of cell phone calls that are spam is less than 0.45.

Figure 6.82 shows a technology solution.



```

> mu <- 225
> sigma <- sqrt(123.75)
> prob <- pnorm(200.5, mu, sigma)
> round(prob, 4)
[1] 0.0138

```

Figure 6.82 Probability calculations using the normal approximation.

TRY IT NOW Go to Exercises 6.155 and 6.157

Section 6.5 Exercises

Concept Check

6.142 True or False The normal approximation to the binomial distribution can be used for any values of n and p .

6.143 True or False Using the normal approximation to a binomial distribution, the probability of a single value is always 0.

6.144 True or False To use the normal approximation to the binomial distribution, the number of trials n must be at least 30.

6.145 True or False The normal approximation to the binomial distribution always provides an overestimate of the exact probability.

6.146 Short Answer If $X \sim B(n, p)$, find the mean and variance of the approximate normal random variable.

6.147 Short Answer When using the normal approximation to the binomial distribution, explain why the continuity correction improves the approximation.

Practice

6.148 Suppose X is a binomial random variable with n trials and probability of success p . Check the nonskewness criterion and find the distribution of the corresponding approximate normal random variable.

- (a) $n = 30, p = 0.40$
- (b) $n = 85, p = 0.55$
- (c) $n = 340, p = 0.38$
- (d) $n = 605, p = 0.75$

6.149 Suppose X is a binomial random variable with $n = 25$ trials and probability of success $p = 0.45$.

- (a) Check the nonskewness criterion.
- (b) Find the distribution of the corresponding approximate normal random variable.
- (c) Carefully sketch the probability histogram for the binomial random variable and the probability density function for the normal random variable on the same coordinate axes.

6.150 Suppose X is a binomial random variable with $n = 100$ trials and probability of success $p = 0.03$.

(a) Check the nonskewness criterion.

(b) Compute the values $\mu \pm 3\sigma$.

(c) Find the distribution of the corresponding approximate normal random variable. Using the results from parts (a) and (b), is the normal distribution a good approximation to the binomial distribution? Justify your answer.

(d) Carefully sketch a portion of the probability histogram for the binomial random variable and the probability density function for the normal random variable near 0 on the same coordinate axes. Explain geometrically why the approximation is or is not accurate.

6.151 Suppose X is a binomial random variable with $n = 25$ trials and probability of success $p = 0.60$. Find each probability using the binomial random variable, the approximate normal random variable without the continuity correction, and the approximate normal random variable with the continuity correction.

- (a) $P(X \leq 14)$
- (b) $P(X < 14)$
- (c) $P(16 \leq X \leq 19)$
- (d) $P(X > 10)$

6.152 Suppose X is a binomial random variable with $n = 600$ trials and probability of success $p = 0.35$. Find the probability using the normal approximation to the binomial distribution with a continuity correction.

- (a) $P(X > 220)$
- (b) $P(X \leq 198)$
- (c) $P(190 < X < 200)$
- (d) $P(X = 212)$

Applications

6.153 Fuel Consumption and Cars Digital Radio UK recently reported that 55% of all new cars sold in the United Kingdom are equipped with a digital radio. Suppose a random sample of 100 new car purchases in the United Kingdom is obtained. Answer each question using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that at least 60 new cars are equipped with a digital radio.

- (b) Find the approximate probability that fewer than 57 new cars are equipped with a digital radio.
- (c) Find the approximate probability that between 45 and 55 (inclusive) new cars are equipped with a digital radio.

6.154 Technology and the Internet Many companies are researching ways in which drones could help improve their business. For example, Amazon could use drones to deliver packages more quickly and Domino's could use drones to get pizza to your door faster and hotter. Despite the many possibilities, a recent survey indicated that 54% of Americans believe that personal and commercial drones should not be allowed to fly near people's homes.² Suppose 160 Americans are selected at random. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that fewer than 80 Americans believe that drones should not be allowed to fly near people's homes.
- (b) Find the approximate probability that at least 100 Americans believe that drones should not be allowed to fly near people's homes.
- (c) Find the approximate probability that exactly 92 Americans believe that drones should not be allowed to fly near people's homes.

6.155 Education and Child Development According to the Nation's Report Card, also known as the National Assessment of Educational Progress (NAEP), only 25% of 12th-grade students are proficient in mathematics.³ Suppose 200 U.S. 12th-graders are selected at random. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that at least 55 12th-graders are proficient in mathematics.
- (b) Find the approximate probability that between 60 and 65 (inclusive) 12th-graders are proficient in mathematics.
- (c) Suppose the NAEP test results for each student are used to find that 36 (of the 200) 12th-grade students are proficient in mathematics. Is there any evidence to suggest that fewer than 25% of 12th-graders are proficient in mathematics? Justify your answer.

6.156 Many science conferences and research papers have focused on global warming in recent years. Several climate model predictions include higher surface temperatures, rising sea levels, and larger subtropical deserts. Global leaders continue to discuss possible actions to stop or slow these trends. In a recent survey, only 54% of Americans said that global warming should be a high priority for political leaders and governments.⁴ Suppose 250 Americans are selected at random and each is asked if global warming should be a priority issue. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that at most 130 Americans believe global warming should be a priority issue.
- (b) Find the approximate probability that fewer than 125 Americans believe that global warming should be a priority issue.

- (c) Find the approximate probability that between 135 and 145 (inclusive) believe that global warming should be a priority issue.

6.157 Medicine and Clinical Studies When filling a prescription, a generic option is often available as an alternative to a brand-name drug. The generic drug is usually equally effective and less expensive. In Canada, 70% of all retail prescriptions are filled using generic drugs.⁵ Suppose 300 prescriptions filled in Canada are randomly selected. Answer each problem using the normal approximation to the binomial distribution.

- (a) Let X be the number of prescriptions filled using a generic drug. Find the approximate distribution of the random variable X .
- (b) Find the approximate probability that at least 220 prescriptions were filled using a generic drug.
- (c) Suppose 200 of the prescriptions were filled using generic drugs. Is there any evidence to suggest that the claim is wrong, that the true proportion of prescriptions filled using generic drugs is less than 0.70? Justify your answer.

6.158 Manufacturing and Product Development There is an ongoing debate about whether to label genetically engineered (GE) foods. In the United States, many processed foods include at least one GE ingredient—for example, GE soybeans, corn, or canola. Suppose 60% of all Americans believe GE foods should be labeled and a random sample of 500 Americans is selected. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that exactly 305 Americans believe GE foods should be labeled.
- (b) Find the approximate probability that at least 310 Americans believe GE foods should be labeled.
- (c) Find the approximate probability that more than 310 Americans believe GE foods should be labeled.

6.159 Marketing and Consumer Behavior “Black Friday” is the day after Thanksgiving and the traditional first day of the Christmas shopping season. Suppose a recent poll suggested that 66% of Black Friday shoppers are actually buying for themselves. A random sample of 130 Black Friday shoppers is obtained. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that fewer than 76 Black Friday shoppers are buying for themselves.
- (b) Find the approximate probability that between 77 and 87 (inclusive) Black Friday shoppers are buying for themselves.
- (c) Suppose 90 Black Friday shoppers are buying for themselves. Is there any evidence to suggest that the claim is wrong, that the true proportion of Black Friday shoppers buying for themselves is greater than 0.66? Justify your answer.

6.160 Psychology and Human Behavior We all need to call a customer service representative at some time. It is often difficult to speak with a live person, and the waiting time can be aggravating. Consequently, 36% of Americans admit that they

have yelled at a customer service representative during the past year. Suppose a random sample of 275 Americans who recently talked with a customer service representative is obtained. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that at least 105 Americans yelled at a customer service representative.
- (b) Find the approximate probability that between 90 and 100 (inclusive) Americans yelled at a customer service representative.
- (c) Suppose 85 Americans actually yelled at a customer service representative. Is there any evidence to suggest that the claim is wrong, that the true proportion of Americans who have yelled at a customer service representative is less than 0.36? Justify your answer.

6.161 Stuck on Band-Aids Despite the fact that a majority of accidents happen in one's home, only 44% of Americans have a first-aid kit in their homes. Suppose a random sample of 300 American homes is obtained. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the probability that fewer than 125 homes have first-aid kits.
- (b) Find the probability that more than 145 homes have first-aid kits.
- (c) Find the probability that exactly 130 homes have first-aid kits.

6.162 Counting Sheep There are many reasons why some people have a difficult time falling asleep or staying asleep—for example, too much caffeine, climate, lack of physical activity, or even a lumpy mattress. In a recent poll, among those people who suffered from insomnia, 51% blamed stress. Suppose a random sample of 80 adults who suffer from insomnia was obtained and asked if they blame stress.

- (a) The sample size ($n = 80$) is relatively small in this example. Why is the normal approximation to the binomial distribution appropriate?
- (b) Find the approximate probability that more than 42 adults blame stress for their insomnia.
- (c) Find the approximate probability that at least 38 adults blame stress for their insomnia.
- (d) Find the approximate probability that fewer than 35 or more than 45 adults blame stress for their insomnia.

6.163 Public Health and Nutrition According to a recent report, only 38% of Canadians were vaccinated against the seasonal flu.⁶ Even though the Centers for Disease Control and Prevention recommends the flu vaccination for everyone older than the age of 6 months, many Canadians don't think the shot is needed. Suppose a random sample of 400 Canadians was obtained and each is asked if he or she was vaccinated against the flu. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that fewer than 152 Canadians were vaccinated against the flu.
- (b) Find the approximate probability that exactly 155 Canadians were vaccinated against the flu.

- (c) Suppose 165 Canadians were actually vaccinated against the flu. Is there any evidence to suggest that the claim in the report is wrong, that the true proportion of Canadians who were vaccinated against the flu is greater than 0.38? Justify your answer.

Extended Applications

6.164 Where's the Remote? Most people misplace the remote control to their television at some point. Although some lost remotes are found in the refrigerator, outside, or in a car, approximately 50% are stuck between sofa cushions. Suppose a random sample of 150 lost (and found) remotes is obtained. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the probability that fewer than 70 lost remotes were found stuck between sofa cushions.
- (b) Suppose at least 80 lost remotes were found stuck between sofa cushions. Find the approximate probability that at least 85 lost remotes were found stuck between sofa cushions.
- (c) Suppose 90 lost remotes were found stuck between sofa cushions. Is there any evidence to suggest that the claim is wrong, that the true proportion of lost remotes found between sofa cushions is greater than 0.50? Justify your answer.

6.165 Economics and Finance Online banking is quick and convenient, but it does present some security risks. Among those people who have an online bank account, 47% use their online bank password for at least one other online site. Suppose a random sample of 100 people who have an online bank account is selected and each is asked if they use their bank password for at least one other online site. Answer each problem using the normal approximation to the binomial distribution.

- (a) Find the approximate probability that fewer than 40 people use their online bank password for at least one other online site.
- (b) Find the approximate probability that at least 50 people use their online bank password for at least one other online site.
- (c) Suppose a second random sample of 100 people who have an online bank account is selected and each is also asked if they use their bank password for at least one other online site. Find the approximate probability that between 42 and 52 (inclusive) people use their online bank password for at least one other online site in both random samples.

6.166 Business and Management In October 2018, Pfizer announced plans to reduce its workforce by 2% through voluntary retirements and layoffs.⁷ The workforce reduction was designed to eliminate some managerial positions to simplify Pfizer's corporate structure. Suppose a random sample of 150 Pfizer employees is obtained.

- (a) Use the binomial distribution to find the exact probability that at most 2 employees will lose their jobs.

- (b) Use the normal approximation to the binomial distribution to find the approximate probability that at most 2 employees will lose their jobs.
- (c) Explain why in this case the normal approximation to the binomial distribution is not very accurate.
- (d) Carefully sketch a graph of the binomial probability histogram and the approximate normal probability density function near the mean. Use this graph to justify your answer to part (c).
- (e) Suppose 7 of the employees actually lose their job. Is there any evidence to suggest that Pfizer's claim is wrong, that the true proportion of employees who lose their job is greater than 0.02? Justify your answer using both the binomial distribution and the normal approximation.

6.167 Education and Child Development The Center for American Progress recently released a report detailing the state of child care in America. The study found that 51% of Americans live in neighborhoods classified as childcare deserts, areas with inadequate access to child care.⁸ Suppose 1000 Americans were selected at random and those living in childcare deserts were identified.

- (a) Use the binomial distribution to find the exact probability that 510 Americans live in a childcare desert.
- (b) Use the normal approximation to the binomial distribution to find the approximate probability that exactly 510 Americans live in a childcare desert.
- (c) Explain why in this case the normal approximation to the binomial distribution is very accurate.
- (d) Suppose 525 Americans live in a childcare desert. Is there any evidence to suggest that the proportion of Americans who live in a childcare desert has increased? Justify your answer using the normal approximation to the binomial distribution.

Challenge Problems

6.168 Biology and Environmental Science Recall that a Poisson random variable is often used to model rare events and is a (discrete) count of the number of times a specific event occurs during a given interval. Suppose X is a Poisson random variable with mean λ . Then $\mu = \lambda$ and $\sigma^2 = \lambda$.

If $\lambda > 10$, then the random variable X is approximately normal with mean $\mu = \lambda$ and variance $\sigma^2 = \lambda$: $X \overset{\sim}{\sim} N(\lambda, \lambda)$. As λ increases, the approximation becomes more accurate. An appropriate continuity correction should be used whenever this normal approximation to a Poisson distribution is applied.

Hops are used in brewing beer; they make up about 5% of the total volume of beer, but contribute about 50% of the taste. Suppose that farmers spray their hops 14 times per year using a wide variety of pesticides and fungicides. Also suppose that a hops farmer is selected at random.

- (a) Use the Poisson distribution to find the probability that the farmer sprays the hops at least 17 times during the year.
- (b) Use the Poisson distribution to find the probability that the farmer sprays the hops fewer than 10 times during the year.
- (c) Suppose the farmer sprays the hops between 8 and 20 times (inclusive) during the year. Use the Poisson distribution to find the probability that the farmer sprays the hops at most 12 times during the year.
- (d) Use the normal approximation to the Poisson distribution with the appropriate continuity correction to answer parts (a), (b), and (c).
- (e) Suppose the farmer sprays the hops 22 times during the year. In addition to avoiding beer made with these hops, is there any evidence to suggest that the mean number of times the hops are sprayed in a year is greater than 14? Justify your answer.

Notes

- 1 The Hill talk, accessed on December 20, 2018, <http://thehilltalk.com/2018/09/22/45-percent-cell-phone-spam-calls/>
- 2 Pew Research Center, accessed on December 21, 2018, <http://www.pewresearch.org/fact-tank/2017/12/19/8-of-americans-say-they-own-a-drone-while-more-than-half-have-seen-one-in-operation/>
- 3 The Nation's Report Card, accessed on December 21, 2018, <https://www.nationsreportcard.gov/>
- 4 Sightline Institute, accessed on December 21, 2018, <https://www.sightline.org/2018/08/06/climate-change-hot-voter-issue-2018-midterms/>
- 5 *Global News*, accessed on December 21, 2018, <https://globalnews.ca/news/4251192/generic-vs-brand-name-drugs-canada/>
- 6 Precision Vaccinations, accessed on December 21, 2018, <https://www.precisionvaccinations.com/less-40-percent-canadians-have-been-getting-annual-flu-shot>
- 7 Reuters, accessed on December 21, 2018, <https://www.reuters.com/article/us-pfizer-jobs/pfizer-to-cut-around-2-percent-of-jobs-through-early-next-year-idUSKCN1MR2MA>
- 8 Center for American Progress, accessed on December 21, 2018, <https://www.americanprogress.org/press/release/2018/12/06/461886/release-unprecedented-report-finds-majority-americans-live-child-care-desert-neighborhoods/>