Section 6.3

Last section, we learned about the normal distribution and did computations involving the standard normal distribution. Now, we will continue doing more computations involving the standard normal distribution and we will also work with non-standard normal distributions as well.

Recall from the last section: if a continuous random variable has a distribution with a graph that is symmetric and bell-shaped and it can be described by the equation below:

$$y = \frac{e^{-\frac{1}{2}(\frac{x-u}{\sigma})^2}}{\sigma\sqrt{2\pi}},$$

we say that it has a **normal distribution**.

The symbols μ and σ represent fixed values for the mean and standard deviation, respectively.

Exercise 1. For a standard normal distribution, what are the values of μ and σ?

Exercise 2. Assume that thermometer readings are normally distributed with a mean of 0°C and a standard deviation of 1°C. A thermometer is randomly selected and tested. Find the probability that the reading is greater than 1.43°C.

We will use the function to find the probability.

Let's draw a sketch:

The left bound for the area is:

The right bound for the area is:

Exercise 3. Assume that thermometer readings are normally distributed with a mean of 0°C and a standard deviation of 1°C. A thermometer is randomly selected and tested. Find the probability that the reading is between 1.20°C and 1.78°C.

We will use the normalcdf function to find the probability.

Let's draw a sketch:

The left bound for the area is:

The right bound for the area is:

Class Exercise 1. Assume that thermometer readings are normally distributed with a mean of 0°C and a standard deviation of 1°C. A thermometer is randomly selected and tested. Find the probability that the reading is less than -0.23°C.

Let's draw a sketch:

The left bound for the area is:

The right bound for the area is:

Class Exercise 2. Assume that thermometer readings are normally distributed with a mean of 0°C and a standard deviation of 1°C. A thermometer is randomly selected and tested. Find the probability that the reading is between -0.14°C and 1.49°C.

Let's draw a sketch:

The left bound for the area is:

The right bound for the area is:

Exercise 4. Assume that thermometer readings are normally distributed with a mean of 0°C and a standard deviation of 1°C. A thermometer is randomly selected and tested. Find the temperature reading corresponding to the 95th percentile.

Here is a sketch:

Class Exercise 3. Assume that thermometer readings are normally distributed with a mean of 0°C and a standard deviation of 1°C. A thermometer is randomly selected and tested. Find the temperature reading corresponding to the 6th percentile.

So far, we have only dealt with standard normal distributions. We will now deal with non-standard normal distributions. Earlier in the semester, we learned about z-scores. Again, here is the formula:

If we convert values to standard z-scores using the formula, then procedures for working with all normal distributions are the same as those for the standard normal distributions.

Exercise 5. The scores for a college mathematics exam are normally distributed with a mean of 78 and a standard deviation of 5. The instructor places "Cat in the Hat" reward stickers on the exams for students who score above 80. What percentage of students receive "Cat in the Hat" stickers?

First, let's draw a sketch of the curve, label the mean, and the specific x value, then *shade* the region representing the desired probability:

For each relevant value x that is a boundary for the shaded region, use the formula to convert that value to the equivalent z score.

Use a calculator to find the area of the shaded region:

Exercise 6. The length of human pregnancies from conception to birth varies according to a distribution that is approximately normal with mean 266 days and standard deviation 16 days. What percent of pregnancies last between 240 days and 272 days?

First, let's draw a sketch of the curve, label the mean, and the specific x value, then *shade* the region representing the desired probability:

For each relevant value x that is a boundary for the shaded region, use the formula to convert that value to the equivalent z score.

Use a calculator to find the area of the shaded region:

Class Exercise 4. A demanding college statistics instructor tells his students that he expects them to spend at least three hours on each of their homework assignments. Some of his students meet his time expectations and some don't. The distribution of time spent on homework for his students can be described by a Normal model with a mean of 3.2 hours and a standard deviation of 0.2 hours. What percent of his students actually spend at least three hours on each homework assignment?

First, let's draw a sketch of the curve, label the mean, and the specific x value, then *shade* the region representing the desired probability:

For each relevant value x that is a boundary for the shaded region, use the formula to convert that value to the equivalent z score.

Use a calculator to find the area of the shaded region:

Class Exercise 5. The level of cholesterol in the blood is important because high cholesterol levels increase the risk of heart disease. The distribution of blood cholesterol levels increase the risk of heart disease. The distribution of blood cholesterol levels in a large population of people of the same age and sex is roughly normal. For 14-year-old boys, the mean is $\mu = 170$ milligrams of cholesterol per deciliter of blood (mg/dl) and the standard deviation is $\sigma = 30$ mg/dl. Levels above 240 mg/dl may require medical attention. What percent of 14-year old boys have more than 240 mg/dl of cholesterol?

First, let's draw a sketch of the curve, label the mean, and the specific x value, then *shade* the region representing the desired probability:

For each relevant value x that is a boundary for the shaded region, use the formula to convert that value to the equivalent z score.

Use a calculator to find the area of the shaded region:

There is another way of finding areas under non-standard normal curves. Suppose we are given a normal curve with mean μ and standard deviation σ. The area under the curve between two values a and b is equal to

$$\text{normalcdf}(a,\, b,\, \mu,\, \sigma).$$

Exercise 7. Solve the "Cat in the Hat" stickers problems using the new method.

Again, here is a sketch:

Remark: With standard normal distributions, we could replace ∞ with 10 when using the normalcdf() function. With non-standard normal distributions, you may need to use a larger number, such as 10,000 or 10,000,000.

Class Exercise 6. Solve the last two class exercises using the second method.

Why did I introduce the z-score method? Isn't the second method less time consuming? The z-score method will be used in later chapters so it is important that you clearly understand the z-score method.

Finally, you can also find percentiles for non-standard normal distributions using your calculator. We use the invNorm function.

Suppose $y = \text{invNorm}(b, \mu, \sigma)$. Then, the area under the normal curve with mean μ and standard deviation σ and to the left of y is b.

Exercise 8. Avoiding an accident when driving can depend on reaction time. That time, measured from the moment the driver first sees the danger until he or she gets his foot on the brake pedal, is thought to follow a Normal model with a mean of 1.5 seconds and a standard deviation of 0.18 seconds. Describe the reaction times of the slowest 25% of all drivers.

First, let's draw a sketch:

Class Exercise 7. A tire manufacturer believes that the treadlife of its snow tires can be described by a Normal model with a mean of 32,000 miles and standard deviation of 2,500 miles. In planning a marketing strategy, a local tire dealer wants to offer a refund to any customer whose tires fail to last a certain number of miles. However, the dealer does not want to take too big a risk. If the dealer is willing to give a refund to no more than 4% of customers, for what mileage can he guarantee these tires to last?

First, let's draw a sketch:

Class Exercise 8. Companies who design furniture for elementary school classrooms produce a variety of sizes for kids of different ages. Suppose the heights of kindergarten children can be described by a Normal model with a mean of 38.2 inches and standard deviation of 1.8 inches. At least how tall are the biggest 10% of kindergarteners?

First, let's draw a sketch:

Homework

C problems: 5 - 29 (every 4th)
B problems: 1, 3
A problems: 31

Section 6.6

Now, we look back at binomial random variables. For a binomial random variable $X \sim B(n, p)$, $\mu = np$ and $\sigma^2 = np(1-p)$.

Here is the probability histogram for a $B(20, 0.5)$ random variable:

Notice how the shape of the histogram resembles a normal distribution curve. It turns out that for **some** binomial random variables, the shape of the probability histogram resembles a normal distribution curve. The shape of the probability histogram resembles the normal distribution curve as long as $np > 5$ **and** $n(1-p) > 5$.

We know that a binomial random variable $X \sim B(n,p)$ has mean $\mu = np$ and variance $\sigma^2 = np(1-p)$. Therefore, the normal random variable approximation to the binomial random variable $X \sim B(n,p)$ is $X \sim N(np, \sqrt{np(1-p)})$.

Normal Approximation Rule: Again, we can use the normal approximation if $np > 5$ and $n(1-p) > 5$. (In other words, the approximation doesn't work well for small n or small p/large p. The cutoff point of 5 is somewhat arbitrary.)

Exercise 9. For which of the following binomial random variables can we use the normal approximation? Answer Yes or No.

(a) $X \sim B(10,\ .5)$

(b) $X \sim B(50,\ 0.99)$

(c) $X \sim B(75,\ 0.25)$

(d) $X \sim B(60,\ 0.01)$

(e) $X \sim B(100,\ 0.5)$

Reminder: The normal approximation to the binomial random variable $X \sim B(n,p)$ is $X \sim N(np, \sqrt{np(1-p)})$.

Exercise 10. A quality control engineer examines parts produced by an automated process. It is known that the defective rate for the parts is 8%. What is the probability that there are 14 or more defective parts in a random sample of 150 such parts?

Let $X =$

What is the distribution of X?

$X \sim$

Can we use the normal approximation for X?

From the reminder before this exercise,

From the question, we are looking for

The probability that there are 14 or more defective parts is

We use the normal approximation again in the next exercise.

Exercise 11. Thirty-eight percent of people in the United States admit that they snoop in other people's medicine cabinets. You randomly select 200 people in the United States and ask each if he or she snoops in other people's medicine cabinets. What is the probability that at most 70 will say yes?

Let $X =$

What is the distribution of X?

$X \sim$

Can we use the normal approximation for X?

Using the normal approximation,

From the question, we are looking for

The probability that at most 70 will say yes is

Class Exercise 9. Ocean fishing for billfish is very popular in the Cozumel region of Mexico. In *World Record Game Fishes* (published by the International Game Fish Association), it was stated that in the Cozumel region, about 44% of strikes (while trolling) result in a catch. Suppose that on a given day a fleet of fishing boats got a total of 24 strikes. What is the probability that the number of fish caught was (a) 12 or fewer?, (b) 5 or more?, and (c) between 5 and 12? (#10)

Class Exercise 10. What are the chances that a person who is murdered actually knew the murderer? The answer to the question explains why a lot of police detective work begins with relatives and friends of the victim! About 64% of people who are murdered actually knew the person who committed the murder. Suppose that a detective file in New Orleans has 63 current unsolved murders. What is the probability that
(a) at least 35 of the victims knew their murderers?, (b) at most 48 of the victims knew their murderers?, (c) fewer than 30 victims did *not* know their murderers?, and (d) more than 20 victims did *not* know their murderers? (#12)

Homework

C problems: 3(a), 3(b), 3(c), 3(d), 7-15 ODD
B problems: 1
A problems: 3(e), 5