



## Working with z-Scores 1

Note: The material in this worksheet comes from two chapters of the Triola textbook. Some of the material is introduced in Chapter 3, but the majority of it is from Chapter 6.

The **normal distribution**, or bell-curve distribution, occurs frequently in nature. We can make predictions about data that is normally distributed because all normal curves are basically the same shape.

If we wanted to know things about circles, for example, we might decide to study everything there is to know about the circle with its centre at  $(0, 0)$  and a radius of 1, just to make everything simple. Every other circle that can be drawn has the same basic shape, but may be a different size, or have its centre in a different place. We could use what we know about our standard circle to calculate things about other circles. In fact, we do exactly this in trigonometry, using the unit circle to tell us about various angles.

The equation for a bell curve is more complicated, so it's worth studying one standard bell curve (with its centre at 0, the mean of the distribution, and a standard deviation of 1), and using what we know about that curve to tell us about other curves.

The table for the **standard normal curve** in your textbook reports the area under the curve *to the left* ( $z < z_0$ ) of various values for the horizontal axis. These values are called **z-scores**. To look up a z-score, look down the rows of the first column in the table for the z-score value to one decimal place, and look across the first row of values until you come to the column that represents the second decimal place of the z-score.

*Example 1:* What is the area under the curve to the left of the z-score of  $-1.25$ ?

*Solution:* This is a negative score, so we need the start of the table. We look down the column with the heading  $z$  for  $-1.2$  since that's the z-score we want to one decimal place. (We don't round z-scores when we look them up.) Then we follow along the line starting with  $-1.2$  until we are in the column with the heading  $.05$ , since 5 is in the hundreds place of the z-score we want. The value we get is  $.1056$ , so that's the area.

### CONVERTING TO A Z-SCORE — CHAPTER 3

We never encounter a situation where we have  $\mu = 0$  and  $\sigma = 1$ . For any statistical problem we encounter involving a normal probability distribution, we must convert our real-world values ( $X$ ) into z-scores ( $Z$ ) to be able to use the table. The z-score translates any x-value of any normal distribution into a measure of how far that value is from the mean, in numbers of standard deviations. This is why there are both negative and positive z-scores: negative z-scores represent x-values below the mean, and positive z-scores represent x-values above the mean.



**Example 2:** In a study of bone brittleness, the ages of people at the onset of osteoporosis followed a normal distribution with a mean age of 71 and a standard deviation of 2.8 years. What z-score would an age of 65 represent in this study?

**Solution:** [Sanity check] Since the age we are interested in, 65, is less than the mean, we expect the z-score to be a negative one.

We use the conversion formula: 
$$z = \frac{x - \mu}{\sigma}$$

The x value for this question is 65,  $\mu = 71$ ,  $\sigma = 2.8$ :

$$z = \frac{65 - 71}{2.8} = -2.14\dots$$

## FINDING PROBABILITIES OF EVENTS USING Z-SCORES — CHAPTER 6

The most basic useful question involving the normal curve is finding the probability of an event using the standard normal table.

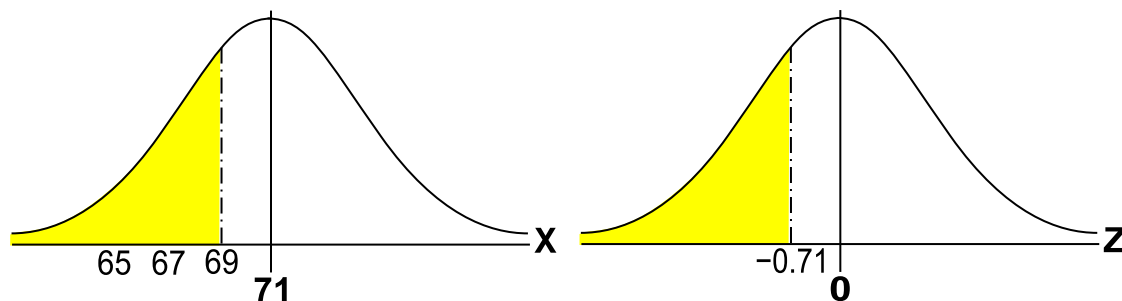
**Example 3:** Using the same bone study as in Example 2, find the probability of a random person getting osteoporosis...

- a) at the age of exactly 69
- b) at an age under 69
- c) at an age of at least 69
- d) at an age between 69 and 72
- e) at an age over 107

**Solution:** First, we convert 69 into a z-score: 
$$z = \frac{x - \mu}{\sigma} = \frac{69 - 71}{2.8} = -0.71\dots$$

a) This question is asking what the probability of getting the disease when you are exactly 69 years old, in other words at the precise instant of your birthday. This is extremely unlikely. The answer is 0. A question of this nature will always ask you about a range of values, and never a single specific value. If it does, it's a trick question.

b) It is important to decide which side of the line  $z = -0.71$  we want to know about.



The question asked about ages less than 69, so we want the left side of the graph. We look up  $-0.71$  in the table and we get .2389. We can write:

$$P(X < 69) = P(Z < -0.71) = 23.89\%$$

Notice that  $P(X < 69)$  means the same as  $P(X \leq 69)$ . Since  $P(X = 69)$  is 0, including it or leaving it out of this calculation won't change the answer.



c) Again, we need to know which side of the line to use. The question asks about ages higher than 69, so we need the right side of the line. However, the table only tells us areas to the left of a value. We need to do another calculation to get the answer.

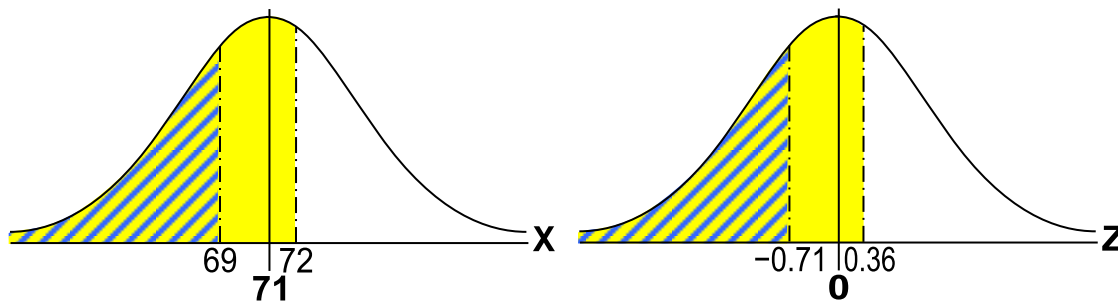
We already know the probability of getting osteoporosis at an age less than 69. Everyone else who contracts osteoporosis must get it at an age of 69 or more. The probability must be  $100\% - 23.89\% = 76.11\%$ . In terms of the normal curve and areas, we can subtract the area that's to the left of the z-score from the entire area under the curve to find the area that's to the right of the z-score. The area under the normal curve is always 1 (whether it's the standard one or not), so the area is  $1 - 0.2389 = 0.7611$ , which gives us the same answer:

$$P(X \geq 69) = P(Z \geq -0.71) = 76.11\%$$

d) For this question, we need an area between two lines. We'll need to convert each x value into a z-score. We know the z-score for 69 years old in this study. For 72 years:

$$z = \frac{x - \mu}{\sigma} = \frac{72 - 71}{2.8} = 0.36\dots$$

Again, we need a trick to calculate the area between the z-scores.



In the diagrams above, we are only interested in the solid (unstriped) area between the lines. We can't get that area directly, but if we take the entire shaded area and subtract the striped area, we'll get what we need. The entire shaded area is the area to the left of 0.36 (which we find in the table as 0.6406), and the striped area is the area to the left of  $-0.71$  (which we know is 0.2389). The area between the lines must be  $0.6406 - 0.2389$  or 0.4017. We write:

$$P(69 < X < 72) = P(-0.71 < Z < 0.36) = 40.17\%$$

e) We want the right side of whatever z-score 107 years old would have.

$$z = \frac{x - \mu}{\sigma} = \frac{107 - 71}{2.8} = 12.86\dots$$

When we try to look this number up, there are no numbers as large as this in the table. In fact, there's nothing higher than 4 or lower than  $-4$ . This means that for z-scores as far above the mean as 12.86, virtually all the area is to the left of the score. For our question, that means there's no realistic chance of getting osteoporosis at such an advanced age. The textbook says to report such a number as 0.0001, but really it's virtually zero (and other sources treat it as approximately zero).



## EXERCISES

A. Using  $\mu = 113.7$  and  $\sigma = 11.6$  convert...

- 1) an x value of 120.5 to a z-score
- 2) an x value of 99.9 to a z-score
- 3) an x value of 108.9 to a z-score
- 4) a z-score of  $-1.64$  to an x value

---

### CHAPTER 6

---

B. What is the area under the standard normal curve...

- 1) to the left of  $z = 1.2$
- 2) to the left of  $z = -0.83$
- 3) to the left of  $z = 0$
- 4) to the right of  $z = -2.9$
- 5) to the right of  $z = 1.8$
- 6) to the right of  $z = 3.32$
- 7) between  $z = -0.37$  and  $0.37$
- 8) between  $z = -2.59$  and  $-0.72$

C. For a study for Stats class, the variable  $X$  represents the number of jelly beans, modelled by a normal distribution. Write a statement in the form  $P(X \dots)$  for the probability that...

- 1) there are more than 2000 jelly beans
- 2) there are no more than 2000 jelly beans
- 3) there are at least 1775 jelly beans
- 4) there are somewhere between 1400 and 1500 jelly beans
- 5) there are at most 2200 jelly beans
- 6) there are at least 1850 jelly beans, but fewer than 2150

D. In a certain school district, the distribution of the heights of eighth-graders who play the tuba is approximately normal, with a mean of 146.1 cm and a standard deviation of 8.4 cm. Determine the probability that a randomly chosen tuba player from the eighth grade...

- 1) is shorter than 150 cm?
- 2) is at least 142 cm tall?
- 3) is 147.33 cm tall?
- 4) is taller than 158.3 cm?
- 5) is between 125 cm and 135 cm tall?
- 6) is between 145 cm and 147 cm tall?
- 7) is less than 110 cm tall?
- 8) is between 160 cm and 190 cm tall?

---

## SOLUTIONS

A. (1) 0.59 (2)  $-1.19$  (3)  $-0.41$  (4) 94.7

B. (1) 0.8849 (2) 0.2033 (3) 0.5000 (4) 0.9981 (5) 0.0359 (6) 0.0005 (7) 0.2886 (8) 0.2310

C. (1)  $P(X > 2000)$  (2)  $P(X \leq 2000)$  (3)  $P(X \geq 1775)$  (4)  $P(1400 < X < 1500)$  (5)  $P(X \leq 2200)$  (6)  $P(1850 \leq X < 2150)$

D. (1) 67.72% (2) 68.79% (3) 0 (4) 7.35% (5)  $0.0934 - 0.0060 = 8.74\%$  (6)  $0.5438 - 0.4483 = 9.55\%$  (7)  $0.0001 = 0.01\%$  (8)  $1.0000 - 0.9505 = 4.95\%$

