

In that case, since the mean equals the total of the observations divided by the sample size, the mean equals

$$\frac{(1)0 + (3)1 + (60)2 + (23)3 + 12(4) + 1(5)}{100} = \frac{245}{100} = 2.45.$$

This calculation has the form

$$0(0.01) + 1(0.03) + 2(0.60) + 3(0.23) + 4(0.12) + 5(0.01),$$

the sum of the possible outcomes times their probabilities. In fact, for any discrete variable y , the mean of its probability distribution has this form.

Mean of a Probability Distribution (Expected Value)

The **mean of the probability distribution** for a discrete variable y is

$$\mu = \sum yP(y).$$

The sum is taken over all possible values of the variable. This parameter is also called the **expected value of y** and is denoted by $E(y)$.

For Table 4.1, for example,

$$\begin{aligned}\mu &= \sum yP(y) = 0P(0) + 1P(1) + 2P(2) + 3P(3) + 4P(4) + 5P(5) \\ &= 0(0.01) + 1(0.03) + 2(0.60) + 3(0.23) + 4(0.12) + 5(0.01) \\ &= 2.45.\end{aligned}$$

This is also the *expected value* of y , $E(y) = \mu = 2.45$. The terminology reflects that $E(y)$ represents what we expect for the average value of y in a long series of observations.

The **standard deviation** of a probability distribution, denoted by σ , measures its variability. The larger the value of σ , the more spread out the distribution. In a rough sense, σ describes how far the variable y tends to fall from the mean of its distribution. The Empirical Rule (Section 3.3) helps us to interpret σ . If a probability distribution is approximately bell shaped, about 68% of the probability falls between $\mu - \sigma$ and $\mu + \sigma$, about 95% falls between $\mu - 2\sigma$ and $\mu + 2\sigma$, and all or nearly all falls between $\mu - 3\sigma$ and $\mu + 3\sigma$. For example, suppose commuting time to work for a randomly selected worker in Toronto has a bell-shaped probability distribution with $\mu = 24$ minutes and $\sigma = 8$ minutes. Then there's about a 95% chance that commuting time falls between $24 - 2(8) = 8$ minutes and $24 + 2(8) = 40$ minutes.

The standard deviation is the square root of the **variance** of the probability distribution. The variance measures the average squared deviation of an observation from the mean. That is, it is the expected value of $(y - \mu)^2$. We shall not need to compute this measure, so we do not study its formula here (Exercise 4.55 shows the formula for σ for the discrete case).

4.3 THE NORMAL PROBABILITY DISTRIBUTION

Some probability distributions are important because they approximate well the distributions of variables in the real world. Some are important because of their uses in statistical inference. This section introduces the **normal probability distribution**, which is important for both reasons. Its bell-shaped curve describes well many histograms of data for variables that are continuous or take a large number of

possible values. It is the most important distribution for statistical inference, because we'll see it is still useful even when the sample data are *not* bell shaped.

Normal Distribution

The **normal distribution** is symmetric, bell shaped, and characterized by its mean μ and standard deviation σ . The probability within any particular number of standard deviations of μ is the same for all normal distributions. This probability equals 0.68 within 1 standard deviation, 0.95 within 2 standard deviations, and 0.997 within 3 standard deviations.

Each normal distribution is specified by two parameters—its mean μ and standard deviation σ . For any real number for μ and any nonnegative number for σ , there is a normal distribution having that mean and standard deviation. Figure 4.3 illustrates. Essentially the entire distribution falls between $\mu - 3\sigma$ and $\mu + 3\sigma$.

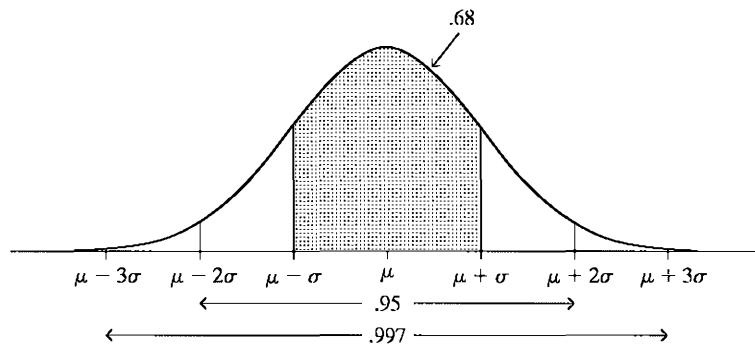


FIGURE 4.3: For Every Normal Distribution, the Probability Equals (Rounded) 0.68 within σ of μ , 0.95 within 2σ of μ , and 0.997 within 3σ of μ

For example, heights of adult females in North America have approximately a normal distribution with mean $\mu = 65.0$ inches and standard deviation $\sigma = 3.5$. The probability is nearly 1.0 that a randomly selected female has height between $\mu - 3\sigma = 65.0 - 3(3.5) = 54.5$ inches and $\mu + 3\sigma = 65.0 + 3(3.5) = 75.5$ inches. Adult male height has a normal distribution with $\mu = 70.0$ and $\sigma = 4.0$ inches. So the probability is nearly 1.0 that a randomly selected male has height between $\mu - 3\sigma = 70.0 - 3(4.0) = 58$ inches and $\mu + 3\sigma = 70.0 + 3(4.0) = 82$ inches. See Figure 4.4.

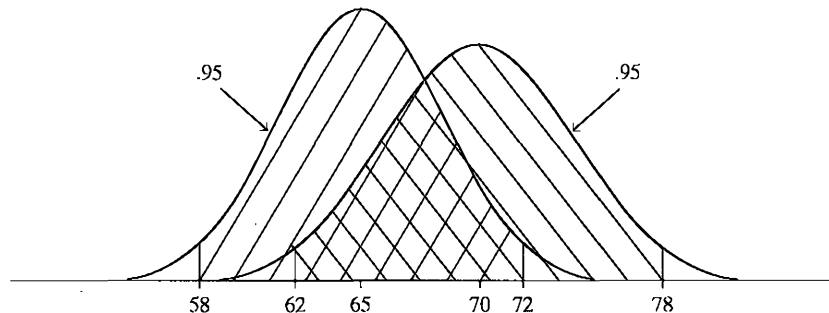


FIGURE 4.4: Normal Distributions for Women's Height ($\mu = 65, \sigma = 3.5$) and for Men's Height ($\mu = 70, \sigma = 4.0$)

Tabulated Normal Tail Probabilities

For the normal distribution, for each fixed number z , the probability falling within z standard deviations of the mean depends only on the value of z . This is the area under the bell-shaped normal curve between $\mu - z\sigma$ and $\mu + z\sigma$. For every normal distribution, this probability is 0.68 for $z = 1$, 0.95 for $z = 2$, and nearly 1.0 for $z = 3$.

For a normal distribution, the probability concentrated within $z\sigma$ of μ is the same for all normal curves even if z is not a whole number—for instance $z = 1.43$ instead of 1, 2, or 3. Table A in Appendix A, also shown next to the inside back cover, determines probabilities for any region of values. It tabulates the probability for the values falling in the right tail, at least z standard deviations above the mean. The left margin column of the table lists the values for z to one decimal point, with the second decimal place listed above the columns.

TABLE 4.2: Part of Table A Displaying Normal Right-Tail Probabilities

z	Second Decimal Place of z									
	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
										
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
										
										

Table 4.2 displays a small excerpt from Table A. The probability for $z = 1.43$ falls in the row labeled 1.4 and in the column labeled .03. It equals 0.0764. This means that for every normal distribution, the right-tail probability above $\mu + 1.43\sigma$ (that is, more than 1.43 standard deviations above the mean) equals 0.0764.

Since the entries in Table A are probabilities for the right half of the normal distribution above $\mu + z\sigma$, they fall between 0 and 0.50. By the symmetry of the normal curve, these right-tail probabilities also apply to the left tail below $\mu - z\sigma$. For example, the probability below $\mu - 1.43\sigma$ also equals 0.0764. The left-tail probabilities, called *cumulative probabilities*, are given by many calculators and software.

Normal Probabilities and the Empirical Rule

The probabilities in Table A apply to the normal distribution and also apply approximately to other bell-shaped distributions. This table yields the probabilities for the Empirical Rule. That rule states that for bell-shaped histograms, about 68% of the data fall within 1 standard deviation of the mean, 95% within 2 standard deviations, and all or nearly all within 3 standard deviations.

For example, the value two standard deviations above the mean has a z -value of 2.00. The normal curve probability listed in Table A opposite $z = 2.00$ is 0.0228. The right-tail probability above $\mu + 2\sigma$ equals 0.0228 for every normal distribution. The left-tail probability below $\mu - 2\sigma$ also equals 0.0228, by symmetry (see Figure 4.5). The total probability more than two standard deviations from the mean is $2(0.0228) = 0.0456$. Since the probability more than two standard deviations from the mean equals 0.0456, the probability between $\mu - 2\sigma$ and $\mu + 2\sigma$ (i.e., within two standard deviations of the mean) equals $1 - 0.0456 = 0.9544$. (Here, we've used rule (1) of the probability rules at the end of Section 4.1, that $P(\text{not } A) = 1 - P(A)$.) When a

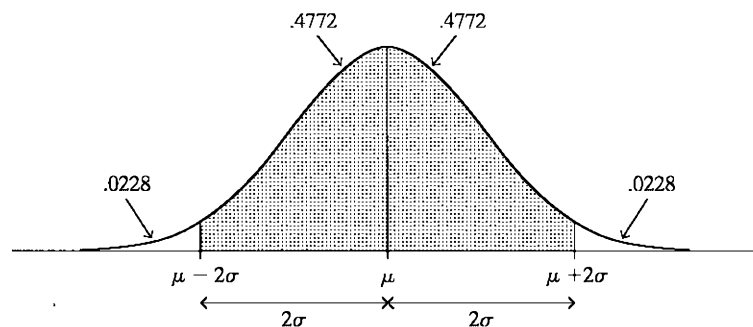


FIGURE 4.5: The Probability within Two Standard Deviations of the Mean for a Normal Distribution Is $1 - 2(0.0228)$, Which Is about 0.95

variable has a normal distribution, 95% of the observations fall within two standard deviations of the mean.

The probability equals 0.50 above the mean, since the normal distribution is symmetric about μ . So the probability between μ and $\mu + 2\sigma$ or between $\mu - 2\sigma$ and μ equals $0.50 - 0.0228 = 0.4772$, also shown in Figure 4.5. Again, we see that the total probability within two standard deviations of the mean equals $2(0.4772) = 0.9544$, or about 95%.

The approximate percentages in the Empirical Rule are the actual percentages for the normal distribution, rounded to two decimal places. For instance, with Table A you can verify that the probability within one standard deviation of the mean of a normal distribution equals 0.68. (*Hint:* Let $z = 1.00$.) The Empirical Rule stated the percentages as being *approximate* rather than *exact*. Why? Because that rule referred to *all approximately bell-shaped distributions*, not just the normal. Not all bell-shaped distributions are normal, only those described by the mathematical formula shown in Exercise 4.56 at the end of the chapter. We won't need that formula, but we will use the probabilities tabulated for it in Table A throughout the text.

Finding z-Values for Certain Tail Probabilities

Many inferential methods use z -values corresponding to certain normal curve probabilities. This entails the reverse use of Table A. Starting with a tail probability, which is listed in the body of Table A, we find the z -value that provides the number of standard deviations that that number falls from the mean.

To illustrate, let's find the z -value having a right-tail probability of 0.025. We look up 0.025 in the body of Table A. It corresponds to $z = 1.96$. This means that a probability of 0.025 falls above $\mu + 1.96\sigma$. Similarly, a probability of 0.025 falls below $\mu - 1.96\sigma$. So a total probability of $0.025 + 0.025 = 0.050$ falls more than 1.96σ from μ . We saw in the previous subsection that 95% of a normal distribution falls within two standard deviations of the mean. More precisely, 0.9544 falls within 2.00 standard deviations, and here we've seen that 0.950 falls within 1.96 standard deviations.

To check that you understand this reasoning, verify that the z -value for a right-tail probability of (1) 0.05 is $z = 1.64$, (2) 0.01 is $z = 2.33$, (3) 0.005 is $z = 2.58$. Show that 90% of a normal distribution falls between $\mu - 1.64\sigma$ and $\mu + 1.64\sigma$.

EXAMPLE 4.3 Finding the 99th Percentile of IQ Scores

Stanford-Binet IQ scores have approximately a normal distribution with mean = 100 and standard deviation = 16. What is the 99th percentile of IQ scores? In other words, what is the IQ score that falls above 99% of the scores?

To answer this, we need to find the value of z such that $\mu + z\sigma$ falls above 99% of a normal distribution. We use the normal curve probability in the right tail beyond the 99th percentile. Then we can use Table A to find the z -value corresponding to that probability. Now, for $\mu + z\sigma$ to represent the 99th percentile, the probability below $\mu + z\sigma$ must equal 0.99, by the definition of a percentile. So 1% of the distribution is above the 99th percentile. The right-tail probability equals 0.01, as Figure 4.6 shows.

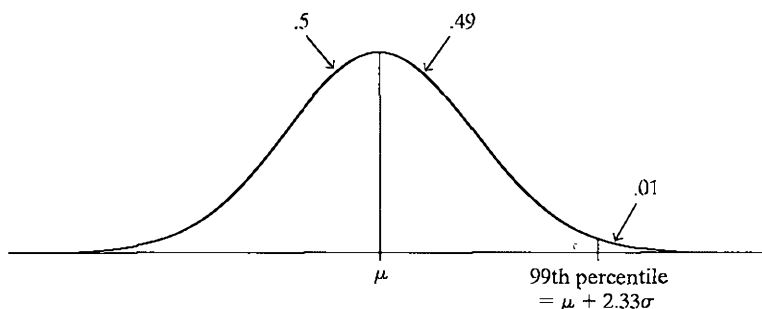


FIGURE 4.6: The 99th Percentile for a Normal Distribution Has 99% of the Distribution below that Point and 1% above It

The body of Table A does not contain the tail probability of exactly 0.0100. The tail probability = 0.0102 corresponds to $z = 2.32$, and tail probability = 0.0099 corresponds to $z = 2.33$. We could interpolate, but it is sufficient to use the z -value rounded to two decimal places. We select the one having probability closer to the desired value of 0.0100. Thus, the 99th percentile is 2.33 standard deviations above the mean. In summary, 99% of any normal distribution is located below $\mu + 2.33\sigma$, no more than 2.33 standard deviations above the mean.

For the IQ scores with mean = 100 and standard deviation = 16, the 99th percentile equals

$$\mu + 2.33\sigma = 100 + 2.33(16) = 137.$$

That is, about 99% of IQ scores fall below 137. ■

To check that you understand the reasoning above, show that the 95th percentile of a normal distribution is $\mu + 1.64\sigma$, and show that the 95th percentile for the IQ distribution equals 126.

z-Score Is Number of Standard Deviations from Mean

The z symbol in a normal table refers to the distance between a possible value y of a variable and the mean μ of its probability distribution, in terms of the *number of standard deviations* that y falls from μ .

EXAMPLE 4.4 z-Scores for College Board Test Scores

Scores on each portion of the Scholastic Aptitude Test (SAT), a college entrance examination, have traditionally been approximately normal with mean $\mu = 500$ and standard deviation $\sigma = 100$. The test score of $y = 650$ has a z -score of $z = 1.50$, because 650 is 1.50 standard deviations above the mean. In other words, $y = 650 = \mu + z\sigma = 500 + z(100)$, where $z = 1.50$. ■

For sample data, Section 3.4 introduced the z -score as a measure of position. Let's review how to find it. The distance between y and the mean μ equals $y - \mu$. The z -score expresses this difference in units of standard deviations.

z-Score

The **z-score** for a value y of a variable is the *number of standard deviations* that y falls from μ . It equals

$$z = \frac{\text{Observation} - \text{Mean}}{\text{Standard Deviation}} = \frac{y - \mu}{\sigma}.$$

To illustrate, when $\mu = 500$ and $\sigma = 100$, an observation of $y = 650$ has the z-score of

$$z = \frac{y - \mu}{\sigma} = \frac{650 - 500}{100} = 1.50.$$

Positive z-scores occur when the number y falls *above* the mean μ . *Negative* z-scores occur when the number y falls *below* the mean. For example, for SAT scores with $\mu = 500$ and $\sigma = 100$, a value of $y = 350$ has a z-score of

$$z = \frac{y - \mu}{\sigma} = \frac{350 - 500}{100} = -1.50.$$

The test score of 350 is 1.50 standard deviations below the mean. The value $y = 350$ falls below the mean, so the z-score is negative.

Table A contains only positive z-values. Since the normal distribution is symmetric about the mean, the left-tail probability below $-z$ equals the right-tail probability above $+z$. Looking up $z = 1.50$ in Table A, we see that the probability that a SAT score falls below 350 is 0.0668, as Figure 4.7 shows. Fewer than 7% of the scores are below 350, and fewer than 7% fall above 650.

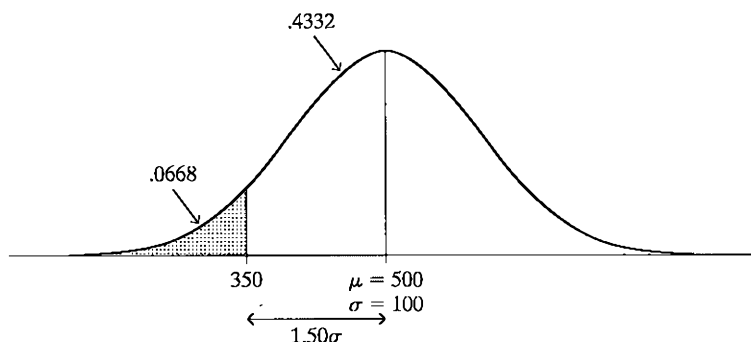


FIGURE 4.7: Normal Distribution for SAT Scores

The next example shows that z-scores provide a useful way to compare positions for different normal distributions.

EXAMPLE 4.5 Comparing SAT and ACT Test Scores

Suppose that when you applied to college, you took a SAT exam, scoring 550. Your friend took the ACT exam, scoring 30. Which score is better?

We cannot compare the test scores of 550 and 30 directly, because they have different scales. We convert them to z-scores, analyzing how many standard deviations each falls from the mean. If SAT has $\mu = 500$ and $\sigma = 100$, a SAT score of $y = 550$ converts to a z-score of

$$z = \frac{(y - \mu)}{\sigma} = \frac{(550 - 500)}{100} = 0.50.$$

The ACT has approximately $\mu = 18$ and $\sigma = 6$, so $\text{ACT} = 30$ converts to a z -score of $(30 - 18)/6 = 2.0$.

The ACT score of 30 is relatively higher than the SAT score of 650, because 30 is 2.0 standard deviations above its mean whereas 650 is only 0.5 standard deviations above its mean. The SAT and ACT scores both have approximate normal distributions. From Table A, $z = 2.0$ has a right-tail probability of 0.0228 and $z = 0.5$ has a right-tail probability of 0.3085. Of all students taking the ACT, only about 2% scored higher than 30, whereas of all students taking the SAT, about 31% scored higher than 650. In this relative sense, the ACT score is higher. ■

Here's a summary of how we've used z -scores:

Using z -Scores to Find Probabilities or y -Values

- If we have a value y and need to find a probability, convert y to a z -score using $z = (y - \mu)/\sigma$, and use a table of normal probabilities to convert it to the probability of interest.
- If we have a probability and need to find a value of y , convert the probability to a tail probability and find the z -score using a normal table, and then evaluate $y = \mu + z\sigma$.

For example, Example 4.6 used the equation $z = (y - \mu)/\sigma$ to determine how many standard deviations an SAT test score fell from the mean. Example 4.3 used the equation $y = \mu + z\sigma$ to find a percentile score for a normal distribution of IQ scores.

The Standard Normal Distribution

Many inferential statistical methods use a particular normal distribution, called the **standard normal distribution**.

Standard Normal Distribution

The **standard normal distribution** is the normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$.

For the standard normal distribution, the number falling z standard deviations above the mean is $\mu + z\sigma = 0 + z(1) = z$. It is simply the z -score itself. For instance, the value of 2 is two standard deviations above the mean, and the value of -1.3 is 1.3 standard deviations below the mean. The original values are the same as the z -scores. See Figure 4.8.

When the values for an arbitrary normal distribution are converted to z -scores, those z -scores are centered around 0 and have a standard deviation of 1. The z -scores have the standard normal distribution.

z -Scores and the Standard Normal Distribution

If a variable has a normal distribution, and if its values are converted to z -scores by subtracting the mean and dividing by the standard deviation, then the z -scores have the standard normal distribution.

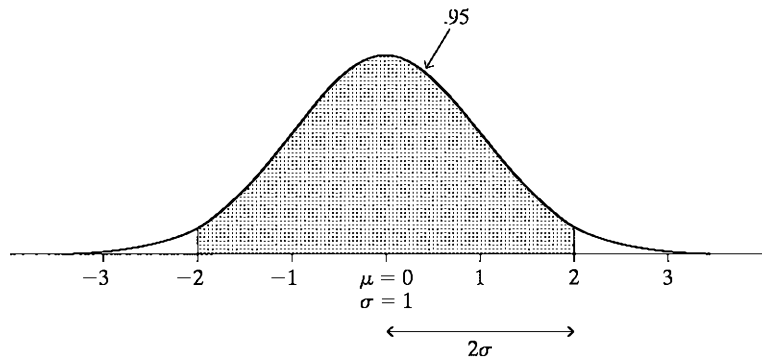


FIGURE 4.8: The Standard Normal Distribution Has Mean 0 and Standard Deviation 1. Its ordinary scores are the same as its *z-scores*.

Suppose we convert each SAT score y to a z -score by using $z = (y - 500)/100$. For instance, $y = 650$ converts to $z = 1.50$, and $y = 350$ converts to $z = -1.50$. Then the entire set of z -scores has a normal distribution with a mean of 0 and a standard deviation of 1. This is the standard normal distribution.

Many inferential methods convert values of statistics to z -scores and then to normal curve probabilities. We use z -scores and normal probabilities often throughout the rest of the book.

4.4 SAMPLING DISTRIBUTIONS DESCRIBE HOW STATISTICS VARY

We've seen that probability distributions summarize probabilities of possible outcomes for a variable. So far, this chapter has treated these distributions as known. In practice, they are rarely known. We use sample data to make inferences about the parameters of those distributions. However, probability distributions with fixed parameter values are useful for many of those inferential methods. Let's look at an example that illustrates the connection between statistical inference and probability calculations with known parameter values.

EXAMPLE 4.6 Predicting an Election from an Exit Poll

Television networks sample voters on election day to help them predict the winners early. For the fall 2006 election for Governor of California, CNN² reported results of an exit poll of 2705 voters. CNN stated that 56.5% reported voting for the Republican candidate, Arnold Schwarzenegger. In this example, the probability distribution for a person's vote would state the probability that a randomly selected voter voted for Schwarzenegger. This equals the proportion of the population of voters who voted for him. When the exit poll was taken, this was an unknown population parameter.

To judge whether this is sufficient information to predict the outcome of the election, the network can ask, "Suppose only half the population voted for Schwarzenegger. Would it then be surprising that 56.5% of the sampled individuals voted for him?" If this would be very unlikely, the network infers that Schwarzenegger received more than half the population votes. The inference about the election outcome is based on finding the probability of the sample result under the supposition that the population parameter, the percentage of voters preferring Schwarzenegger, equals 50%. ■

²www.cnn.com/ELECTION/2006