Some Key Ingredients for Inferential Statistics

The Normal Curve, Sample versus Population, and Probability

Chapter Outline

Ordinarily, behavioral and social scientists do a research study to test some theoretical principle or the effectiveness of some practical procedure. For example, an educational researcher might compare reading speeds of students taught with two different methods to examine a theory of teaching. A sociologist might examine the effectiveness of a program of neighborhood meetings intended to promote water conservation. Such studies are carried out with a particular group of research participants. But the researchers usually want to make more general conclusions about the theoretical principle or procedure being studied. These conclusions go *beyond* the particular group of research participants studied. To do this, researchers use inferential statistics.

TIP FOR SUCCESS

Before beginning this chapter, be sure you have mastered the shapes of distributions and the concepts of mean, standard deviation, and *Z* scores.

In this chapter, we consider three topics: the normal curve, sample versus population, and probability.

The Normal Curve

The graphs of many of the distributions of variables that behavioral and social scientists study (as well as many other distributions in nature) follow a unimodal, roughly symmetrical, bell-shaped distribution. These bell-shaped histograms or frequency polygons approximate a precise and important distribution called the **normal distribution** or, more simply, the **normal curve.** The normal curve is a mathematical (or theoretical) distribution. Researchers often compare the actual distributions of the variables they are studying (that is, the distributions they find in research studies) to the normal curve. They don't expect the distributions of their variables to match the normal curve *perfectly* (since the normal curve is a theoretical distribution), but researchers often check whether their variables *approximately* follow a normal curve. An example of the normal curve is shown in Figure 1.

Why the Normal Curve Is So Common in Nature

Take, for example, the number of different letters a particular person can remember accurately on various testings (with different random letters each time). On some testings, the number of letters remembered may be high, on others low, and on most somewhere in between. That is, the number of different letters a person can recall on various testings probably approximately follows a normal curve. Suppose that theperson has a basic ability to recall, say, seven letters in this kind of memory task. Nevertheless, on any particular testing, the actual number recalled will be affected by various influences—noisiness of the room, the person's mood at the moment, a combination of random letters unwittingly confused with a familiar name, and so on.

These various influences add up to make the person recall more than seven on some testings and less than seven on others. However, the particular combination of such influences that come up at any testing is essentially random. Thus, on most testings, positive and negative influences should cancel out. The chances are not very good of all the random negative influences happening to come together on a testing when none of the random positive influences show up. Thus, in general, the person remembers a middle amount, an amount in which all the opposing influences cancel each other out. Very high or very low scores are much less common.

This creates a unimodal distribution with most of the scores near the middle and fewer at the extremes. It also creates a distribution that is symmetrical, because the number of letters recalled is as likely to be above as below the middle. Being a unimodal symmetrical curve does not guarantee that it will be a normal curve; it could

normal distribution Frequency distribution following a normal curve.

normal curve Specific, mathematically defined, bell-shaped frequency distribution that is symmetrical and unimodal; distributions observed in nature and in research commonly approximate it.

Figure 1 A normal curve.

be too flat or too pointed. However, it can be shown mathematically that in the long run, if the influences are truly random and the number of different influences being combined is large, a precise normal curve results. Mathematical statisticians call this principle the *central limit theorem.*

Although *many* of the distributions of variables that behavioral and social scientists study approximately follow a normal curve, this is not *always* the case. However, for now we focus on the more common situations in which the variables we are studying do approximately follow a normal curve.

The Normal Curve and the Percentage of Scores between the Mean and 1 and 2 Standard Deviations from the Mean

The shape of the normal curve is standard. Thus, there is a known percentage of scores above or below any particular point. For example, exactly 50% of the scores in a normal curve are below the mean, because in any symmetrical distribution, half the scores are below the mean. More interestingly, as shown in Figure 2, approximately 34% of the scores are always between the mean and 1 standard deviation from the mean.

Consider IQ scores. On many widely used intelligence tests, the mean IQ is 100, the standard deviation is 16, and the distribution of IQs is roughly a normal curve (see Figure 3). Knowing about the normal curve and the percentage of scores between the mean and 1 standard deviation above the mean tells you that about 34% of people have IQs between 100, the mean IQ, and 116, the IQ score that is 1 standard deviation above the mean. Similarly, because the normal curve is symmetrical, about 34% of people have IQs between 100 and 84 (the score that is 1 standard deviation below the mean), and 68% (34% + 34%) have IQs between 84 and 116.

There are many fewer scores between 1 and 2 standard deviations from the mean than there are between the mean and 1 standard deviation from the mean. It turns out that about 14% of the scores are between 1 and 2 standard deviations above the mean (see Figure 2). (Similarly, about 14% of the scores are between 1 and 2 standard deviations below the mean.) Thus, about 14% of people have IQs between 116 (1 standard deviation above the mean) and 132 (2 standard deviations above the mean).

You will find it very useful to remember the 34% and 14% figures. These figures tell you the percentage of people above and below any particular score whenever that score is an exact number of standard deviations above or below the mean. You can also reverse this approach and figure out a person's number of standard deviations from the mean from a percentage. Suppose a laboratory test showed that a particular archaeological sample of ceramics was in the top 2% of all the samples in the excavation for containing iron, and it is known that the distribution of iron in such samples

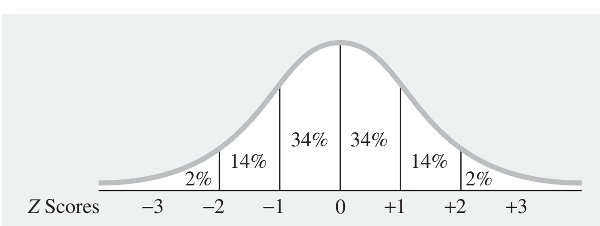


Figure 2 Normal curve with approximate percentages of scores between the mean and 1 and 2 standard deviations above and below the mean.

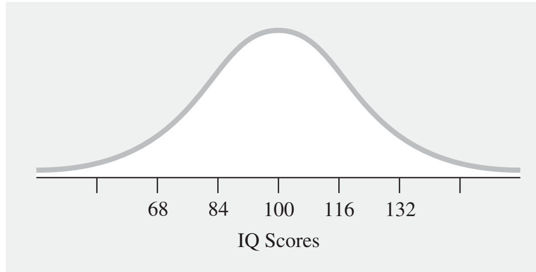


Figure 3 Distribution of IQ scores on many standard intelligence tests (with a mean of 100 and a standard deviation of 16).

roughly follows a normal curve. In this situation, the sample must have a level of iron that is at least 2 standard deviations above the mean level of iron. This is because a total of 50% of the scores are above the mean. There are 34% between the mean and 1 standard deviation above the mean and another 14% between 1 and 2 standard deviations above the mean. That leaves 2% of scores (that is, 50% − 34% − 14% = 2%) that are 2 standard deviations or more above the mean.

Remember that a Z score is the number of standard deviations a score is above or below the mean—which is just what we are talking about here. Thus, if you knew the mean and the standard deviation of the iron concentrations in the samples, you could figure out the raw score (the actual level of iron in this sample) that is equivalent to being 2 standard deviations above the mean. You would do this using methods of changing Z scores (in this case, a Z score of $+2$) to raw scores (and vice versa), namely,

$$X = (Z)(SD) + M \text{ and } Z = (X - M)/SD.$$

The Normal Curve Table and *Z* Scores

The 50%, 34%, and 14% figures are important practical rules for working with a group of scores that follow a normal distribution. However, in many research and applied situations, behavioral and social scientists need more accurate information. Because the normal curve is a precise mathematical curve, you can figure the *exact percentage* of scores between any two points on the normal curve (not just those that happen to be right at 1 or 2 standard deviations from the mean). For example, exactly 68.59% of scores have a Z score between $+.62$ and -1.68; exactly 2.81% of scores have a Z score between $+.79$ and $+.89$; and so forth.

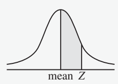
You can figure these percentages using calculus, based on the formula for the normal curve (which you can look up in a mathematical statistics text). However, you can also do this much more simply (which you are probably glad to know!). Statisticians have worked out tables for the normal curve that give the percentage of scores between the mean (a Z score of 0) and any other Z score (as well as the percentage of scores in the tail for any Z score).

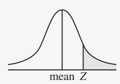
normal curve table Table showing percentages of scores associated with the normal curve; the table usually includes percentages of scores between the mean and various numbers of standard deviations above the mean and percentages of scores more positive than various numbers of standard deviations above the mean.

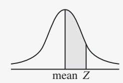
Table 1 shows the first part of the **normal curve table.**[1] The first column in the table lists the Z score. The second column, labeled "% Mean to Z," gives the percentage of

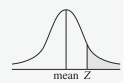
[1]The exact percentage of scores between any two Z scores can also be calculated using statistics or spreadsheet software (for example, using the normal curve function in Microsoft Excel).

Table 1 Normal Curve Areas: Percentage of the Normal Curve Between the Mean and the *Z* Scores Shown and Percentage of Scores in the Tail for the *Z* Scores Shown. (First part of table only. Highlighted values are examples from the text.)









Z	% Mean to *Z*	% in Tail	*Z*	% Mean to *Z*	% in Tail
.00	.00	50.00	.45	17.36	32.64
.01	.40	49.60	.46	17.72	32.28
.02	.80	49.20	.47	18.08	31.92
.03	1.20	48.80	.48	18.44	31.56
.04	1.60	48.40	.49	18.79	31.21
.05	1.99	48.01	.50	19.15	30.85
.06	2.39	47.61	.51	19.50	30.50
.07	2.79	47.21	.52	19.85	30.15
.08	3.19	46.81	.53	20.19	29.81
.09	3.59	46.41	.54	20.54	29.46
.10	3.98	46.02	.55	20.88	29.12
.11	4.38	45.62	.56	21.23	28.77
.12	4.78	45.22	.57	21.57	28.43
.13	5.17	44.83	.58	21.90	28.10
.14	5.57	44.43	.59	22.24	27.76
.15	5.96	44.04	.60	22.57	27.43
.16	6.36	43.64	.61	22.91	27.09
.17	6.75	43.25	.62	23.24	26.76
.18	7.14	42.86	.63	23.57	26.43
.19	7.53	42.47	.64	23.89	26.11
.20	7.93	42.07	.65	24.22	25.78
.21	8.32	41.68	.66	24.54	25.46

TIP FOR SUCCESS

Notice that the table repeats the basic three columns twice on the page. Be sure to look across to the appropriate column you need.

scores between the mean and that *Z* score. The shaded area in the curve at the top of the column gives a visual reminder of the meaning of the percentages in the column. The third column, labeled "% in Tail," gives the percentage of scores in the tail for that *Z* score. The shaded tail area in the curve at the top of the column shows the meaning of the percentages in the column. Notice that the table lists only positive *Z* scores. This is because the normal curve is perfectly symmetrical. Thus, the percentage of scores between the mean and, say, a *Z* of +.98 (which is 33.65%) is exactly the same as the percentage of scores between the mean and a *Z* of −.98 (again, 33.65%); and the percentage of scores in the tail for a *Z* score of +1.77 (3.84%) is the same as the percentage of scores in the tail for a *Z* score of −1.77 (again, 3.84%). Notice that for each *Z* score, the "% Mean to *Z*" value and the "% in Tail" value sum to 50.00. This is because exactly 50% of the scores are above the mean for a normal curve. For example, for the *Z* score of .57, the % Mean to *Z* value is 21.57% and the % in Tail value is 28.43%, and 21.57% + 28.43% = 50.00%.

TIP FOR SUCCESS

Remember that negative *Z* scores are scores below the mean and positive *Z* scores are scores above the mean.

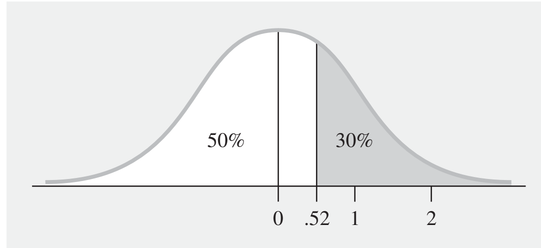


Figure 4 Distribution of creativity test scores showing area for top 30% and *Z* score where this area begins.

Suppose you want to know the percentage of scores between the mean and a *Z* score of .64. You just look up .64 in the "*Z*" column of the table and the "% Mean to *Z*" column tells you that 23.89% of the scores in a normal curve are between the mean and this *Z* score. These values are highlighted in Table 1.

You can also reverse the process and use the table to find the *Z* score for a particular percentage of scores. For example, imagine that 30% of ninth-grade students had a creativity score higher than Samantha's. Assuming that creativity scores follow a normal curve, you can figure out her *Z* score as follows: if 30% of students scored higher than she did, then 30% of the scores are in the tail above her score. This is shown in Figure 4. So, you would look at the "% in Tail" column of the table until you found the percentage that was closest to 30%. In this example, the closest is 30.15%. Finally, look at the "*Z*" column to the left of this percentage, which lists a *Z* score of .52 (these values of 30.15% and .52 are highlighted in Table 1). Thus, Samantha's *Z* score for her level of creativity is .52. If you know the mean and standard deviation for ninth-grade students' creativity scores, you can figure out Samantha's actual raw score on the test by changing her *Z* score of .52 to a raw score using the usual formula, $X = (Z)(SD) + (M)$.

Steps for Figuring the Percentage of Scores above or below a Particular Raw Score or *Z* Score Using the Normal Curve Table

Here are the five steps for figuring the percentage of scores.

❶ **If you are beginning with a raw score, first change it to a *Z* score.** Use the usual formula, $Z = (X - M)/SD$.

❷ **Draw a picture of the normal curve, decide where the *Z* score falls on it, and shade in the area for which you are finding the percentage.**

❸ **Make a rough estimate of the shaded area's percentage based on the 50%–34%–14% percentages.** You don't need to be very exact; it is enough just to estimate a range in which the shaded area has to fall, figuring it is between two particular whole *Z* scores. This rough estimate step is designed not only to help you avoid errors (by providing a check for your figuring), but also to help you develop an intuitive sense of how the normal curve works.

❹ **Find the exact percentage using the normal curve table, adding 50% if necessary.** Look up the *Z* score in the "*Z*" column of Table 1 of the appendix "Tables" and find the percentage in the "% Mean to *Z*" column or "% in Tail" next to it. If you want the percentage of scores between the mean and this *Z* score, the percentage in the table is your final answer. However, sometimes you need to add 50% to the percentage in the

TIP FOR SUCCESS

When marking where the *Z* score falls on the normal curve, be sure to put it in the right place above or below the mean according to whether it is a positive or negative *Z* score.

table. You need to do this if the Z score is positive and you want the total percentage below this Z score, or if the Z score is negative and you want the total percentage above this Z score. However, you don't need to memorize these rules. It is much easier to make a picture for the problem and reason out whether the percentage you have from the table is correct as is, or if you need to add 50%.

❺ **Check that your exact percentage is within the range of your rough estimate from Step ❸.**

Examples

Here are two examples using IQ scores where $M = 100$ and $SD = 16$.

Example 1: If a person has an IQ of 125, what percentage of people have higher IQs?

❶ **If you are beginning with a raw score, first change it to a *Z* score.** Using the usual formula, $Z = (X - M)/SD$, $Z = (125 - 100)/16 = +1.56$.

❷ **Draw a picture of the normal curve, decide where the *Z* score falls on it, and shade in the area for which you are finding the percentage.** This is shown in Figure 5 (along with the exact percentages figured later).

❸ **Make a rough estimate of the shaded area's percentage based on the 50%–34%–14% percentages.** If the shaded area started at a Z score of 1, it would have 16% above it. If it started at a Z score of 2, it would have only 2% above it. So, with a Z score of 1.56, it has to be somewhere between 16% and 2%.

❹ **Find the exact percentage using the normal curve table, adding 50% if necessary.** In Table 1 of the appendix "Tables," 1.56 in the "Z" column goes with 5.94 in the "% in Tail" column. Thus, 5.94% of people have IQ scores higher than 125. This is the answer to our problem. (There is no need to add 50% to the percentage.)

❺ **Check that your exact percentage is within the range of your rough estimate from Step ❸.** Our result, 5.94%, is within the 16% to 2% range we estimated.

Example 2: If a person has an IQ of 90, what percentage of people have higher IQs than this person?

❶ **If you are beginning with a raw score, first change it to a *Z* score.** Using the usual formula, $Z = (90 - 100)/16 = -.63$.

❷ **Draw a picture of the normal curve, decide where the *Z* score falls on it, and shade in the area for which you are finding the percentage.** This is shown in Figure 6 (along with the exact percentages figured later).

❸ **Make a rough estimate of the shaded area's percentage based on the 50%–34%–14% percentages.** You know that 34% of the scores are between

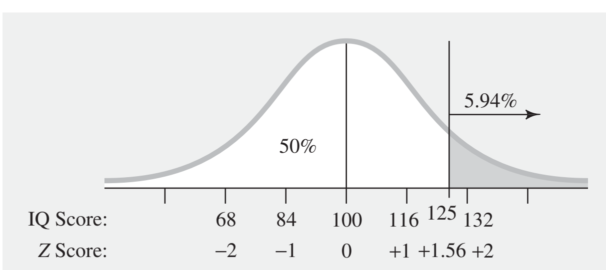


Figure 5 Distribution of IQ scores showing percentage of scores above an IQ score of 125 (shaded area).

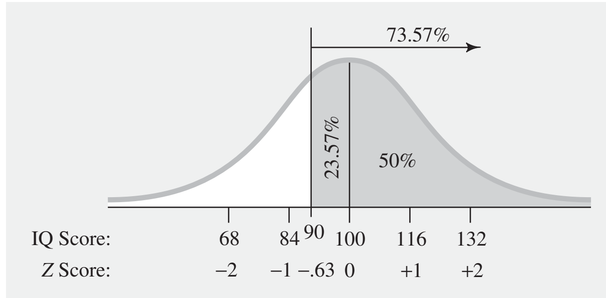


Figure 6 Distribution of IQ scores showing percentage of scores above an IQ score of 90 (shaded area).

the mean and a *Z* score of -1. Also, 50% of the curve is above the mean. Thus, the *Z* score of $-.63$ has to have between 50% and 84% of scores above it.

❹ **Find the exact percentage using the normal curve table, adding 50% if necessary.** The table shows that 23.57% of scores are between the mean and a *Z* score of .63. Thus, the percentage of scores above a *Z* score of $-.63$ is the 23.57% between the *Z* score and the mean plus the 50% above the mean, which is 73.57%.

❺ **Check that your exact percentage is within the range of your rough estimate from Step ❸.** Our result of 73.57% is within the 50% to 84% range we estimated.

Figuring *Z* Scores and Raw Scores from Percentages Using the Normal Curve Table

Going from a percentage to a *Z* score or raw score is similar to going from a *Z* score or raw score to a percentage. However, you reverse the procedure when figuring the exact percentage. Also, any necessary changes from a *Z* score to a raw score are done at the end.

Here are the five steps:

❶ **Draw a picture of the normal curve and shade in the approximate area for your percentage using the 50%–34%–14% percentages.**

❷ **Make a rough estimate of the *Z* score where the shaded area stops.**

❸ **Find the exact *Z* score using the normal curve table (subtracting 50% from your percentage if necessary before looking up the *Z* score).** Looking at your picture, figure out either the percentage in the shaded tail or the percentage between the mean and where the shading stops. For example, if your percentage is the bottom 35%, then the percentage in the shaded tail is 35%. Figuring the percentage between the mean and where the shading stops will sometimes involve subtracting 50% from the percentage in the problem. For example, if your percentage is the top 72%, then the percentage from the mean to where that shading stops is 22% (that is, $72\% - 50\% = 22\%$).

Once you have the percentage, look up the closest percentage in the appropriate column of the normal curve table ("% Mean to *Z*" or "% in Tail") and find the *Z* score for that percentage. That *Z* will be your answer—except it may be negative. The best way to tell if it is positive or negative is by looking at your picture.

❹ **Check that your exact *Z* score is within the range of your rough estimate from Step ❷.**

❺ **If you want to find a raw score, change it from the *Z* score.** Use the usual formula, $X = (Z)(SD) + M$.

Examples

Here are three examples. Once again, we will use IQ for our examples, where $M = 100$ and $SD = 16$.

Example 1: What IQ score would a person need to be in the top 5%?

❶ **Draw a picture of the normal curve and shade in the approximate area for your percentage using the 50%–34%–14% percentages.** We wanted the top 5%. Thus, the shading has to begin above (to the right of) 1 *SD* (there are 16% of scores above 1 *SD).* However, it cannot start above 2 *SD* because only 2% of all the scores are above 2 *SD*. But 5% is a lot closer to 2% than to 16%. Thus, you would start shading a small way to the left of the 2 *SD* point. This is shown in Figure 7.

❷ **Make a rough estimate of the *Z* score where the shaded area stops.** The *Z* score has to be between +1 and +2.

❸ **Find the exact *Z* score using the normal curve table (subtracting 50% from your percentage if necessary before looking up the *Z* score).** We want the top 5%, which means we can use the "% in Tail" column of the normal curve table. Looking in that column, the closest percentage to 5% is 5.05% (or you could use 4.95%). This goes with a *Z* score of 1.64 in the "*Z*" column.

❹ **Check that your exact *Z* score is within the range of your rough estimate from Step ❷.** As we estimated, +1.64 is between +1 and +2 (and closer to 2).

❺ **If you want to find a raw score, change it from the *Z* score.** Using the formula, $X = (Z)(SD) + M = (1.64)(16) + 100 = 126.24$. In sum, to be in the top 5%, a person would need an IQ of at least 126.24.

Example 2: What IQ score would a person need to be in the top 55%?

❶ **Draw a picture of the normal curve and shade in the approximate area for your percentage using the 50%–34%–14% percentages.** You want the top 55%. There are 50% of scores above the mean. So, the shading has to begin below (to the left of) the mean. There are 34% of scores between the mean and 1 *SD* below the mean, so the score is between the mean and 1 *SD* below the mean. You would shade the area to the right of that point. This is shown in Figure 8.

❷ **Make a rough estimate of the *Z* score where the shaded area stops.** The *Z* score has to be between 0 and −1.

❸ **Find the exact *Z* score using the normal curve table (subtracting 50% from your percentage if necessary before looking up the *Z* score).** Being in the top 55% means that 5% of people have IQs between this IQ and the mean (that is, $55\% - 50\% = 5\%$). In the normal curve table, the closest percentage to 5% in the "% Mean to *Z*" column is 5.17%, which goes with a *Z* score of .13. Because you are below the mean, this becomes −.13.

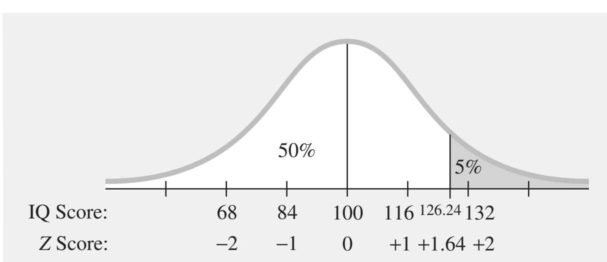


Figure 7 Finding the *Z* score and IQ raw score for where the top 5% starts.

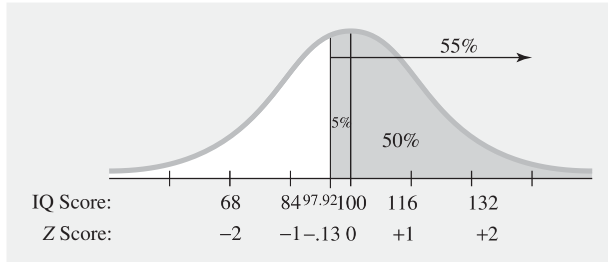


Figure 8 Finding the IQ score for where the top 55% start.

❹ **Check that your exact *Z* score is within the range of your rough estimate from Step** ❷. As we estimated, $-.13$ is between 0 and -1.

❺ **If you want to find a raw score, change it from the *Z* score.** Using the usual formula, $X = (-.13)(16) + 100 = 97.92$. So, to be in the top 55% on IQ, a person needs an IQ score of 97.92 or higher.

Example 3: What range of IQ scores includes the 95% of people in the middle range of IQ scores?

This kind of problem, of finding the middle percentage, may seem odd at first. However, it is actually a very common situation used in procedures. Think of this kind of problem in terms of finding the scores that go with the upper and lower ends of this percentage. Thus, in this example, you are trying to find the points where the bottom 2.5% ends and the top 2.5% begins (which, out of 100%, leaves the middle 95%).

❶ **Draw a picture of the normal curve, and shade in the approximate area for your percentage using the 50%–34%–14% percentages.** Let's start where the top 2.5% begins. This point has to be higher than 1 *SD* (16% of scores are higher than 1 *SD*). However, it cannot start above 2 *SD* because there are only 2% of scores above 2 *SD*. But 2.5% is very close to 2%. Thus, the top 2.5% starts just to the left of the 2 *SD* point. Similarly, the point where the bottom 2.5% comes in is just to the right of -2 *SD*. The result of all this is that we will shade in the area starting just above -2 *SD* and continue shading up to just below $+2$ *SD*. This is shown in Figure 9.

❷ **Make a rough estimate of the *Z* score where the shaded area stops.** You can see from the picture that the *Z* score for where the shaded area stops above the mean is just below $+2$. Similarly, the *Z* score for where the shaded area stops below the mean is just above -2.

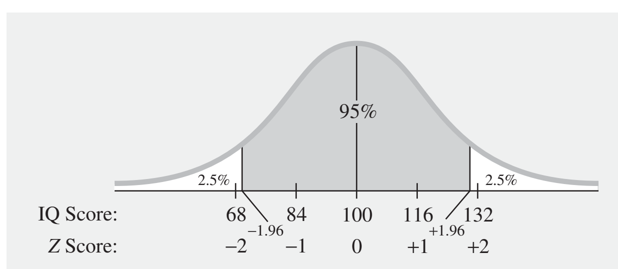


Figure 9 Finding the IQ scores for where the middle 95% of scores begin and end.

❸ **Find the exact *Z* score using the normal curve table (subtracting 50% from your percentage if necessary before looking up the *Z* score).** Being in the top 2.5% means that 2.5% of the IQ scores are in the upper tail. In the normal curve table, the closest percentage to 2.5% in the "% in Tail" column is exactly 2.50%, which goes with a *Z* score of +1.96. The normal curve is symmetrical. Thus, the *Z* score for the lower tail is −1.96.

❹ **Check that your exact *Z* score is within the range of your rough estimate from Step ❷.** As we estimated, +1.96 is between +1 and +2 and is very close to +2, and −1.96 is between −1 and −2 and very close to −2.

❺ **If you want to find a raw score, change it from the *Z* score.** For the high end, using the usual formula, $X = (1.96)(16) + 100 = 131.36$. For the low end, $X = (-1.96)(16) + 100 = 68.64$. In sum, the middle 95% of IQ scores run from 68.64 to 131.36.

How are you doing?

1. Why is the normal curve (or at least a curve that is symmetrical and unimodal) so common in nature?
2. Without using a normal curve table, about what percentage of scores on a normal curve are (a) above the mean, (b) between the mean and 1 *SD* above the mean, (c) between 1 and 2 *SD* above the mean, (d) below the mean, (e) between the mean and 1 *SD* below the mean, and (f) between 1 and 2 *SD* below the mean?
3. Without using a normal curve table, about what percentage of scores on a normal curve are (a) between the mean and 2 *SD* above the mean, (b) below 1 *SD* above the mean, (c) above 2 *SD* below the mean?
4. Without using a normal curve table, about what *Z* score would a person have who is at the start of the top (a) 50%, (b) 16%, (c) 84%, (d) 2%?
5. Using the normal curve table, what percentage of scores are (a) between the mean and a *Z* score of 2.14, (b) above 2.14, (c) below 2.14?
6. Using the normal curve table, what *Z* score would you have if (a) 20% are above you, (b) 80% are below you?

Answers

1. It is common because any particular score is the result of the random combination of many effects, some of which make the score larger and some of which make the score smaller. Thus, on average these effects balance out to produce scores near the middle, with relatively few scores at each extreme, because it is unlikely for most of the increasing or decreasing effects to come out in the same direction.
2. (a) Above the mean: 50%, (b) between the mean and 1 *SD* above the mean: 34%, (c) between 1 and 2 *SD* above the mean: 14%, (d) below the mean: 50%, (e) between the mean and 1 *SD* below the mean: 34%, (f) between 1 and 2 *SD* below the mean: 14%.
3. (a) Between the mean and 2 *SD* above the mean: 48%, (b) below 1 *SD* above the mean: 84%, (c) above 2 *SD* below the mean: 98%.
4. Above start of top (a) 50%: 0, (b) 16%: 1, (c) 84%: −1, (d) 2%: 2.
5. (a) Between the mean and a *Z* score of 2.14: 48.38%, (b) above a *Z* score of 2.14: 1.62%, (c) below a *Z* score of 2.14: 98.38%.
6. (a) 20% above you: .84, (b) 80% below you: .84.

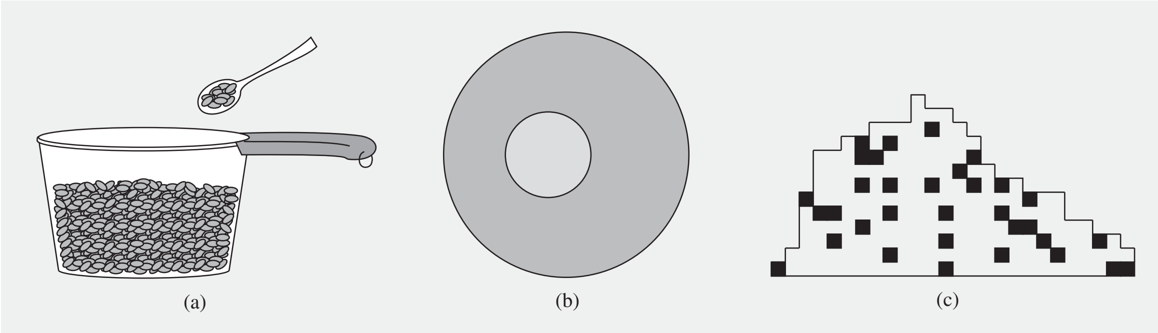


Figure 10 Populations and samples: In (a), the entire pot of beans is the population and the spoonful is a sample. In (b), the entire larger circle is the population and the circle within it is the sample. In (c), the histogram is of the population and the particular shaded scores together make up the sample.

Sample and Population

We are going to introduce you to some important ideas by thinking of beans. Suppose you are cooking a pot of beans and taste a spoonful to see if they are done. In this example, the pot of beans is a **population,** the entire set of things of interest. The spoonful is a **sample,** the part of the population about which you actually have information. This is shown in Figure 10a. Figures 10b and 10c are other ways of showing the relation of a sample to a population.

In behavioral and social science research, we typically study samples not of beans but of individuals to make inferences about some larger group. A sample might consist of 50 Canadian women who participate in a particular experiment; but the population might be intended to be all Canadian women. In an opinion survey, 1,000 people might be selected from the voting-age population and asked for whom they plan to vote. The opinions of these 1,000 people are the sample. The opinions of the larger voting public, to which the pollsters hope to apply their results, is the population (see Figure 11).[2]

Why Samples Instead of Populations Are Studied

If you want to know something about a population, your results would be most accurate if you could study the *entire population* rather than a *subgroup* from that population. However, in most research situations this is not practical. More important, the whole point of research is usually to be able to make generalizations or predictions about events beyond our reach. We would not call it research if we tested three particular cars to see which gets better gas mileage—unless you hoped to say something about the gas mileage of those models of cars in general. In other words, a researcher might do an experiment on the effect of a particular new method of teaching geography using 40 eighth-grade students as participants in the experiment. But the purpose of the experiment is not to find out how these *particular 40 students* respond to the experimental condition. Rather, the purpose is to learn something about how *eighth-grade students in general* respond to the new method of teaching geography.

population Entire group of people to which a researcher intends the results of a study to apply; the larger group to which inferences are made on the basis of the particular set of people (sample) studied.

sample Scores of the particular group of people studied; usually considered to be representative of the scores in some larger population.

[2]Strictly speaking, *population* and *sample* refer to scores (numbers or measurements), not to the people who have those scores. In the first example, the sample is really the scores of the 50 Canadian women, not the 50 women themselves, and the population is really what the *scores* would be if all Canadian women were tested.

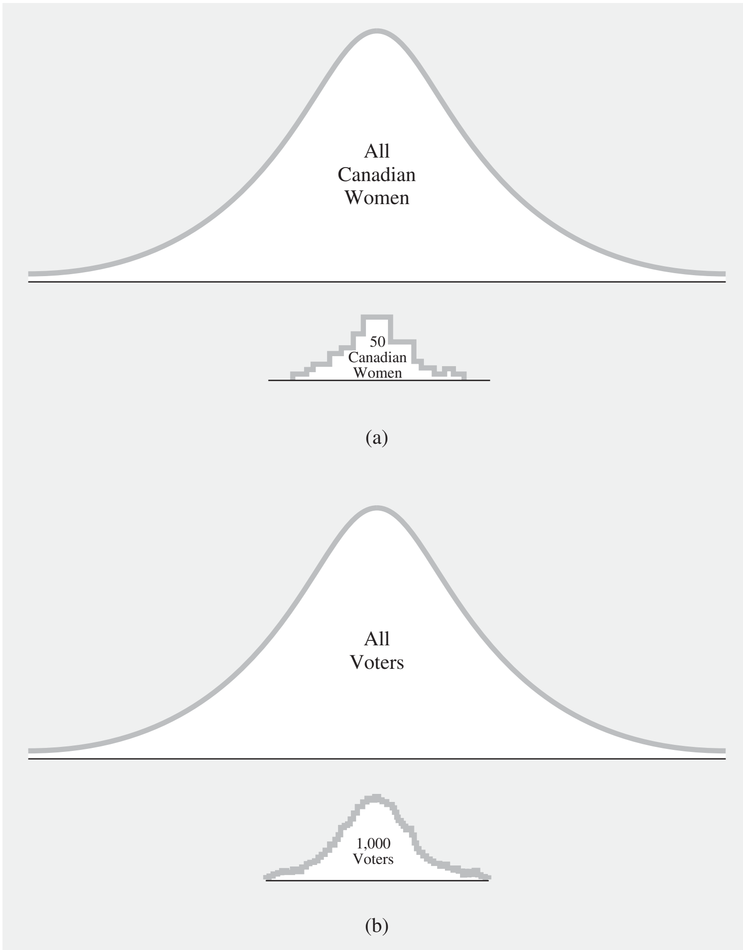


Figure 11 Additional examples of populations and samples. In (a), the population is the scores of all Canadian women and the sample is the scores of the 50 particular Canadian women studied. In (b), the population is the voting preferences of the entire voting-age population and the sample is the voting preferences of the 1,000 voting-age people who were surveyed.

The strategy in almost all behavioral and social science research is to study a sample of individuals who are believed to be representative of the general population (or of some particular population of interest). More realistically, researchers try to study people who do not differ from the general population in any systematic way that should matter for that topic of research.

The sample is what is studied, and the population is an unknown about which researchers draw conclusions based on the sample. Most of what you learn in the rest of this text is about the important work of drawing conclusions about populations based on information from samples.

random selection Method for selecting a sample that uses truly random procedures (usually meaning that each person in the population has an equal chance of being selected); one procedure is for the researcher to begin with a complete list of all the people in the population and select a group of them to study using a table of random numbers.

Methods of Sampling

Usually, the ideal method of picking out a sample to study is called **random selection.** The researcher starts with a complete list of the population and randomly selects some of them to study. An example of a random method of selection would be to put each name on a table tennis ball, put all the balls into a big hopper, shake it up, and have a blindfolded person select as many as are needed. (In practice, most researchers use a computer-generated list of random numbers.)

It is important not to confuse truly random selection with what might be called **haphazard selection;** for example, just taking whoever is available or happens to be first on a list. When using haphazard selection, it is surprisingly easy to pick accidentally a group of people that is very different from the population as a whole. Consider a survey of attitudes about your statistics instructor. Suppose you give your questionnaire only to other students sitting near you in class. Such a survey would be affected by all the things that influence where students choose to sit, some of which have to do with just what your survey is about—how much they like the instructor or the class. Thus, asking students who sit near you would likely result in opinions more like your own than a truly random sample would.

Unfortunately, it is often impractical or impossible to study a truly random sample. One problem is that we do not usually have a complete list of the full population. Another problem is that not everyone a researcher approaches agrees to participate in the study. Yet another problem is the cost and difficulty of tracking down people who live in remote places. For these and other reasons as well, behavioral and social scientists use various approximations to truly random samples that are more practical. However, researchers are consistently careful to rule out, as much as possible in advance, any systematic influence on who gets selected. Once the study has begun, researchers are constantly on the alert for any ways in which the sample may be systematically different from the population. For example, in much experimental research in education and psychology, it is common to use volunteer college students as the participants in the study. This is done for practical reasons and because often the topics studied in these experiments (for example, how short-term memory works) are thought to be relatively consistent across different types of people. Yet even in these cases, researchers avoid, for example, selecting people with a special interest in their research topic. Such researchers are also very aware that their results may not apply beyond college students, volunteers, people from their region, and so forth.

Methods of sampling is a complex topic that is discussed in detail in research methods textbooks (also see Box 1).

Statistical Terminology for Samples and Populations

haphazard selection Procedure of selecting a sample of individuals to study by taking whoever is available or happens to be first on a list.

population parameter Actual value of the mean, standard deviation, and so on, for the population (usually population parameters are not known, though often they are estimated based on information in samples).

sample statistic Descriptive statistic, such as the mean or standard deviation, figured from the scores in a particular group of people studied.

The mean, variance, and standard deviation of a population are called **population parameters.** A population parameter usually is unknown and can be estimated only from what you know about a sample taken from that population. You do not taste all the beans, just the spoonful. "The beans are done" is an inference about the whole pot.

In this text, when referring to the population mean, standard deviation, or variance, even in formulas, we use the word "Population"[3] before the M, SD^2, or SD. The mean, variance, and standard deviation you figure for the scores in a sample are called **sample statistics.** A sample statistic is figured from known information. Sample statistics are what we have been calculating all along. Sample statistics use the symbols we have been using all along: M, SD^2, and SD.

[3]In statistics writing, it is common to use Greek letters for population parameters. For example, the population mean is μ ("mu") and the population standard deviation is σ (lowercase "sigma"). However, we have not used these symbols in this text, wanting to make it easier for students to grasp the formulas without also having to deal with Greek letters.

How are you doing?

1. Explain the difference between the population and a sample for a research study.
2. Why do behavioral and social scientists usually study samples and not populations?
3. Explain the difference between random sampling and haphazard sampling.
4. Explain the difference between a population parameter and a sample statistic.

Answers

1. The population is the entire group to which results of a study are intended to apply. The sample is the particular, smaller group of individuals actually studied.
2. Behavioral and social scientists usually study samples and not populations because it is not practical in most cases to study the entire population.
3. In random sampling, the sample is chosen from among the population using a completely random method, so that any particular individual has an equal chance of being included in the sample. In haphazard sampling, the researcher selects individuals who are easily available or are convenient to study.
4. A population parameter is about the population (such as the mean of all the scores in the population); a sample statistic is about a particular sample (such as the mean of the scores of the people in the sample).

Probability

The purpose of most research in the behavioral and social sciences is to examine the truth of a theory or the effectiveness of a procedure. But scientific research of any kind can only make that truth or effectiveness seem more or less likely; it cannot give us the luxury of knowing for certain. Probability is very important in science. In particular, probability is very important in *inferential statistics,* the methods behavioral and social scientists use to go from results of research studies to conclusions about theories or applied procedures.

Probability has been studied for centuries by mathematicians and philosophers. Yet even today, the topic is full of controversy. Fortunately, however, you need to know only a few key ideas to understand and carry out the inferential statistical procedures that we teach.[4] These few key points are not very difficult; indeed, some students find them to be quite intuitive.

In statistics, we usually define **probability** as the expected relative frequency of a particular outcome. An **outcome** is the result of an experiment (or just about any situation in which the result is not known in advance, such as a coin coming up heads or it raining tomorrow). *Frequency* is how many times something happens. The *relative*

probability Expected relative frequency of a particular outcome; the proportion of successful outcomes to all outcomes.

outcome Term used in discussing probability for the result of an experiment (or almost any event, such as a coin coming up heads or it raining tomorrow).

[4]There are, of course, many probability topics that are not related to statistics and other topics related to statistics that are not important for the kinds of statistics used in the behavioral and social sciences. For example, computing joint and conditional probabilities, which is covered in many statistics books, is not covered here because it is rarely seen in published research in the behavioral and social sciences and is not necessary for an intuitive grasp of the logic of the major inferential statistical methods covered here.

BOX 1 Surveys, Polls, and 1948's Costly "Free Sample"

It is time to make you a more informed reader of polls in the media. Usually the results of properly done public polls are accompanied, somewhere in fine print, by a statement like "From a telephone poll of 1,000 American adults taken on June 4 and 5. Sampling error ± 3%." What does a statement like this mean?

The Gallup poll is as good an example as any (Gallup, 1972; see also www.galluppoll.com), and there is no better place to begin than in 1948, when all three of the major polling organizations—Gallup, Crossley (for Hearst papers), and Roper (for *Fortune*)—wrongly predicted Thomas Dewey's victory over Harry Truman for the U.S. presidency. Yet Gallup's prediction was based on 50,000 interviews and Roper's on 15,000. By contrast, to predict Barack Obama's 2008 victory, Gallup used only 3,050. Since 1952, the pollsters have never used more than 8,144—but with very small error and no outright mistakes. What has changed?

The method used before 1948, and never repeated since, was called "quota sampling." Interviewers were assigned a fixed number of persons to interview, with strict quotas to fill in all the categories that seemed important, such as residence, sex, age, race, and economic status. Within these specifics, however, they were free to interview whomever they liked. Republicans generally tended to be easier to interview. They were more likely to have telephones and permanent addresses and to live in better houses and better neighborhoods. In 1948, the election was very close, and the Republican bias produced the embarrassing mistake that changed survey methods forever.

Since 1948, all survey organizations have used what is called a "probability method." Simple random sampling is the purest case of the probability method, but simple random sampling for a survey about a U.S. presidential election would require drawing names from a list of all the eligible voters in the nation—a lot of people. Each person selected would have to be found, in diversely scattered locales. So instead, "multistage cluster sampling" is used. The United States is divided into seven size-of-community groupings, from large cities to rural open country; these groupings are divided into seven geographic regions (New England, Middle Atlantic, etc.), after which smaller equal-sized groups are zoned, and then city blocks are drawn from the zones, with the probability of selection being proportional to the size of the population or number of dwelling units. Finally, an interviewer is given a randomly selected starting point on the map and is required to follow a given direction, taking households in sequence.

Actually, telephoning is often the favored method for polling today. Phone surveys cost about one-third of door-to-door polls. Since most people now own phones, this method is less biased than in Truman's time. Phoning also allows computers to randomly dial phone numbers and, unlike telephone directories, this method calls unlisted numbers. However, survey organizations in the United States typically do not call cell phone numbers. Thus, U.S. households that use a cell phone for all calls and do not have a home phone are not usually included in telephone opinion polls. Most survey organizations consider the current cell-phone-only rate to be low enough not to cause large biases in poll results (especially since the demographic characteristics of individuals without a home phone suggest that they are less likely to vote than individuals who live in households with a home phone). However, anticipated future increases in the cell-phone-only rate will likely make this an important issue for opinion polls. Survey organizations are also starting to explore the use of other polling methods, such as email and the Internet.

Whether by telephone or face to face, there will be about 35% nonrespondents after three attempts. This creates yet another bias, dealt with through questions about how much time a person spends at home, so that a slight extra weight can be given to the responses of those reached but usually at home less, to make up for those missed entirely.

Now you know quite a bit about opinion polls, but we have left two important questions unanswered: Why are only about 1,000 included in a poll meant to describe all U.S. adults, and what does the term *sampling error* mean?

expected relative frequency In figuring probabilities, number of successful outcomes divided by the number of total outcomes you would expect to get if you repeated an experiment a large number of times.

frequency is the number of times something happens relative to the number of times it could have happened. That is, relative frequency is the proportion of times something happens. (A coin might come up heads 8 times out of 12 flips, for a relative frequency of 8/12, or 2/3.) **Expected relative frequency** is what you expect to get in the long run, if you repeat the experiment many times. (In the case of a coin, in the long run you expect to get 1/2 heads.) This is called the **long-run relative-frequency interpretation of probability.**

Figuring Probabilities

Probabilities are usually figured as the proportion of successful possible outcomes: the number of possible successful outcomes divided by the number of *all* possible outcomes. That is,

$$\text{Probability} = \frac{\text{Possible successful outcomes}}{\text{All possible outcomes}}$$

Consider the probability of getting heads when flipping a coin. There is one possible successful outcome (getting heads) out of two possible outcomes (getting heads or getting tails). This makes a probability of 1/2, or .5. In a throw of a single die, the probability of a 2 (or any other particular side of the six-sided die) is 1/6, or .17. This is because there is one possible successful outcome out of six possible outcomes of any kind. The probability of throwing a die and getting a number 3 or lower is 3/6, or .5. There are three possible successful outcomes (a 1, a 2, or a 3) out of six possible outcomes.

Now consider a slightly more complicated example. Suppose a class has 200 people in it, and 30 are seniors. If you were to pick someone from the class at random, the probability of picking a senior would be 30/200, or .15. This is because there are 30 possible successful outcomes (getting a senior) out of 200 possible outcomes.

Steps for Figuring Probabilities

There are three steps for figuring probabilities:

❶ **Determine the number of possible successful outcomes.**
❷ **Determine the number of all possible outcomes.**
❸ **Divide the number of possible successful outcomes (Step ❶) by the number of all possible outcomes (Step ❷).**

Let's apply these three steps to the probability of getting a number 3 or lower on a throw of a die:

❶ **Determine the number of possible successful outcomes.** There are three outcomes of 3 or lower: 1, 2, or 3.
❷ **Determine the number of possible outcomes.** There are six possible outcomes in the throw of a die: 1, 2, 3, 4, 5, or 6.
❸ **Divide the number of possible successful outcomes (Step ❶) by the number of all possible outcomes (Step ❷).** $3/6 = .5$.

Range of Probabilities

A probability is a proportion, the number of possible successful outcomes to the total number of possible outcomes. A proportion cannot be less than 0 or greater than 1. In terms of percentages, proportions range from 0% to 100%. Something that has no chance of happening has a probability of 0, and something that is certain to happen has a probability of 1. Notice that when the probability of an event is 0, the event is completely *impossible;* it cannot happen. But when the probability of an event is low, say .05 (5%) or even .01 (1%), the event is *improbable* or *unlikely,* but not impossible.

long-run relative-frequency interpretation of probability Understanding of probability as the proportion of a particular outcome that you would get if the experiment were repeated many times.

TIP FOR SUCCESS

To change a proportion into a percentage, multiply by 100. So, a proportion of .13 is equivalent to $.13 * 100 = 13\%$. To change a percentage into a proportion, divide by 100. So, 3% is a proportion of $3/100 = .03$.

"*p*" for Probability

Probability usually is symbolized by the letter ***p***. The actual probability number is usually given as a decimal, though sometimes fractions are used. A 50–50 chance is usually written as $p = .5$, but it could also be written as $p = 1/2$ or $p = 50\%$. It is also common to see a probability written as being less than some number using the "less than" sign. For example, $p < .05$ means "the probability is less than .05."

Probability, *Z* Scores, and the Normal Distribution

So far, we have mainly discussed probabilities of specific events that might or might not happen. We also can talk about a range of events that might or might not happen. The throw of a die coming out 3 or lower is an example (it includes the range 1, 2, and 3). Another example is the probability of selecting someone on a city street who is between the ages of 30 and 40.

If you think of probability in terms of the proportion of scores, probability fits in well with frequency distributions. In the frequency distribution shown in Figure 12, 3 of the total of 50 people scored 9 or 10. If you were selecting people from this group of 50 at random, there would be 3 chances (possible successful outcomes) out of 50 (all possible outcomes) of selecting one that was 9 or 10, so $p = 3/50 = .06$.

You can also think of the normal distribution as a probability distribution. With a normal curve, the percentage of scores between any two *Z* scores is known. The percentage of scores between any two *Z* scores is the same as the probability of selecting a score between those two *Z* scores. As you saw, approximately 34% of scores are between the mean and 1 standard deviation from the mean. You should therefore not be surprised to learn that the probability of a score being between the mean and a *Z* score of +1 (1 standard deviation above the mean) is .34 (that is, $p = .34$).

In a previous example in the normal curve section of this chapter, we figured that 95% of the scores in a normal curve are between a *Z* score of +1.96 and a *Z* score of −1.96 (see Figure 9). Thus, if you were to select a score at random from a distribution that follows a normal curve, the probability of selecting a score *between* *Z* scores of +1.96 and −1.96 is .95 (that is, a 95% chance). This is a very high probability.

p Probability.

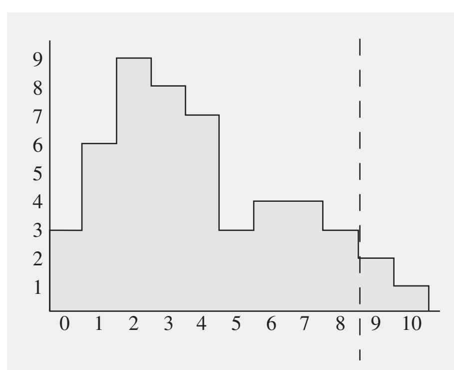


Figure 12 Frequency distribution (shown as a histogram) of 50 people in which $p = .06$ (3/50) of randomly selecting a person with a score of 9 or 10.

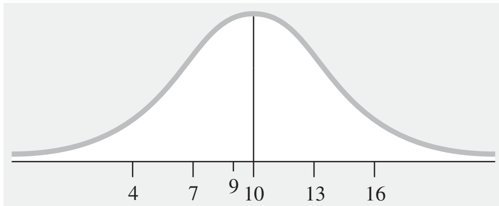


Figure 13 A normal distribution with a mean of 10 and a standard deviation of 3.

Also, the probability of selecting a score from such a distribution that is either *greater than* a *Z* score of +1.96 or *less than* a *Z* score of −1.96 is .05 (that is, a 5% chance). This is a very low probability. It helps to think about this visually. If you look back to Figure 9, the .05 probability of selecting a score that is either *greater than* a *Z* score of +1.96 or *less than* a *Z* score of −1.96 is represented by the unshaded tail areas in the figure.

Probability, Samples, and Populations. Probability is also relevant to samples and populations. We will use an example to give you a sense of the role of probability in samples and populations. Imagine you are told that a sample of one person has a score of 4 on a certain measure. However, you do not know whether this person is from a population of women or of men. You are told that a population of women has scores on this measure that are normally distributed with a mean of 10 and a standard deviation of 3. How likely do you think it is that your sample of one person comes from this population of women? As you can see in Figure 13, there are very few scores as low as 4 in a normal distribution that has a mean of 10 and a standard deviation of 3. So, there is a very low likelihood that this person comes from the population of women. Now, what if the sample person had a score of 9? In this case, there is a much greater likelihood that this person comes from the population of women because there are many scores of 9 in that population. This kind of reasoning provides an introduction to the process of *hypothesis testing*.

How are you doing?

1. The probability of an event is defined as the expected relative frequency of a particular outcome. Explain what is meant by (a) relative frequency and (b) outcome.
2. Suppose you have 400 pennies in a jar and 40 of them are more than 9 years old. You then mix up the pennies and pull one out. (a) What is the number of possible successful outcomes? (b) What is the number of all possible outcomes? (c) What is the probability of getting one that is more than 9 years old?
3. Suppose people's scores on a particular measure are normally distributed with a mean of 50 and a standard deviation of 10. If you were to pick a person completely at random, what is the probability you would pick someone with a score on this measure higher than 60?
4. What is meant by $p < .01$?

Answers

1. (a) Relative frequency is the number of times something happens in relation to the number of times it could have happened. (b) An outcome is the result of an experiment—what happens in a situation where what will happen is not known in advance.
2. (a) The number of possible successful outcomes is 40. (b) The number of all possible outcomes is 400. (c) The probability of getting one of the 40 pennies that are more than 9 years old is: $p = 40/400 = .10$.
3. The probability you would pick someone with a score on this test higher than 60 is $p = .16$ (since 16% of the scores are more than 1 standard deviation above the mean).
4. The probability is less than .01.

Normal Curves, Samples and Populations, and Probabilities In Research Articles

You need to understand the topics we covered in this chapter to learn what comes next. However, the topics of this chapter are rarely mentioned directly in research articles (except articles *about* methods or statistics). Sometimes you will see the normal curve mentioned, usually when a researcher is describing the pattern of scores on a particular variable.

Probability is also rarely discussed directly, except in relation to statistical significance. In almost any article you look at, the Results section will have many descriptions of various methods associated with statistical significance, followed by something like "$p < .05$" or "$p < .01$." The p refers to probability, but the probability of what?

Finally, you will sometimes see a brief mention of the method of selecting the sample from the population. For example, Bennett, Wolin, Puleo, Mâsse, and Atienza (2009) used data from a national survey to examine awareness of national physical activity recommendations for health promotion among adults. They described the method of their study as follows:

> The Health Information National Trends Survey (HINTS) is a probability-based sample of the US population contacted via a random-digit dial (RDD) telephone survey conducted biennially (since 2003) by the National Cancer Institute (NCI). Data from the 2005 HINTS ($n = 5586$) were used in this analysis; the development and details and the sampling design have been described in greater detail elsewhere . . . . Briefly, during a period of 25 wk in 2005, trained interviewers collected data using a list-assisted, RDD sample of all telephone exchanges in the United States. One adult (aged 18 yr and older) per household was selected to complete an extended interview. (p. 1850)

Whenever possible, researchers report the proportion of individuals approached for the study who actually participated in the study. This is called the *response rate.* Bennett and colleagues reported that the response rate for the extended screening interview was 61.25%, and the rate was calculated based on guidelines from the American Association for Public Opinion Research (2001).

Researchers sometimes also check whether their sample is similar to the population as a whole, based on any information they may have about the overall population.

For example, Schuster and colleagues (2001) conducted a national survey of stress reactions of U.S. adults after the September 11, 2001, terrorist attacks. In this study, the researchers compared their sample to 2001 census records and reported that the "sample slightly overrepresented women, non-Hispanic whites, and persons with higher levels of education and income" (p. 1507). Schuster and colleagues went on to note that overrepresentation of these groups "is typical of samples selected by means of random-digit dialing" (pp. 1507–1508).

Learning Aids

Summary

1. The scores on many variables in behavioral and social science research approximately follow a bell-shaped, symmetrical, unimodal distribution called the *normal curve* (or *normal distribution*). Because the shape of this curve follows an exact mathematical formula, there is a specific percentage of scores between any two points on a normal curve.
2. A useful working rule for normal curves is that 50% of the scores are above the mean, 34% are between the mean and 1 standard deviation above the mean, and 14% are between 1 and 2 standard deviations above the mean.
3. A normal curve table gives the percentage of scores between the mean and any particular Z score, as well as the percentage of scores in the tail for any Z score. Using this table, and knowing that the curve is symmetrical and that 50% of the scores are above the mean, you can figure the percentage of scores above or below any particular Z score. You can also use the table to figure the Z score for the point where a particular percentage of scores begins or ends.
4. A *sample* is an individual or group that is studied—usually as representative of a larger group or *population* that cannot be studied in its entirety. Ideally, the sample is selected from a population using a strictly random procedure. The mean, variance, and so forth of a sample are called *sample statistics;* when of a population, they are called *population parameters.*
5. Most behavioral and social scientists consider the probability of an event to be its expected relative frequency. Probability is figured as the proportion of successful outcomes to total possible outcomes. It is symbolized by p and has a range from 0 (event is impossible) to 1 (event is certain). The normal curve provides a way to know the probabilities of scores' being within particular ranges of values.
6. Research articles rarely discuss normal curves (except briefly when the variable being studied seems not to follow a normal curve) or probability (except in relation to statistical significance). Procedures of sampling are sometimes described, particularly when the study is a survey, and the representativeness of a sample may also be discussed.

normal distribution
normal curve
normal curve table
population
sample
random selection
haphazard selection
population parameters
sample statistic
probability (p)
outcome
expected relative frequency
long-run relative-frequency interpretation of probability

Example Worked-Out Problems

Figuring the Percentage above or below a Particular Raw Score or Z Score

Suppose a test of sensitivity to violence is known to have a mean of 20, a standard deviation of 3, and a normal curve shape. What percentage of people have scores above 24?

Answer

❶ **If you are beginning with a raw score, first change it to a *Z* score.** Using the usual formula, $Z = (X - M)/SD$, $Z = (24 - 20)/3 = 1.33$.

❷ **Draw a picture of the normal curve, decide where the *Z* score falls on it, and shade in the area for which you are finding the percentage.** This is shown in Figure 14.

❸ **Make a rough estimate of the shaded area's percentage based on the 50%–34%–14% percentages.** If the shaded area started at a *Z* score of 1, it would include 16%. If it started at a *Z* score of 2, it would include only 2%. So with a *Z* score of 1.33, it has to be somewhere between 16% and 2%.

❹ **Find the exact percentage using the normal curve table, adding 50% if necessary.** In Table 1 of the appendix "Tables," 1.33 in the "*Z*" column goes with 9.18% in the "% in Tail" column. This is the answer to our problem: 9.18% of people have a higher score than 24 on the sensitivity to violence measure. (There is no need to add 50% to the percentage.)

❺ **Check that your exact percentage is within the range of your rough estimate from Step ❸.** Our result, 9.18%, is within the 16% to 2% range estimated.

Note: If the problem involves *Z* scores that are all exactly 0, 1, or 2 (or −1 or −2), you can work the problem using the 50%–34%–14% figures and without using the normal curve table. However, you should still draw a figure and shade in the appropriate area.

Figuring Z Scores and Raw Scores from Percentages

Consider the same situation: A test of sensitivity to violence is known to have a mean of 20, a standard deviation of 3, and a normal curve shape. What is the minimum score a person needs to be in the top 75%?

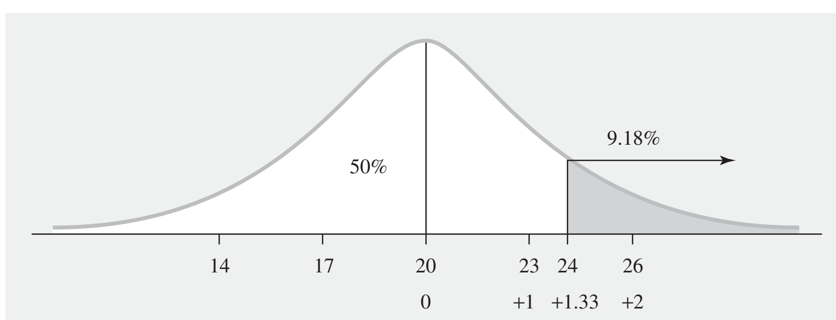


Figure 14 Distribution of sensitivity to violence scores showing percentage of scores above a score of 24 (shaded area).

Answer

❶ **Draw a picture of the normal curve and shade in the approximate area for your percentage using the 50%–34%–14% percentages.** The shading has to begin between the mean and 1 *SD* below the mean. (There are 50% above the mean and 84% above 1 *SD* below the mean.) This is shown in Figure 15.

❷ **Make a rough estimate of the *Z* score where the shaded area stops.** The *Z* score has to be between 0 and −1.

❸ **Find the exact *Z* score using the normal curve table (subtracting 50% from your percentage if necessary before looking it up).** Since 50% of people have IQs above the mean, for the top 75% you need to include the 25% below the mean (that is, 75% − 50% = 25%). Looking in the "% Mean to *Z*" column of the normal curve table, the closest figure to 25% is 24.86, which goes with a *Z* of .67. Since we are interested in below the mean, we want −.67.

❹ **Check that your exact *Z* score is similar to your rough estimate from Step ❷.** −.67 is between 0 and −1.

❺ **If you want to find a raw score, change it from the *Z* score.** Using the usual formula, $X = (Z)(SD) + M$, $X = (-.67)(3) + 20 = -2.01 + 20 = 17.99$. That is, to be in the top 75%, a person needs to have a score on this test of at least 18.

Note: If the problem instructs you not to use a normal curve table, you should be able to work the problem using the 50%–34%–14% figures. However, you should still draw a figure and shade in the appropriate area.

Outline for Writing Essays on the Logic and Computations for Figuring a Percentage from a *Z* Score and Vice Versa

1. Note that the normal curve is a mathematical (or theoretical) distribution, describe its shape (be sure to include a diagram of the normal curve), and mention that many variables in nature and in the behavioral and social sciences approximately follow a normal curve.
2. If necessary (that is, if required by the question), explain the mean and standard deviation.
3. Describe the link between the normal curve and the percentage of scores between the mean and any *Z* score. Be sure to include a description of the normal curve table and explain how it is used.

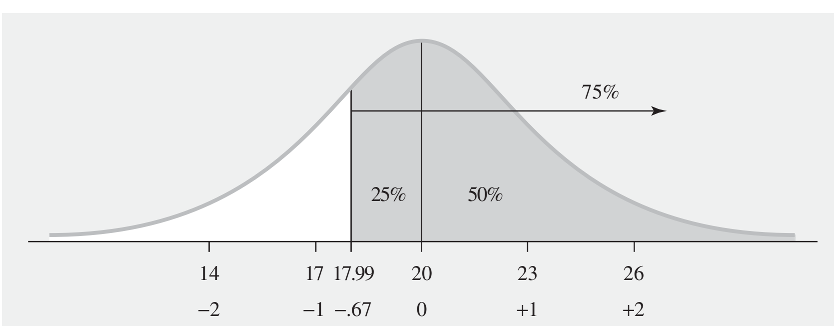


Figure 15 Finding the sensitivity to violence raw score for where the top 75% start.

4. Briefly describe the steps required to figure a percentage from a Z score or vice versa (as required by the question). Be sure to draw a diagram of the normal curve with appropriate numbers and shaded areas marked on it from the relevant question (e.g., the mean, 1 and 2 standard deviations above/below the mean, shaded area for which percentage or Z score is to be determined).

Finding a Probability

A candy dish has four kinds of fruit-flavored candy: 20 apple, 20 strawberry, 5 cherry, and 5 grape. If you close your eyes and pick one piece of candy at random, what is the probability it will be either cherry or grape?

Answer

❶ **Determine the number of possible successful outcomes.** There are 10 possible successful outcomes—5 cherry and 5 grape.
❷ **Determine the number of all possible outcomes.** There are 50 possible outcomes overall: $20 + 20 + 5 + 5 = 50$.
❸ **Divide the number of possible successful outcomes (Step ❶) by the number of all possible outcomes (Step ❷).** $10/50 = .2$. Thus, the probability of picking either a cherry or grape candy is .2.

These problems involve figuring. Most real-life statistics problems are done on a computer with special statistical software. Even if you have such software, do these problems by hand to ingrain the method in your mind. In this chapter, in addition, it is important to draw the diagrams for the normal curve problems.

All data are fictional unless an actual citation is given.

Set I (for answers, see the end of this chapter)

1. Suppose the people living in a particular city have a mean score of 40 and a standard deviation of 5 on a measure of concern about the environment. Assume that these concern scores are normally distributed. Using the 50%–34%–14% figures, approximately what percentage of people in this city have a score (a) above 40, (b) above 45, (c) above 30, (d) above 35, (e) below 40, (f) below 45, (g) below 30, (h) below 35?
2. Using the information in problem 1 and the 50%–34%–14% figures, what is the minimum score a person has to have to be in the top (a) 2%, (b) 16%, (c) 50%, (d) 84%, and (e) 98%?
3. An education researcher has been studying test anxiety using a particular measure, which she administers to students prior to midterm exams. On this measure, she has found that the distribution follows a normal curve. Using a normal curve table, what percentage of students have Z scores (a) below 1.5, (b) above 1.5, (c) below -1.5, (d) above -1.5, (e) above 2.10, (f) below 2.10, (g) above .45, (h) below -1.78, and (i) above 1.68?
4. In the previous problem, the measure of test anxiety has a mean of 15 and a standard deviation of 5. Using a normal curve table, what percentage of students have scores (a) above 16, (b) above 17, (c) above 18, (d) below 18, (e) below 14?

5. In the test anxiety example of problems 3 and 4, using a normal curve table, what is the lowest score on the test anxiety measure a person has to have to be in (a) the top 40%, (b) the top 30%, (c) the top 20%?
6. Using a normal curve table, give the percentage of scores between the mean and a Z score of (a) .58, (b) .59, (c) 1.46, (d) 1.56, (e) $-.58$.
7. Assuming a normal curve, (a) if a student is in the bottom 30% of the class on Spanish ability, what is the highest Z score this person could have? (b) If the person is in the bottom 3%, what would be the highest Z score this person could have?
8. Assuming a normal curve, (a) if a person is in the top 10% of the country on mathematics ability, what is the lowest Z score this person could have? (b) If the person is in the top 1%, what would be the lowest Z score this person could have?
9. Consider a test of coordination that has a normal distribution, a mean of 50, and a standard deviation of 10. (a) How high a score would a person need to be in the top 5%? (b) Explain your answer to someone who has never had a course in statistics.
10. A research article is concerned with the level of self-esteem of Australian high school students. The methods section emphasizes that a "random sample" of Australian high school students was surveyed. Explain to a person who has never had a course in statistics what this means and why it is important.
11. Tanski and colleagues (2009) conducted a telephone survey of smoking and exposure to smoking in movies among U.S. adolescents. In the method section of their article, they explained that "a nationally representative sample of 6522 US adolescents, 10 to 14 years of age, was recruited through a random-digit dial survey" (p. 136). Explain to a person who has never had a course in statistics or research methods what this means and why it is important.
12. The following numbers of employees in a company received special assistance from the personnel department last year:

Drug/alcohol	10
Family crisis counseling	20
Other	20
Total	50

If you were to select a score at random from the records for last year, what is the probability that it would be in each of the following categories? (a) drug/alcohol, (b) family, (c) drug/alcohol or family, (d) any category except "other," (e) any of the three categories? (f) Explain your answer to part (a) to someone who has never had a course in statistics.

Set II

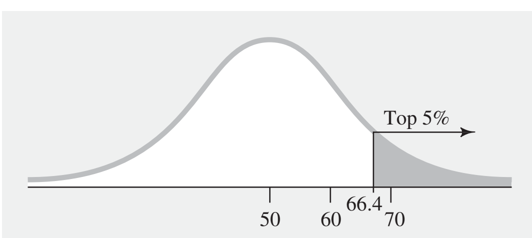
13. Consider a test that has a normal distribution, a mean of 100, and a standard deviation of 14. How high a score would a person need to be in the top (a) 1%, (b) the top 5%?
14. The length of time it takes to complete a word puzzle is found to be normally distributed with a mean of 80 seconds and a standard deviation of 10 seconds. Using the 50%–34%–14% figures, approximately what percentage of scores (on time to complete the word puzzle) will be (a) above 100, (b) below 100, (c) above 90, (d) below 90, (e) above 80, (f) below 80, (g) above 70, (h) below 70, (i) above 60, and (j) below 60?
15. Using the information in problem 14 and the 50%–34%–14% figures, what is the longest time to complete the word puzzle a person can have and still be in the bottom (a) 2%, (b) 16%, (c) 50%, (d) 84%, and (e) 98%?

16. Suppose that the scores of architects on a particular creativity test are normally distributed. Using a normal curve table, what percentage of architects have Z scores (a) above .10, (b) below .10, (c) above .20, (d) below .20, (e) above 1.10, (f) below 1.10, (g) above $-.10$, and (h) below $-.10$?
17. In the example in problem 16, using a normal curve table, what is the minimum Z score an architect can have on the creativity test to be in the (a) top 50%, (b) top 40%, (c) top 60%, (d) top 30%, and (e) top 20%?
18. In the example in problem 16, assume that the mean is 300 and the standard deviation is 25. Using a normal curve table, what scores would be the top and bottom score to find (a) the middle 50% of architects, (b) the middle 90% of architects, and (c) the middle 99% of architects?
19. Suppose that you are designing an instrument panel for a large industrial machine. The machine requires the person using it to reach 2 feet from a particular position. The reach from this position for adult women is known to have a mean of 2.8 feet with a standard deviation of .5. The reach for adult men is known to have a mean of 3.1 feet with a standard deviation of .6. Both women's and men's reach from this position is normally distributed. If this design is implemented, (a) what percentage of women will not be able to work on this instrument panel? (b) What percentage of men will not be able to work on this instrument panel? (c) Explain your answers to a person who has never had a course in statistics.
20. Suppose you want to conduct a survey of the voting preference of undergraduate students. One approach would be to contact every undergraduate student you know and ask them to fill out a questionnaire. (a) What kind of sampling method is this? (b) What is a major limitation of this kind of approach?
21. A large study of concerns about terrorism among individuals living in Los Angeles (Eisenmann et al., 2009) recruited participants using ". . . a random digit dial survey of the Los Angeles population October 2004 through January 2005 in 6 languages." These are procedures in which phone numbers to call potential participants are randomly generated by a computer. Explain to a person who has never had a course in statistics (a) why this method of sampling might be used and (b) why it may be a problem if not everyone called agreed to be interviewed.
22. Suppose that you were going to conduct a survey of visitors to your campus. You want the survey to be as representative as possible. (a) How would you select the people to survey? (b) Why would that be your best method?
23. You are conducting a survey at a college with 800 students, 50 faculty members, and 150 administrators. Each of these 1,000 individuals has a single listing in the campus phone directory. Suppose you were to cut up the directory and pull out one listing at random to contact. What is the probability it would be (a) a student, (b) a faculty member, (c) an administrator, (d) a faculty or administrator, and (e) anyone except an administrator? (f) Explain your answer to part (d) to someone who has never had a course in statistics.
24. You apply to 20 graduate programs, 10 of which are in the United States, 5 of which are in Canada, and 5 of which are in Australia. You get a message from home that you have a letter from one of the programs to which you applied, but nothing is said about which one. Give the probability it is from (a) a program in the United States, (b) a program in Canada, (c) from any program other than in Australia. (d) Explain your answer to (c) to someone who has never had a course in statistics.

Answers to Set I Practice Problems

1. (a) 50%; (b) 16%; (c) 98%; (d) 84%; (e) 50%; (f) 84%; (g) 2%; (h) 16%.
2. (a) 50; (b) 45; (c) 40; (d) 35; (e) 30.
3. (a) From the normal curve table, 43.32% (.4332) have Z scores between the mean and 1.5. By definition, 50% have Z scores below the mean. Thus, the total percentage below 1.5 is $50\% + 43.32\% = 93.32\%$; (b) 6.68%; (c) 6.68%; (d) 93.32%; (e) 1.79%; (f) 98.21%; (g) 32.64%; (h) 3.75%; (i) 4.65%.
4. (a) $Z = (16 - 15)/5 = .2$; from the normal curve table, percent in the tail for a Z score of $.2 = 42.07\%$; (b) 34.46%; (c) 27.43%; (d) 72.57%; (e) 42.07%.
5. (a) Top 40% means 40% in the tail; the nearest Z score from the normal curve table for 40% in the tail is .25; a Z score of .25 equals a raw score of $(.25)(5) + 15 = 16.25$; (b) 17.6; (c) 19.2.
6. (a) 21.90%; (b) 22.24%; (c) 42.79%; (d) 44.06%; (e) 21.90%.
7. (a) The bottom 30% means 30% in the tail; the nearest Z score from the normal curve table for 30% in the tail is $-.52$; (b) -1.88.
8. (a) The top 10% means 10% in the tail; the nearest Z score for 10% in the tail is 1.28; (b) 2.33.
9. (a) Needed $Z = 1.64$; this corresponds to a raw score of $(1.64)(10) + 50 = 66.4$; (b) The scores for many things you measure, in nature and the behavioral and social sciences, tend approximately to follow the particular mathematical pattern shown below, called a normal curve. In a normal curve, most of the scores are near the middle, with fewer numbers of scores at each extreme. Because the normal curve is mathematically defined, the precise proportion of scores in any particular section of it can be calculated, and these have been listed in special tables. (Then explain mean, standard deviation, and Z scores.) The normal curve table tells the percentage of scores in the normal curve between the mean and any particular Z score; it also tells the percentage of scores in the tail for (that is, more extreme than) any particular Z score.

   The coordination test scores are known to follow a normal curve. Thus, you can look up in the table the Z score for the point on the normal curve at which 5% of the scores are in the tail. This is a Z score of 1.64 (actually, there is not an exact point on the table for 5% in the tail, so I could have used either 1.64 or 1.65). With a standard deviation of 10, a Z score of 1.64 is 16.4 points above the mean. Adding that to the mean of 50 makes the score needed to be in the top 5% turn out to be 66.4.



10. A *sample* is a group of people studied that represents the entire group to which the results are intended to apply, called the *population.* (In this example, the population is all Australian high school students.) You study a sample because it would be impractical or impossible to test the entire population. One way of ensuring that the sample is not systematically unrepresentative is to select the sample randomly. This does not mean haphazardly. For example, just taking the students who are easily available to test would be haphazard sampling. But this would not be a good method because whatever factors make them easily available—such as living in a nearby town—might make them unrepresentative of the population as a whole. An example of a truly random method would be to acquire a list of all the high school students in Australia, number each student, and then use a table of random numbers to pick as many as are to be surveyed.
11. Similar to 10 above.
12. (a) 10/50: $p = 10/50 = .2$; (b) .4; (c) $(10 + 20)/50 = .6$; (d) .6; (e) 1. (f) The probability of a particular thing happening is usually defined as the number of possible ways the thing could happen (the number of *possible successful outcomes*) divided by the number of possible ways things like this could happen (the number of *all possible outcomes*). For part (a) there are 10 different drug/alcohol people you might pick out of a total of 50 people you are picking from. Thus, the probability is $10/50 = .2$.

Introduction to Hypothesis Testing

Chapter Outline

In this chapter, we introduce the crucial topic of **hypothesis testing.** A **hypothesis** is a prediction intended to be tested in a research study. The prediction may be based on informal observation (as in clinical or applied settings), on related results of previous studies, or on a broader *theory* about what is being studied. You can think of a **theory** as a set of principles that attempt to explain one or more facts, relationships, or events. A theory usually gives rise to various specific hypotheses that can be tested in research studies.

This chapter focuses on the basic logic for analyzing results of a research study to test a hypothesis. The central theme of hypothesis testing has to do with the important distinction between sample and population. Hypothesis testing is a systematic procedure for deciding whether the results of a research study, which examines a sample, support a hypothesis that applies to a population. Hypothesis testing is the central theme in most behavioral and social science research.

hypothesis testing Procedure for deciding whether the outcome of a study (results for a sample) supports a particular theory or practical innovation (which is thought to apply to a population).

hypothesis A prediction, often based on observation, previous research, or theory, that is tested in a research study.

theory A set of principles that attempt to explain one or more facts, relationships, or events; behavioral and social scientists often derive specific predictions (hypotheses) from theories that are then tested in research studies.