6

THE NORMAL DISTRIBUTION AND STANDARD SCORES

6.1 THE IMPORTANCE OF THE NORMAL DISTRIBUTION

The *normal distribution*, also known as the Gaussian curve or the normal probability curve, is the most fundamentally important distribution in statistics. The normal curve is used extensively in all subsequent chapters. Its use in this chapter will be illustrated by describing the performance of an individual or group using standard scores. Other more important applications of the normal distribution will be evident in the following chapters. Measures of *skewness* and *kurtosis*, which describe and quantify the extent to which a distribution deviates from a true normal distribution, will also be considered.

Historical Background

The study of the normal distribution dates from at least the seventeenth century. It was noticed, for example, that if an object were weighed repeatedly, the observed weights were not identical; there was some variation among the measurements. If enough measurements were taken, the distribution of the observations displayed a regular pattern, a pattern now recognized to be the normal distribution. Errors of observation of many kinds were found to follow this same pattern. In fact, the distribution was initially called the "normal curve of errors."

6.2 GOD LOVES THE NORMAL CURVE

It was soon discovered that observations other than measurement error resulted in normal, or approximately normal, curves. If ten fair coins were tossed, the number of heads recorded in the toss, and the procedure repeated many times (actually an infinite number), the distribution in Figure 6.1 would result. Note that the *expected value* (Section 5.13) for the number of heads is 5, which is the mean, μ, of the theoretical distribution shown in Figure

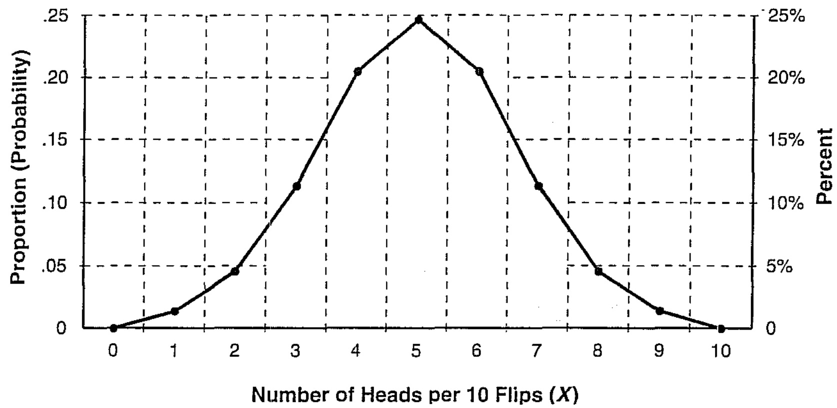


FIGURE 6.1 The Distribution of the Number of Heads in Ten Fair Coins Tossed Many Times.

6.1. In normal distributions the mean is also the mode—the value of μ occurs more frequently than any other score. Figure 6.1 shows that almost 25% of the sets of ten tosses results in five heads; but for 75% of the sets of ten flips, the number of heads is not five, but varies symmetrically about five; four and six heads were each observed in more than 20% of the sets. The distribution is symmetrical and approximately normal, but note that it does not result from errors of measurement, but from the laws of chance. No collection of empirical observations would look exactly like a perfect normal distribution because the latter is a mathematical abstraction. For example, the distribution of number of heads in Figure 6.1 is *discrete* (i.e., has gaps), not *continuous*: there are no points between 4 and 5, or between 5 and 6, for instance. The true normal distribution is continuous; there are no gaps. Variables like time, distance, and weight are continuous and can be measured so there are virtually no gaps. If the number of coins flipped were increased to 100, and the number of heads recorded in many repeated tosses, then the distribution of the number of heads would approach the mathematical normal distribution much more closely. Of course, the approximation would be even better if 1,000 fair coins were tossed.

Late in the nineteenth century Francis Galton, an Englishman, took systematic measurements of many physical, psychological, and psychomotor characteristics on large samples of persons and found that the distributions of the measurements were very close approximations to the normal distribution. Figure 6.2 illustrates his findings using the heights of 8,585 adult men born in Great Britain during the nineteenth century. Note how closely the distribution approximates a normal curve.

It is fortunate that the measurements of many variables in all disciplines have distributions that are good approximations of the normal distribution, for example, reaction times for children at ten years of age, size of wings of a given species of butterfly, heights for a given variety of plant, daily high temperature on January 15 for the past century. Refined measures of most cognitive, psychomotor, and many affective and other human

FIGURE 6.2 Frequency Polygon for Heights of 8,585 Adult Men Born in Great Britain During the Nineteenth Century ($\bar{X}$ = 67.02", s = 2.564").

characteristics have empirical distributions that are approximately normal. Indeed, God loves the normal curve.

No set of empirical observations is ever perfectly described by the normal distribution. Even if a variable were perfectly normally distributed, the observed distribution would never be perfectly normal. There is sampling error inherent in any finite set of data that will result in some departure from the mathematical curve, but the shape is often extremely close to the theoretical normal curve. The discrepancy is frequently so small that it can be disregarded for practical purposes.

Building on the work of Pascal (1623–1662), de Fermat (1601–1665), and Bernoulli (1654–1705), Abraham DeMoivre (1667–1754) was able to show that the mathematical curve that approximates the curve that connects the tops of the lines in Figures 6.1 and 6.2 is described by the following equation:

$$u = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}[(X-\mu)/\sigma]^2} \qquad (6.1)$$

where u is the height of the curve above any given value of the variable, X;
π is the ratio of the circumference of a circle to its diameter; $\pi = 3.14159 \ldots$;
e is the base of the system of natural logarithms, $e = 2.71828 \ldots$; and
μ and σ are the mean and the standard deviation of the variable, X.

Equation 6.1 is the equation for the *normal distribution*. We shall have little to do directly with Equation 6.1 as such, although it was used to generate the normal curve table, Table A in the Appendix.

The empirical distribution of IQ scores is almost perfectly normal between 70 and 130. Although the commonly observed empirical bell-shaped curves of errors, height, IQ, and other variables have piqued the curiosity of scientists of many different stripes, *the prominence of the normal distribution in inferential statistics is primarily due to its mathematical properties*. No other distribution has such desirable properties with which the mathematical statistician can do magic. Many technical problems in statistics have been solved only by assuming the observations in the population are normally distributed. Specific instances will appear in later chapters.

The ubiquity of the normal curve sometimes leads to the mistaken notion that there is a necessary link between it and almost any good set of data, but many variables are definitely *not* normally distributed. For example, many sociological variables, such as social class, socioeconomic status, income, level of education, and family size, are skewed. Certain social and political attitudes, such as attitude toward abortion, have bimodal distributions. Such variables as age, ethnicity, religion, and college major obviously are not normally distributed.

The graph of Equation 6.1 yields the familiar, symmetric, bell-shaped curve known as the *normal curve*. One speaks of a normal curve, because Equation 6.1 imparts a characteristic shape to the graph. All normal curves have the following properties: unimodal, symmetry, points of inflection[1] at $\mu \pm \sigma$, tails that approach (but never quite touch) the horizontal axis[2] as they deviate from μ. Not all curves with these characteristics are normal (as you will see in Section 6.9). The normal curve has a smooth, altogether handsome countenance—a thing of beauty.

6.3 THE STANDARD NORMAL DISTRIBUTION AS A STANDARD REFERENCE DISTRIBUTION: *z*-SCORES

A raw score of 42 on a test means little, but to know that this score is 1½ standard deviations (1.5σ) above the mean tells us that it is quite high relative to the other scores in the distribution. If the mean and standard deviation are known, the individual scores can be pictured relative to the entire set of scores in the distribution. Observations expressed in standard deviation units from the mean are termed *z-scores*. For IQ scores where μ is 100 and σ is 15, an IQ score of 130 can be transformed to a *z-score* of 2. It is two standard deviations above the mean. A z-score of –2 is two standard deviations *below* the mean, or equivalent to an IQ score of 70. Equation 6.2 defines a z-score:

$$z_i = \frac{X_i - \mu}{\sigma} = \frac{x_i}{\sigma} \qquad (6.2)$$

[1]A point of inflection is the precise point at which a smooth curve changes from concave (down-sloping) to convex (up-sloping). It is found using calculus; it is the point at which the second derivative of the equation equals zero.

[2]The mathematician would say that "the curve approaches the X-axis *asymptotically*" or "the X-axis is the asymptote of the curve."

In other words, a z-score tells us how many standard deviations the given score is above or below the mean. It is very informative when scores are expressed in terms of standard deviations from the mean μ, that is, z-scores. For almost any application of the normal curve, one wants to know how many standard deviations a score lies above or below the mean. Knowing this, questions about the area between points or scores, X_1 and X_2 (or heights of the curve above any point), can be answered by reference to the standard normal curve (see Table A in the Appendix). The *shape* of a curve does not change when a *constant* is added or subtracted from each score or when each score is multiplied or divided by a *constant*. Thus, when μ is subtracted from each score and these differences are divided by σ (see Equation 6.2), the shape of the distribution is not changed.

Any set of scores with mean μ and standard deviation σ can be transformed into a different set of scores with a mean of 0, and a standard deviation of 1; the transformed score, z, describes how many standard deviations the score falls above or below the mean of the distribution. These points are summarized in the following statement: If X is normally distributed with mean μ and standard deviation σ, then $z_i = (X_i - \mu)/\sigma$ has a normal distribution with a mean of 0, and standard deviation of 1. By making use of the definition of a z-score in Equation 6.2, Equation 6.1 can be simplified (Heffernan, 1988):

$$u = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{z^2}{2}\right)} \qquad \textbf{(6.3)}$$

The normal curve in Equation 6.3 and Figure 6.3 is a special one because it has been chosen as a standard. It is known variously as the *unit normal curve* or the standard (or standardized) normal distribution.

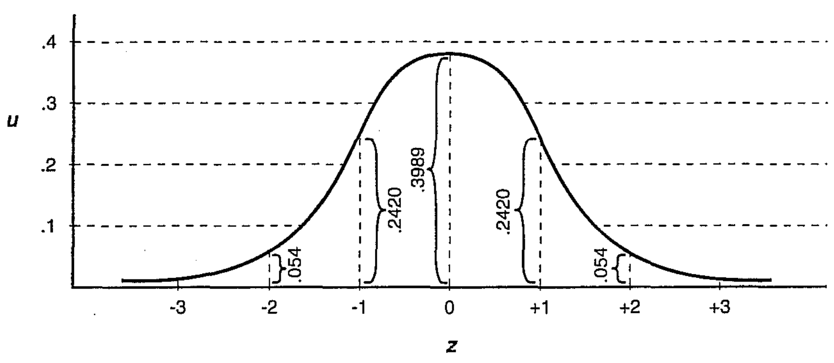


FIGURE 6.3 The Standard Normal Distribution, $\mu = 0$ and $\sigma = 1$.

6.4 ORDINATES OF THE NORMAL DISTRIBUTION

In rare circumstances, it is necessary to find the ordinate u, the height of the curve, at a given value of z. Solving Equation 6.3 for u when z is given is far too inconvenient (unless one is proficient with a scientific hand calculator that has a variable exponent function). Table A in the Appendix gives the ordinate u for values of z in the standard normal curve. The highest point on the curve is above the point $z = 0$ (see Figure 6.3 and Table A); when $z = 0$ is inserted in Equation 6.3, the height (ordinate) u is .3989. Notice in Figure 6.3 that the ordinate u equals .2420 at $z \pm 1.0$, and that $u = .0540$ at $z \pm 2.0$. For practice, locate these values in Table A.

6.5 AREAS UNDER THE NORMAL CURVE

In many applications of statistics, it is necessary to know the area of the normal curve (i.e., the proportion of the distribution) that falls below a particular value of z. To find the proportion of scores that falls below any particular score in a normal distribution, the score is converted to a z-score. The proportion is then read from the normal curve table, Table A. For example, in Figure 6.2, what proportion of the men were shorter than 70 inches? Stated differently, what is the percentile rank of 70 inches?[3] Assuming a normal distribution with a mean of 67.02" and a standard deviation of 2.56", $X_1 = 70''$ expressed as a z-score, using Equation 6.2, is:

$$z_1 = \frac{70 - 67.02}{2.56} = \frac{2.98}{2.56} = 1.16$$

One then finds $z = 1.16$ in Table A and in the adjacent column reads p, the proportion of the curve that falls below $z_1 = 1.16$. For $z_1 = 1.16$, $p_1 = .8770$; thus, 87.7% of the area (or observations) in a normal distribution falls below a z-score of 1.16 (see Figure 6.4). Or, stated differently, only 1 –.8770 = .1230 (12.3%) of the men were taller than 70".

What proportion of the heights in Figure 6.2 falls between 66" and 70"? If the proportion of heights below $X_2 = 66''$ is subtracted from the proportion below 70", the difference is the proportion between 66" and 70". A height of 66" corresponds to a z-score of –.40 [$z_2 = (66 - 67.02)/2.56 = -.398$, which rounds to –.40]. From Table A, the area below $z_2 = -.40$ is found to be .3446 (see Figure 6.4). Of the .8770 of the heights below 70", .3446 are below 66"; therefore, .8770 – .3446 = .5324 (or 53.24%) of the cases fall between 66" and 70".[4]

The steps just illustrated can be summarized as follows: The area between X_1 and X_2

[3] The heights in Figure 6.2 are not representative of the United States population, for which $\mu = 69.7''$ and $\sigma = 2.6''$ in 1976.

[4] Chebyshev's inequality proves that the proportion of the area of any distribution that is beyond the points $\pm z$ is less than $1/z^2$. Thus there is never more than $1/2^2 = .25$ of the distribution that is more than two standard deviations ($z = 2$) from the mean. For symmetric unimodal distributions, the maximum area beyond $\pm z$ is $4/(9z^2)$ (Dixon and Massey, 1983). Thus, there can be no more than $4/[9(2)^2] = 1/9$ or 11.1% of the area falling in the tails beyond the points $z = -2$ and $z = +2$ in symmetric unimodal distributions.

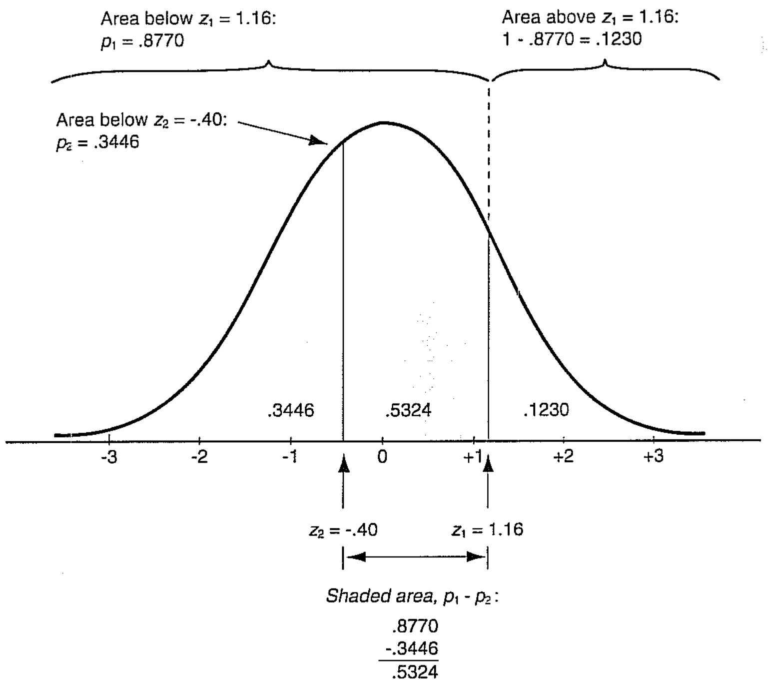


FIGURE 6.4 The Proportion of the Area under the Normal Curve Falling between Two Values of z.

in the normal distribution with mean μ and standard deviation σ is the area between $z_1 = (X_1 - \mu)/\sigma$ and $z_2 = (X_2 - \mu)/\sigma$ in the standard normal distribution.

6.6 OTHER STANDARD SCORES

It is easier to interpret observations when they are expressed as standard scores rather than as raw scores. With *standard scores*, the mean and standard deviation always have the *same fixed values*. To know that an IQ score on the Wechsler Intelligence Scale is 120 means little unless one knows also that $\mu = 100$; in addition, to know that $\sigma = 15$ enables the score of 120 to be interpreted much more meaningfully.

The z-score scale ($\mu = 0$, $\sigma = 1$) is the most widely used standard score scale in statistics, but any observation expressed in standard deviation units from the mean is a *standard score*. Most standardized tests of intelligence, achievement, interest, and personality report performance in standard scores. Such measures rarely use z-scores because other standard-score scales that do not involve negative numbers or decimals are preferred.

A z-score can be converted to any other standard score (S) using the general formula in Eq. 6.4:

$$S_i = \mu_S + (\sigma_S)z_i \quad \textbf{(6.4)}$$

where S_i is the new standard score equivalent to z_i;
μ_S is the mean of the new standard-score scale;
σ_S is the standard deviation of the new standard-score scale; and
z_i is the z-score for the ith observation.

6.7 *T*-SCORES

T-scores[5] are widely used to report performance on standardized tests and inventories. *T*-scores are standard scores with a mean (μ_T) of 50 and a standard deviation (σ_T) of 10. To convert z-scores to *T*-scores, Equation 6.4 becomes

$$T_i = 50 + 10z_i \quad \textbf{(6.5)}$$

An example will illustrate certain advantages of standard scores. Suppose a ten-year-old boy is 46" tall and weighs 76 pounds. Are his height and weight commensurate? Who knows without norms? However, expressed as *T*-scores (30 and 70, respectively) his weight problem becomes readily apparent—he is at the 2nd percentile in height, but the 98th percentile in weight!

If a student in grade 5.1 (first month of grade 5) obtained an IQ score of 130 and grade-equivalent scores of 6.4 and 6.1 on the standardized reading and arithmetic tests, respectively, how does her achievement compare with her measured scholastic aptitude (IQ)? The corresponding *T*-scores of 70, 60, and 60 show that the student's relative superiority above the mean on the intelligence test was twice as great as her degree of exceeding the mean on the reading and arithmetic tests.[6]

Figure 6.5 shows the relation of z-scores, *T*-scores, and several other standard-score scales. Observe that converting raw scores to standard scores does not alter the shape of the distribution or change the percentile ranks of any observation. Standard scores have the advantage of having a common mean and standard deviation that facilitates interpretation.

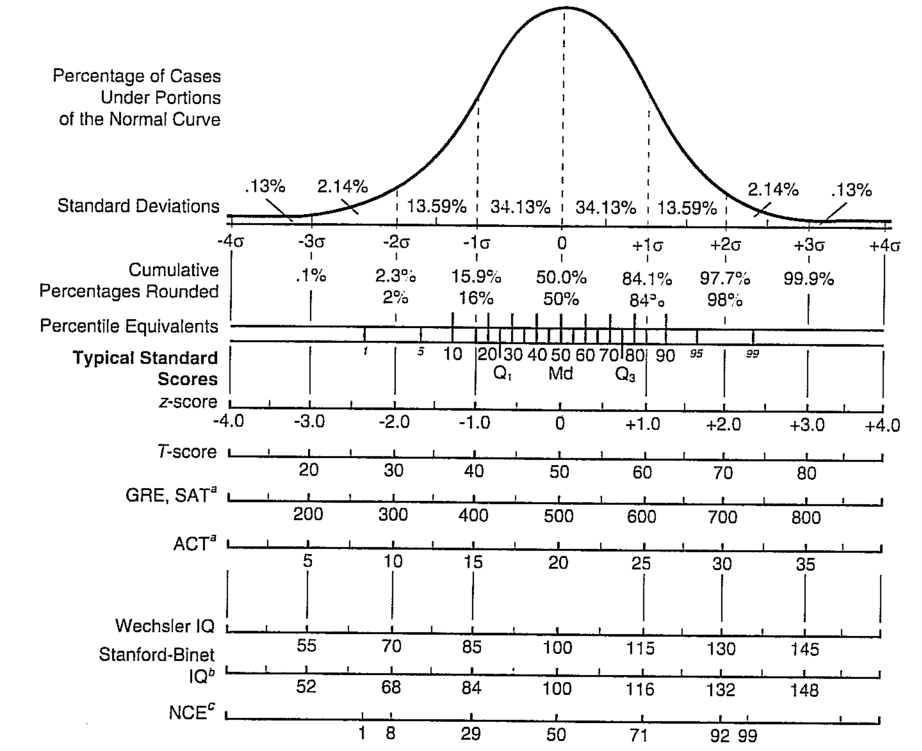
Notice that the frequently mentioned Wechsler IQ scale is a standard-score scale with $\mu = 100$ and $\sigma = 15$. The scale employed by the historic Stanford-Binet Intelligence Scale (form L-M) differs little ($\mu = 100$, $\sigma = 16$).[7] An IQ score of 145 on the Wechsler has the same z-score and percentile rank as a Stanford-Binet score of 148.

---

[5]The *T*-scale (named in honor of the early educational psychologist, Edward Lee Thorndike) was originally proposed as a *normalized* standard score (Section 6.11); in most current applications *T*-scores are not normalized, but are simply a linear transformation (Section 7.9) of raw scores.

[6]If more explanation and practice using the normal distribution and standard scores are desired, you may find the programmed instruction in Chapters 2 and 3 of Hopkins, Stanley, and Hopkins (1990) helpful.

[7]Prior to the 1960 revision of the Stanford-Binet, performance was expressed as a ratio IQ: IQ = 100(Mental Age/Chronological Age). The standard deviation of IQ scores fluctuated from one age to another, consequently the associated percentile rank for a given IQ score varied considerably from age to age. Now, virtually all cognitive aptitude tests use standard scores.



Distribution of scores of many standardized educational and psychological tests approximate the form of the normal curve shown at the top of this chart. Below it are shown other standard scores that are used by certain tests.

The zero (0) at the center of the baseline shows the location of the mean (average) raw score on a test, and the symbol σ (sigma) marks off the scale of raw scores in standard deviation units.

Most systems are based on the standard deviation unit. Among these standard score scales, the z-score and the T-score are general systems that have been applied to a variety of tests. The others are special variations used with College Entrance Examination Board tests, the Graduate Records Examination, and other intelligence and ability scales.

Tables of norms, whether in percentile or standard score form, have meaning only with reference to a specified test applied to a specified reference population. The chart does not permit one to conclude, for instance, that a percentile rank of 84 on one test necessarily is equivalent to a z-score of +1.0 on another; this is true only when each test yields essentially a normal distribution of scores and when both scales are based on identical or very similar groups of people.

[a]Score points (norms pertain to university students and not the general population). (GRE = Graduate Record Examination, SAT = Scholastic Aptitude Test of the College Entrance Examination Board, ACT = American College Testing Assessment.) Certain of these tests are not rescaled to better allow comparisons over time. Consequently, current means are lower than means given above.

[b]Standard-score IQ's with $\sigma = 16$ are also used on certain other intelligence tests.

[c]The NCE ("Normal Curve Equivalent") scale is an ill-conceived normalized scale used in the evaluation of certain federally funded educational programs. The NCE scale has $\mu = 50$ and $\sigma = 21$; and the NCE unit is 1/98 of the distance between the 1st and 99th percentiles, expressed in z-score units. The NCE scale invites the confusion of NCE standard scores with percentiles.

FIGURE 6.5 Illustrations of Various Standard Score Scales. (Adapted from Test Service Bulletin No. 48, The Psychological Corporation, New York, by permission of The Psychological Corporation.)

One may wonder why statisticians bothered to invent standard scores; why not just provide percentiles, which are easier to interpret? For all the clarity and simplicity of percentile scores, they are unsatisfactory for many statistical operations such as averaging and correlation (Chapter 7). The difference between the heights of two women at the 50th and 55th percentiles is much smaller than the height difference between two women at the 90th and 95th percentiles. Compare the z-scores in the normal distribution at the 50th and 55th versus the 90th and 95th percentiles: P_{50} corresponds to a z of 0 and P_{55} to a z of .126, a difference of $.126\sigma$; whereas, P_{90} corresponds to a z of 1.282 and P_{95} to a z of 1.645, a difference $.363\sigma$; and the difference between P_{90} and P_{95} is almost three (2.88) times greater than the difference between P_{50} and P_{55}! In terms of the heights in Figure 6.2, the difference between P_{50} and P_{55} is only .33 inches, whereas the difference between P_{90} and P_{95} is .93 inches. Stated differently, in IQ units, P_{50} and P_{55} differ by less than two (1.89) IQ points, whereas P_{90} and P_{95} differ by more than five (5.45) points. Standard scores avoid this distortion and lend themselves readily to meaningful summary statistical calculations.

6.8 AREAS UNDER THE NORMAL CURVE IN SAMPLES

This chapter has assumed that the population mean (μ) and standard deviation (σ) of a normal distribution are known. If the mean and standard deviation are estimated from a sample (i.e., $\bar{X}$ and s are used since μ and σ are not known) and inferences are made to the population, the proportion found in Table A is not precise, but only an approximation; the accuracy of the approximation is determined by how accurately $\bar{X}$ and s estimate μ and σ. When the random sample contains 100 or more scores, the z_i-value for an observation (X_i) using $\bar{X}$ and s will differ from the true z-value (i.e., the z-value using μ and σ) by .1 or less in most situations.[8] This degree of precision is adequate for most purposes. One should be wary of using Table A for inferential purposes if $\bar{X}$ and s are based on very small samples and when the frequency distribution is not normal.

6.9 SKEWNESS

A complete description of a distribution should include not only its central tendency and variability, but also the degree of asymmetry or skewness. The nature and extent of skewness is visually apparent from well-constructed frequency polygons and histograms, but these are rarely available in published research. Besides, mere observation of a distribution is imprecise and cannot be communicated accurately in words or numbers. There are two common measures of skewness. Recall (Section 4.15) that skewness influences the mean, median, and mode in a predictable way. In positively skewed distributions, the mean will have the largest value, and the mode the lowest; the relationship is reversed with negatively skewed distributions.

Figure 6.6 was constructed to illustrate various degrees of skewness. All the curves were converted to standard scores so they all have the same mean and standard deviation, but differ in skewness. Of course, the differences among the mean, median, and mode in-

[8]Technically only $(X_i - \mu)/\sigma$ is a z-ratio; $(X_i - \bar{X})/s$ is termed a t-ratio. If n is large, there will be little difference between the z and the t associated with an observation. The t-distribution is widely used in statistics and will be used extensively beginning in Chapter 11.

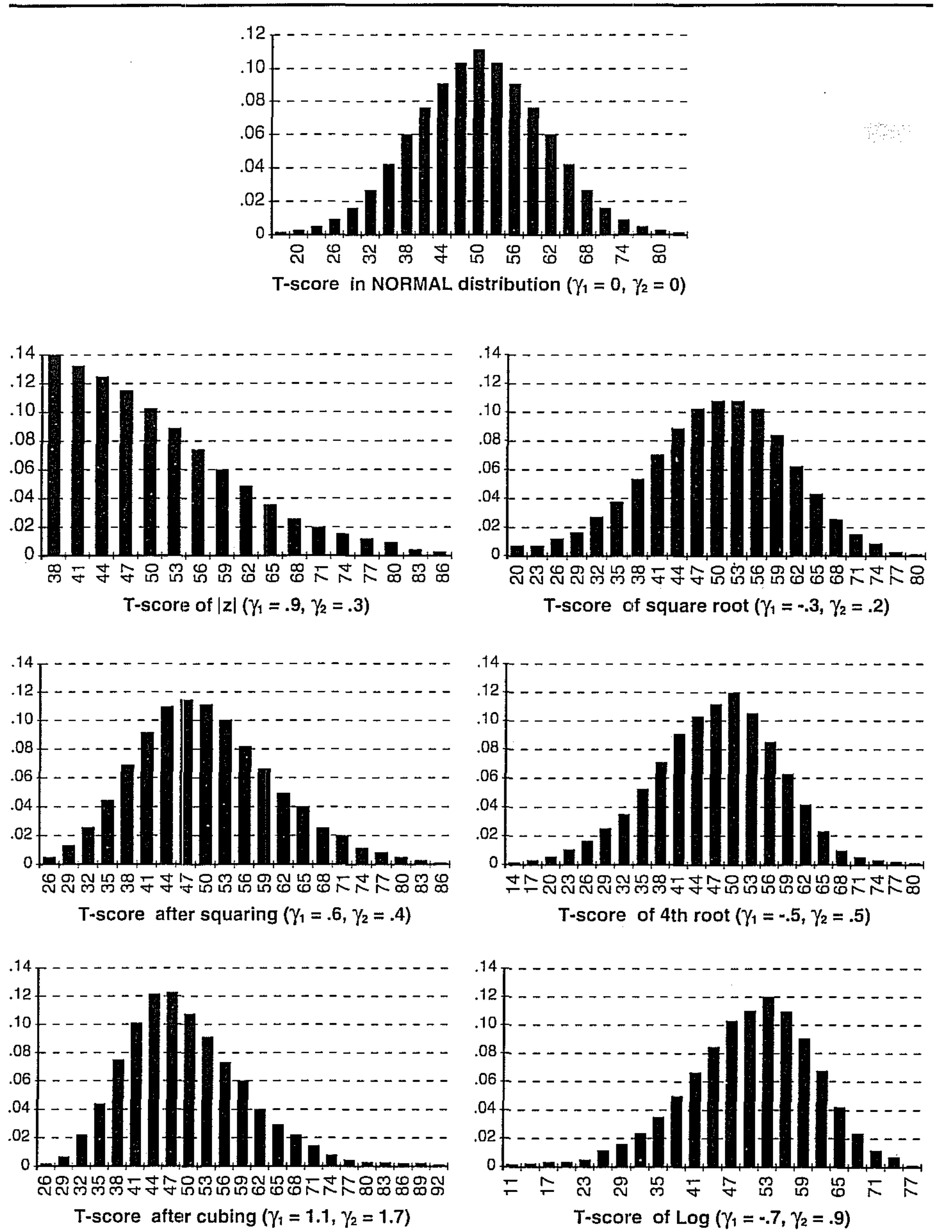


FIGURE 6.6 Distributions Illustrating Various Degrees of Positive and Negative Skewness.

crease as the magnitude of the skewness increases. Indeed, Karl Pearson suggested Equation 6.6 as a useful and an easily interpreted measure of skewness in the population:

$$\Omega = \frac{\mu - \text{Mode}}{\sigma} \tag{6.6}$$

Notice that the skewness index, Ω, describes the distance from the mean to the mode parameters in standard deviation units; it is the z-score of the mode with a change in sign. If $\Omega = .5$, the mean is $.5\sigma$ above the mode.

When used inferentially, the formula is usually revised so that the median can replace the mode because the sample median has much less sampling error than the mode of a sample. Recall that the distance between the mean and mode is approximately three times the distance between the mean and the median for regular unimodal distributions (see Equation 4.10, Section 4.15). Equation 6.7 provides a useful alternative for estimating Ω from the sample mean and median:

$$sk = \frac{3(\bar{X} - Md)}{s} \tag{6.7}$$

Pearson proposed a second measure of skewness, γ_1, which is preferred for inferential use.[9] If each of the N scores in a frequency distribution is transformed to a z-score and then cubed $[(z_i)(z_i)(z_i) = z_i^3]$, the mean of the z_i^3 scores is γ_1:

$$\gamma_1 = \frac{\Sigma_i z_i^3}{N} \tag{6.8}$$

Note that γ_1 is a parameter; when s is used rather than σ in the computation of the z-scores, γ_1 is estimated by the statistic $\hat{\gamma}_1$.[10] The various curves in Figure 6.6 illustrate several degrees of positive and negative skewness. The curves were obtained by performing mathematical transformations (Section 6.11) on the normally distributed T-scores in the top curve. The six curves with various degrees of skewness provide some frame of reference for interpreting the γ_1 skewness index:[11] the curves on the left illustrate various degrees of positive skewness ($\gamma_1 > 0$) and the curves on the right are negatively skewed. Ω is more informative than γ_1 for visualizing the degree of skewness and reconstructing the frequency distribution; γ_1 is superior for inferential purposes (Snedecor and Cochran, 1980, pp. 78–79). Note that Ω can be estimated for the curves in Figure 6.6.[12]

---

[9] This is the measure used by BMD, SPSS, SAS and virtually all other computer programs.

[10] The circumflex, "^" above a Greek letter denotes not the parameter, but an estimate of the parameter, for example, $s = \hat{\sigma}$ (say, "sigma hat").

[11] The curves were created by applying a mathematical transformation to scores from a normal distribution. For example, if scores from a normal distribution are squared, the γ_1 changes from 0 to .6.

[12] Since all distributions are expressed using T-scores ($\mu = 50$, $\sigma = 10$), and the mode is the value with the greatest frequency, $\Omega = (\text{Mode} - 50)/10$.

Although it is not yet common practice, researchers should routinely report the degree of skewness to describe the shape of the distributions of interest (Hopkins & Weeks, 1990). This information is also useful for the interpretation of certain measures (e.g., correlation coefficients) and statistical tests (e.g., homogeneity of variance, Section 12.10) that can be affected by skewness.

6.10 KURTOSIS

Up to this point, three properties or features of groups of scores have been described: central tendency, variability, and symmetry. A fourth property, *kurtosis*, completes the set of characteristics of distributions of scores that are of interest in analyzing data. One may wish to know whether there are more or fewer extreme scores[13] than expected in a normal distribution. The customary measure of kurtosis, γ_2, is the mean of the distribution of z_i^4 scores (i.e., z's raised to the fourth power) minus the constant 3 (which is the mean z_i^4 value for the normal curve).[14]

$$\gamma_2 = \frac{\Sigma_i z_i^4}{N} - 3 \qquad (6.9)$$

Figure 6.7 gives three distributions (on the left) that have the same means and standard deviations (i.e., expressed in T-scores), but have negative kurtosis. They are termed *platykurtic* distributions. (The prefix "platy" means flat or broad.) These curves have fewer extreme scores than found in a normal distribution. Note that γ_2 for a symmetrical dichotomous distribution is −2 and −1.2 for a rectangular (uniform) distribution. When the kurtosis is based on a sample of n observations, the kurtosis index is a statistic and denoted by $\hat{\gamma}_2$.

The three curves on the right are from the t-distribution family with 5, 10, and 25 degrees of freedom; this distribution will be used extensively beginning in Chapter 11. The t-distributions have more extremely high or low scores than does the normal distribution, which gives γ_2 a positive value. In the normal distribution, the value of γ_2 is 0; it is said to be *mesokurtic* ("meso" means intermediate). Distributions in which the kurtosis index is positive are described as *leptokurtic* ("lepto" means slender or narrow). Highly skewed distributions tend to be leptokurtic because they have more scores that are far from the mean than does the normal distribution.

Ordinarily, there is far less interest in the kurtosis of a distribution as a descriptive statistic than in its central tendency, variability, and skewness. Kurtosis is important for evaluating the accuracy of certain statistical tests (e.g., see Section 13.8).

[13]Shiffler (1988) has shown that the largest possible z-score for any observation is limited by sample size n: $z_{max} = (n-1)/\sqrt{n}$. For example, if $n = 4$, $z_{max} = 1.5$; if $n = 100$, $z_{max} = 9.9$.

[14]In some sources −3 is absent from the formula for kurtosis; without the −3 in Equation 6.9, the normal distribution would have a kurtosis index of 3.

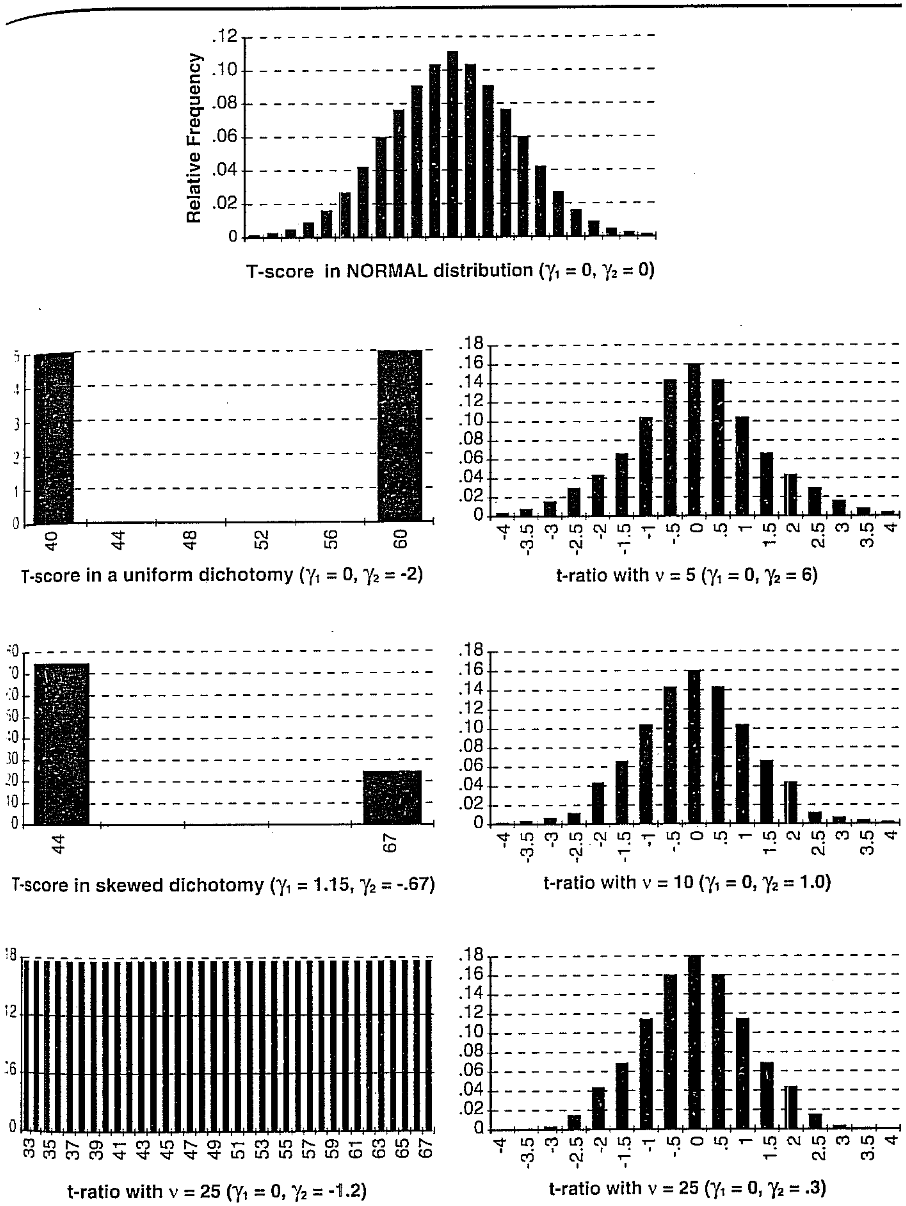


FIGURE 6.7 Symmetric Curves Differing in Kurtosis.

6.11 TRANSFORMATIONS[15]

Many statistical methods assume a normal distribution in the populations being studied. Although many of these methods work quite well even when the assumption is not satisfied, it is sometimes desirable to convert the original scores to another metric in which the distribution is more nearly normal. It often happens that a certain change of scale, such as using the square root, reciprocal, or logarithm of the observations, will result in less skewness than in the original observation. Transformations can sometimes convert non-linear relationships between two variables into a linear relationship (Section 8.28).

Figure 6.6 illustrates how the shape of a normal distribution is altered by six mathematical transformations. Notice that if the mathematical process is reversed, the normal distribution would be reproduced (except for the distribution of absolute values, $|z|$). For example, if we started with a distribution that has slight negative skewness like the *T*-scores of the square root transformation (immediately below and to the right of the normal distribution), and square its members, a normal distribution will result. In other words, negative skewness can be reduced by raising the scores to a power greater than 1. Likewise, the square root and log transformations will reduce the positive skewness in a distribution. If the square root transformation is applied to the distribution "*T*-score after squaring," the normal distribution at the top of Figure 6.6 would result. Thus, root and log transformations reduce positive skewness and power transformations reduce negative skewness.

Transformations are not an end in themselves; but certain other statistical procedures may require that a variable be normally distributed, or have a linear relationship with another variable, and that may require a transformation. The best transformation is often difficult to find, and success in finding a good transformation is frequently a matter of trial and error. Additional guidelines can be found in Kirk (1982, pp. 81–84); Winer, Brown, and Michels (1991); Lee (1975, pp. 288–291); Dixon and Massey (1983, Chapter 16); Snedecor and Cochran (1980, Chapter 15); and Tukey (1977, Chapter 3).

6.12 NORMALIZED SCORES

When it can be assumed that the variable being measured is normally distributed, but the observations are not normally distributed because of faulty measurement, the observed distribution is sometimes *normalized*, that is, the distribution is forced to approximate the normal distribution as closely as possible. This transformation is monotonic (the rank order of X_i's is maintained), but is non-linear (the relative distances between scores are not maintained, see Section 7.12).

Usually normalized scores are expressed using the *T*-score scale ($\mu = 50$, $\sigma = 10$). Normalized *T*-scores are obtained by first converting the original scores to percentiles, then converting each percentile to the *T*-score corresponding to that percentile in a normal distribution. In other words, this sequence is followed:

$$X_i \rightarrow P_i \rightarrow z_i \rightarrow T_i \qquad \textbf{(6.10)}$$

For example, suppose a raw score of 37 is at the 10th percentile in the original distribution. From Table A in the Appendix, one can see that the z-score that is associated with the 10th percentile in the normal distribution is −1.282. The *T*-score that corresponds to $z = -1.282$ (Equation 6.5) is $T = 50 + 10(-1.282) = 37.18$ or 37 (*T*-scores are usually rounded to the nearest whole number.) Unless the observed distribution deviates substantially from the normal distribution in skewness or kurtosis, normalized *T*-scores will differ little from non-normalized *T*-scores.[16]

[15]More explicitly, this should read non-linear transformations (i.e., transformations that change the shape of a distribution). The z-scale and *T*-scale (Equations 6.2 and 6.5) are linear transformations of X (Section 7.9).

6.13 CHAPTER SUMMARY

The measurements of many variables in the social and behavioral sciences have distributions that are closely approximated by the normal distribution. In addition, many distributions used in inferential statistics are normally distributed.

The normal distribution is symmetrical, unimodal, and bell-shaped. There is a known proportion of the curve below any z-score in a normal distribution. These proportions can be found from Table A by expressing the observation as a z-score, the number of standard deviations that the observation falls above or below the mean [$z_i = (X_i - \mu)/\sigma$].

Besides z-scores, there are other widely used standard-score scales. The most popular is the *T*-scale that sets $\mu = 50$ and $\sigma = 10$.

Skewness and kurtosis indices describe two ways in which a distribution differs from the normal distribution. Non-normal distributions can often be made to be more nearly normal by using certain mathematical transformations on the scores.

6.14 CASE STUDY

In previous chapters, we have studied the variables in the case study with respect to central tendency and variability. We did observe skewness in certain of the variables. In this chapter, we quantify the degree of skewness, γ_1, and kurtosis, γ_2.

The skewness index, $\hat{\gamma}_1$, is a measure of asymmetry; ID # has a skewness index of .00 because the distribution is perfectly rectangular and symmetrical. The distribution of height is only slightly asymmetrical. Notice that all the other distributions are skewed positively; the degree of the skewness is substantial for the SBP, weight, and coronary variables. Note that a dichotomous variable like coronary will be skewed to the extent that the frequencies in the two categories are unequal; they are very unequal here, hence, the large value for $\hat{\gamma}_1$.

Table 6.1 gives $\hat{\gamma}_1$ and $\hat{\gamma}_2$ indices when the scores on each variable are converted to *T*-scores. Although for all variables the mean *T*-score is 50 and the standard deviation is 10, the shape of the distribution has not changed; thus, the skewness and kurtosis indices are

[16]On rare occasions, one may need to know the mean and standard deviation within a segment of the normal distribution. The mean z-score for a given section of the normal distribution between the points z_a and z_b (where $z_b > z_a$) is the difference in the corresponding ordinates u_a and u_b at z_a and z_b divided by p, the proportion of cases falling within the segment. Thus, $\mu_z = (u_a - u_b)/p$. To find the mean of scores falling between the median and Q_3 expressed as a z-score, do the following: From Table A, read that the ordinates corresponding to z_a and z_b are $u_a = .3989$ and $u_b = .3178$. Hence, $\mu_z = (.3989 - .3178)/.25 = .3244$. The same procedure can be used to find the mean of the scores in the tail of the normal distribution (Kelley, 1939). The mean z-score in the top quarter of normally distributed scores is $\mu_z = (.3178 - 0)/.25 = 1.27$: conversely, the mean of scores below Q_1 is $\mu_z = (0 - .3178)/.25 = -1.27$.

The variance of z-scores within a section of the normal distribution is $\sigma_z^2 = [p(z_a u_a - z_b u_b) - (u_a - u_b)^2]/p^2$. The variance of z-scores between the median and Q_3 is $\sigma_z^2 = \{.25[(0)(.3989) - (.674)(.3178)] - (.3989 - .3178)^2\}/(.25)^2$ $= .0380$, and $\sigma_z = .195$. For the top or bottom quarter of the normal distribution, $\sigma_z = .490$.

TABLE 6.1 Descriptive Information for Chapman Case Study Variables ($n = 200$)

			T-scores		$\sqrt{T}$		$LOG_{10}T$	
Variable	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
Age	.28	–.80	.28	–.80	.07	–.96	–.15	–.97
Systolic B.P. (SBP)	1.51	3.50	1.51	3.50	1.23	2.47	.96	1.70
Diastolic B.P. (DBP)	.45	.76	.45	.76	.22	.58	–.01	.55
Cholesterol Level	.41	.42	.41	.42	.05	.09	–.31	.16
Height (in.)	.10	–.35	.10	–.35	.05	–.34	.01	–.33
Weight (lbs.)	.84	1.54	.84	1.54	.55	.91	.28	.52
Coronary? (0=N, 1=Y)	2.29	3.25	2.29	3.25	2.29	3.25	NA	NA
ID #	.00	–1.2	.00	–1.2	–.56	–.63	–.56	–.625

identical to those for the raw score distributions. This type of conversion from one scale to a different scale that does not change the configuration of the frequency distribution is termed a *linear transformation*.[17]

The positive skewness in SBP could be reduced if certain *non-linear* transformations are employed, for example, if the actual blood pressure values are replaced by their square roots or logs.[18] We will illustrate the effect of two transformations—square root and logarithms.[19] In Table 6.1, skewness and kurtosis indices are given for the distributions of the square root (SQRT) transformation, $\sqrt{X}$. Note that the SQRT transformation reduces positive skewness, for example, for age the skewness was reduced from .28 to .07. The SQRT transformation will introduce negative skewness for symmetrical distributions like ID #. Since skewed distributions are leptokurtic ($\hat{\gamma}_2 > 0$), the SQRT transformation also reduces the leptokurtosis (i.e., $\hat{\gamma}_2$ is decreased) in the distributions. The transformations have no effect on a dichotomous distribution, but they will change its mean and variance.

The effect of the square root transformation on the shape of the distribution of cholesterol level is illustrated in Figure 6.8. The printout is from the computer program SPSS FREQUENCIES, which will superimpose a normal distribution backdrop if requested. (This program also produces a variety of statistical information including measures of central tendency, variability, skewness, and kurtosis.) Although the differences between the two distributions are not striking, a study of both distributions will reveal that the outlier ($X = 520$) that contributes greatly to the positive skewness and kurtosis, becomes less extreme after the square root transformation.[20]

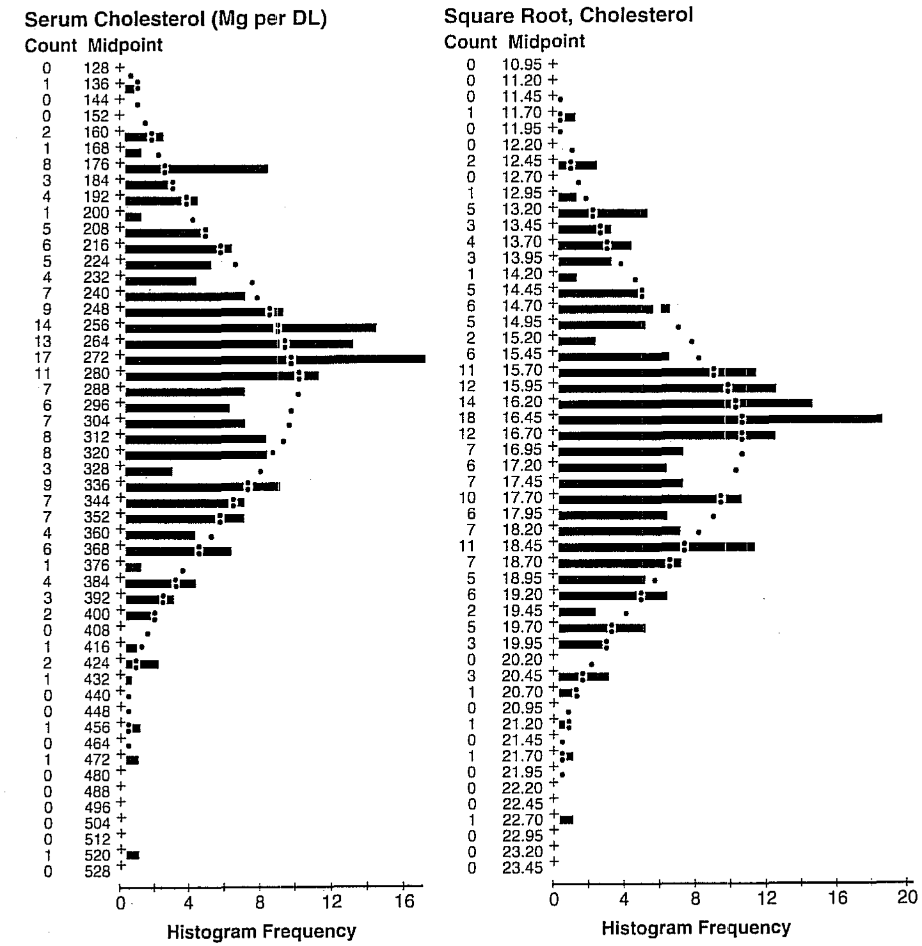
Table 6.1 gives $\hat{\gamma}_1$ and $\hat{\gamma}_2$ indices after a second non-linear transformation has been applied to the raw scores, that is, when each X_i is replaced by $\log_{10}X_i$. This LOG transformation reduces positive skewness, or increases negative skewness, even more than the SQRT

[17]The topic of linear transformations is treated more extensively in Chapter 7, Section 7.12.

[18]Note in Figure 6.6 how the shape of the distribution changes depending on the particular transformation that is applied.

[19]It is unlikely that transformations would be used in the actual study, especially based on what we know about the variables at this point; they are included primarily for illustrative purposes.

[20]Confirm that in the raw score distribution, the outlier 520 in the distribution with a mean of 285.11 and a standard deviation of 65.04 is 3.61 (= z) standard deviations above the mean. After the square root transformation, the mean and standard deviation were 16.775 and 1.927, respectively, and $\sqrt{520} = 22.804$, which is 3.13 (= z) standard deviations from the mean.



Mean	285.110	Std Err	4.599	Median	276.000
Mode	264.000	Std Dev	65.041	Variance	4230.309
Kurtosis	.420	S E Kurt	.342	Skewness	.407
S E Skew	.172	Range	385.000	Minimum	135.000
Maximum	520.000	Sum	57022.000		

Valid Cases 200 Missing Cases 0

Mean	16.775	Std Err	.136	Median	16.613
Mode	16.248	Std Dev	1.927	Variance	3.713
Kurtosis	.094	S E Kurt	.342	Skewness	.048
S E Skew	.172	Range	11.185	Minimum	11.619
Maximum	22.804	Sum	3355.092		

Valid Cases 200 Missing Cases 0

FIGURE 6.8 Frequency Distributions (with Superimposed Normal Distribution) for Cholesterol and Square Root Transformation of Cholesterol Variables.

transformation. Note that for age, the original value of $\hat{\gamma}_1$ was .28; $\hat{\gamma}_1$ decreased to .07 with the SQRT transformation, and $\hat{\gamma}_1$ became slightly negative (–.15) when the LOG transformation was applied. If our goal was to coerce the distributions into approximate normality as closely as possible, we would apply neither transformation to coronary nor ID #. We would choose the SQRT over the LOG transformation for age and cholesterol level, and the LOG transformation for all others. Based on our current analyses, we would not use transformations; other considerations pertaining to transformations will be considered in Chapter 7.

We could remove most of the skewness in all the variables that have an underlying continuum if we normalized (Section 6.12) the distributions, but normalizing is appropriate only when there is a problem with the measurement scale, which is not the case here.

6.15 SUGGESTED COMPUTER EXERCISE

Using speadsheet or statistical software, compute skewness indices for all the variables in the HSB case study data set including ID, but not career (CAR) and RACE. Note that the variable ID has a rectangular distribution, thus has no skewness. Note that the dichotomous variable School Type (SCTYP) is highly skewed, but SEX is only slightly skewed. Note that the use of *T*-scores does not necessarily indicate that a variable is normally distributed, for example, writing (WRTG). Use a transformation on WRTG to see if you can reduce its skewness (since WRTG is negatively skewed, create another variable that is the square or cube of the WRTG scores). See what happens to the skewness of WRTG if the square root or log transformations are used. Recompute the skewness to see how the skewness is affected by the transformations. Obtain histograms and examine the change in appearance before and after the transformations.

MASTERY TEST

Information on certain standardized intelligence and achievement tests is given. Answer questions 1 to 10 assuming the scores are normally distributed.

Iowa Test of Basic Skills Grade-Equivalent Scores

		Reading			Arithmetic
	Wechsler IQ	*Grade 3*	*Grade 5*	*Grade 8*	*Grade 5*
μ	100	3.0	5.0	8.0	5.0
σ	15	1.0	1.4	1.9	1.1

1. An IQ score above 115 is obtained by what percent of the population?
2. If a fifth-grade student obtains a percentile rank of 84 in reading, what is the grade-equivalent score?
3. What is the grade-equivalent score for the same relative performance as in question 2 (P_{84}) in arithmetic at grade 5?
4. Jack was reading at 6.1 when he entered grade 8. If his Wechsler IQ is equivalent to the same percentile rank, what is it?
5. If Jack's score in question 4 is valid, he reads better than about what percentage of children in his grade?

6. Upon entering grade 3, approximately what percent of third-grade children:
   (a) obtains a reading grade-equivalent score of 3.0 or better?
   (b) obtains a score of 4.0 or better?
   (c) obtains a score of 5.0 or better?
7. On the reading test, what percent of beginning third-grade students (3.0) score no higher than the average beginning second-grade students (2.0)?
8. At grade 5, is a grade-equivalent score of 6.0 relatively better (i.e., does it have a higher percentile equivalent) in arithmetic than in reading?
9. In reading, what percentages of third-grade students score below grade-equivalent scores of 2.0, 3.0, 4.0, and 5.0, respectively?
10. How much reading gain in grade-equivalent units is required during the five years between grades 3.0 and 8.0 to:
    (a) maintain a percentile equivalent of 50?
    (b) maintain a percentile rank of 84?
11. If $X_3 = 176$ with $\mu = 163$ and $\sigma = 26$, express X_3 as:
    (a) a z-score. (b) a *T*-score. (c) a percentile equivalent.
12. If IQ's were perfectly normally distributed, how many persons in the United States would have IQ's exceeding 175? (Assume $\mu = 100$, $\sigma = 15$, and $N = 250{,}000{,}000$.)
13. What percentage of IQ scores would fall between:
    (a) 90 and 110? (b) 80 and 120? (c) 75 and 125?
14. If men's heights are distributed normally, approximately how many men in 10,000 will be 6′6″ or taller? (Use $\mu = 69.7''$, $\sigma = 2.6''$)
15. Which of these is not a characteristic of a normal distribution?
    (a) symmetrical (b) unimodal (c) skewed (d) mesokurtic
16. Which of these reflects the poorest performance on a test? Assume a normal distribution.
    (a) P_{10} (b) $z = -1.5$ (c) $T = 30$
17. With a sample of 1,000 representative observations, which of these is probably least accurately characterized by the normal distribution?
    (a) scores on a musical aptitude test
    (b) number of baby teeth lost by age eight
    (c) size of reading vocabulary of twelve-year-old children
    (d) number of times attended a religious service in the past year
    (e) scores on an inventory measuring interest in politics
18. If raw scores are changed to z-scores, would the shape of the distribution be changed?
19. If z-scores are multiplied by 10, the standard deviation increases from ____ to ____.
20. What is the variance in a distribution expressed as (a) z-scores? (b) *T*-scores?
21. Small changes in z-scores near the mean (e.g., from 0 to .5) correspond to large or small changes in percentile equivalents. Large z-score changes near the extremes (e.g., 2.0 to 2.5) correspond to large or small changes in percentile equivalents.
22. If for a class of gifted children $\mu = 140$, $\sigma = 10$, and skewness $sk = .6$ (Equation 6.7), estimate the mode and median of the distribution of scores.
23. What is the skewness index, sk, for the distribution of cholesterol levels in the Chapman study? (Use Equation 6.7 and Table 5.2.)
24. The square root and log transformations will reduce ____ (positive or negative) skewness in a distribution. (See Figure 6.6)

PROBLEMS AND EXERCISES

1. Find the area under the normal curve which lies:
   (a) above $z = 1.00$ (e) between $z_1 = 0$ and $z_2 = 3.00$
   (b) below $z = 2.00$ (f) above $z = -.50$
   (c) above $z = 1.64$ (g) between $z_1 = -1.50$ and $z_2 = 1.50$
   (d) below $z = -1.96$
2. Find the ordinates of these z-scores in the normal distribution:
   (a) $z = 2.25$ (b) $z = -.15$
3. Find the z-scores that are exceeded by the following proportions of the area under the normal distribution:
   (a) .50 (b) .16 (c) .84 (d) .05 (e) .005 (f) .995 (g) .10
4. If in the general population of children, Stanford-Binet IQ's have a nearly normal distribution with mean 100 and standard deviation 16 (see Figure 6.5), find the percentile equivalent of each of the following IQ's:
   (a) 100 (b) 120 (c) 75 (d) 95 (e) 140
5. Suppose Mary obtained the following percentiles on five subtests on the *McCarthy Scales of Children's Abilities*:

Subtest	*Percentile*
Verbal	98
Perceptual	99.9
Quantitative	50
Memory	84
Motor	16

Use Figure 6.5 to answer the exercises below.

(a) If Mary's Motor performance improved by 1σ, the percentile equivalent would increase from 16 to ____, or ____ percentile units.
(b) If the Verbal score improved by 1σ, the percentile equivalent on the Verbal tests would increase from 98 to ____, or ____ percentile units.
(c) In standard deviation units, is the size of the difference between Mary's performance on the Verbal and Perceptual tests the same as the difference between her Motor and Quantitative scores?
(d) If expressed in T-scores, would the change from P_{16} to P_{50} in exercise (a) be equal to the change from P_{98} to $P_{99.9}$ in exercise (b)?

6. The manual for the *Metropolitan Achievement Tests* (MAT) contains no report of standard deviations for the grade equivalent (GE) scales but does give percentile ranks as indicated:

Percentile for Fall of Grade 5	*Grade Equivalents*	
	Reading	*Math*
84	9.0	6.8
50	5.0	5.0
16	3.0	3.4

(a) Estimate the GE standard deviation for the reading and math tests. $[\sigma \approx (P_{84} - P_{16})/2]$

(b) Which distribution is more severely skewed?
(c) Would the mean GE be greater on the reading or the math test? Explain.
(d) Using the estimated standard deviation on the math test, approximately what percent of beginning fifth-grade students obtain GE scores above 6.0 on the MAT? Assuming normal distributions, compare this figure with the corresponding figure for the ITBS Arithmetic Test (see data preceding question 1 on the Mastery Test).

7. "Grading on the normal curve" was popular in some circles a few decades ago. The most common method used the following conversion. Using this system, what percent of A's, B's, C's, D's, and F's are expected with a normal distribution of scores?

Grade	*z-score*
A	above 1.5
B	.5 to 1.5
C	–.5 to .5
D	–1.5 to –.5
F	below –1.5

8. If many naive examinees guess randomly on each of the 100 items on a true-false test, the mean would be expected to be $50 = \mu$, with $\sigma = 5$. What percent of examinees would be expected to earn scores of 65 or more?
9. Each of eleven students in a class was asked to respond to a sociometric measure in which they identified the three persons who had showed the most leadership ability. The scores (number of nominations) for each student are given below. View the group as a population, not as a sample.

X_i	x_i	z_i	z_i^2	z_i^3	z_i^4
9	6	2.4	5.76	13.824	33.1776
5	2	.8	.64	.512	.4096
5	2	—	—	—	—
4	1	.4	—	—	—
3	0	.0	.00	.000	.0000
2	–1	–.4	.16	—	—
2	–1	–.4	.16	–.064	.0256
1	–2	–.8	.64	–.512	.4096
1	–2	–.8	.64	–.512	.4096
1	–2	–.8	.64	–.512	.4096
0	–3	___	1.44	–1.728	2.0736
$\Sigma X_i =$ ___	$\Sigma x_i =$ ___	$\Sigma z_i =$ ___	$\Sigma z_i^2 =$ ___	$\Sigma z_i^3 =$ ___	$\Sigma z_i^4 =$ ___

(a) What are the mean, median, and mode of the distribution?
(b) Supply the missing z-score ($\sigma = 2.5$).
(c) Supply the missing values in the z_i^2, z_i^3, and z_i^4 columns, and find the sums of each column.
(d) Compute the skewness (γ_1) and kurtosis (γ_2) indices of the distribution.

10. Suppose a student can qualify for \$100 in additional state aid designated for special remedial reading by scoring 2.0 or more grade equivalents below his current grade level status. For a typical, representative school district with approximately 4,000 students per grade level, how much more state aid would the district receive for its fifth graders given that the Metropolitan (MAT) was given rather than the Iowa (ITBS)? The standard deviations are 3.0 and 1.4 for the

MAT and ITBS, respectively; use 5.0 for both means. Assume normality and round z-scores to the second decimal place.

11. Knowing $\sigma = 15$ on the Wechsler IQ scale, estimate Q.
12. Some academic departments use the *Miller Analogies Test* for selecting graduate students. Although only raw scores are reported, for students applying for graduate study the mean and standard deviation are approximately 48 and 17, respectively. If applicants at a certain prestigious university are expected to be in the upper 10%, what is the minimum raw score expected on the Miller?
13. Prove that $\Sigma_i z_i^2 = N$.

ANSWERS TO MASTERY TEST

1. 16%
2. 6.4
3. 6.1
4. 85
5. 16%
6. (a) 50%, (b) 16%, (c) 2%
7. 16%
8. yes
9. 16%, 50%, 84%, and 98%
10. (a) 5.0 (b) 5.9
11. (a) $z = (176 - 163)/26 = .50$
    (b) $T = 50 + 10(.50) = 55$
    (c) P_{69}
12. $z = (175 - 100)/15 = 5.0$;
    $(.0000002867)(250{,}000{,}000) \approx 80$
13. (a) .7486 – .2514 =.4972 or about 50%
    (b) .9082 – .0918 = .8164 or 82%
    (c) .9525 – .0475 = .9050 or 91%
14. $z = (78 - 69.7)/2.6 = 3.19$;
    $(.00071)(10{,}000) = 7.1$ or about 7
15. (c)
16. (c)
17. (d)
18. no
19. 1.0 to 10
20. (a) $(1)^2 = 1$; (b) $(10)^2 = 100$
21. large, small
22. Mode ≈ 134, Md ≈ 138 (See Equation 4.10)
23. .42
24. positive

ANSWERS TO PROBLEMS AND EXERCISES

1. (a) .1587 (b) .9772 (c) .0505 (d) .025 (e).4987 (f) .6915 (g) .8664
2. (a) .0317 (b) .3945
3. (a) 0.00 (b) +1.00 (c) –1.00 (d) +1.645 (e) +2.58 (f) –2.58 (g) +1.28
4. (a) 50 (b) 89 (c) 6 (d) 38 (e) 99
5. (a) 50, 34 (b) 99.9, 1.9 (c) yes, 1σ in each instance (d) yes, T-score increase of 10 in each instance
6. (a) $\sigma_R \approx 3.0$, $\sigma_M \approx 1.7$ (b) reading (c) reading; greater positive skewness (d) MAT: $z = .59$, 28% above 6.0; ITBS: $z = .91$, 18% above 6.0
7. A: 7%, B: 24%, C: 38%, D: 24%, F: 7%
8. .0013 or .13%, or roughly one student in 1,000
9. (a) mean = 3.0, median = 2, mode = 1; (b) .8, –1.2 (c); (c) $\Sigma X = 33$, $\Sigma x = 0$, $\Sigma z = 0$, $\Sigma z^2 = 10.88$, $\Sigma z^3 = 11.52$, $\Sigma z^4 = 37.3760$; (d) $\gamma_1 = 11.52/11 = 1.05$, $\gamma_2 = 37.376/11 - 3 = .40$
10. MAT: $z = -.67$, (.2514)(4,000)(100) = \$100,560; ITBS: $z = -1.43$, (.0764)(4,000)(100)= \$30,560, or \$70,000 more using MAT
11. From Table A, $Q = .674\sigma$, $Q = 10$
12. 70
13. $\Sigma_i z_i^2 = \Sigma_i (X_i - \mu)^2/\sigma^2 = (1/\sigma^2)\Sigma_i (X_i - \mu)^2 = (1/\sigma^2)(N\sigma^2) = N$

7

CORRELATION: MEASURES OF RELATIONSHIP BETWEEN TWO VARIABLES

7.1 INTRODUCTION

Measures of correlation are used to describe the relationship between two variables. In addition, correlation is an important part of many other statistical techniques. In this chapter, we will present the meaning, use, and computation of common measures of relationship.

Behavioral research frequently assesses the degree of association between two variables. The variables may be on many different kinds of observational units: persons, classes, schools, sites, cities, or the like. For example: Is absenteeism related to socioeconomic status (SES) for high school students? Is class size related to achievement growth for first grade classes? Do competitive cultures have a greater incidence of peptic ulcers? Can GPA be better predicted from SES than from IQ? To answer questions such as these, measures of relationship (correlation coefficients) are needed.

Most persons have a general understanding of correlation. Two variables are correlated if high scores on one variable tend to "go together" with high scores on the second variable. Likewise, if low scores on variable X tend to be accompanied by low scores on variable Y, then the variables X and Y are correlated. The degree of correlation between variables can be described by such terms as strong, low, positive, or moderate, but these terms are not very precise. If a coefficient of correlation is computed between the two sets of scores, the relationship is described more precisely. A coefficient of correlation is a statistical summary of the degree and direction of relationship or association between two variables.

7.2 THE CONCEPT OF CORRELATION

There is a substantial, but by no means perfect, positive correlation between annual income and the taxes paid to IRS. Husbands and wives tend to be alike in age, amount of education, and many other ways. The sons of tall fathers tend to be taller than average, and the sons of