CHAPTER 7

Nature of Variables and Data

"USA Today *has come out with a new survey. Apparently, three out of four people make up 75% of the population.*"

David Letterman

OBJECTIVES

By reading and studying this chapter, you should acquire the competency to do the following:

- Explain the nature of a construct and explain why constructs are considered variables.
- Describe the difference between a construct and a property.
- Identify the level of precision (NOIR) for a range of different variables.
- Explain the limitation imposed by the precision category level of a variable and how its central tendency is determined.
- Read and interpret data presentations involving variance and standard deviations.
- Describe appropriate graphic presentations for all levels of scale precision.
- Find *z*-scores and describe how they apply to areas under the Gaussian normal curve.
- Describe commonly employed standard score transformations.

INTRODUCTION AND MAJOR THEMES

"Before the curse of statistics fell upon mankind, we lived a happy, innocent life, full of merriment and go and informed by fairly good judgment."

Hilaire Belloc

The best therapists are simultaneously aware of many **variables** when working with clients. Some of these variables are psychological constructs, and others are categorical information. All variables do not have the same degree of precision, and care must be exercised when summarizing and reporting data. The different levels of precision of variables imply that a limited range of graphics and arithmetic procedures can be used to describe and summarize them.[1]

CONSTRUCTS

The point was made in Chapter 1 that much of what we understand about the world is the product of what we were told or taught. Another major source of new learning is what we observe others experiencing. This latter source of knowledge about the world includes what the **social learning** theorist Albert Bandura described as **vicarious learning** (Bandura, 1976; Bandura & McDonald, 1963).

Professional counselors are highly trained observers and listeners who learn a great deal about clients through these channels. These observations are mentally clustered into **constructs** held about clients. In the mental health fields, practitioners observe, analyze, and assist clients in modifying aspects of some of these constructs. Constructs are observable clusters of characteristics that may be drawn from an individual's social circumstances, history, beliefs, or ideologies. They are not physical characteristics and lack any physical dimensions. These **psychological constructs** include dimensions such as anger, creativity, jealousy, emotional intelligence, attention-deficit/hyperactivity disorder, love, dependency, enmity, sangfroid, rectitude, and thousands more. While a construct is not a concrete or physical entity, its presence in people varies in degree from person to person. This **variability** implies the noun *variable* is an appropriate description of the degree to which such constructs are expressed by individuals. We employ indirect methods to assess the degree to which an individual exhibits one or more of these constructs (variables). These methods involve formal and informal measurements and observations. They can involve a wide range of techniques, including field-based observation, checklists, inventories, in-office discussions, group process observation, standardized psychometric evaluation tools, and many others.

PROPERTIES AND ATTRIBUTES

In addition to using constructs, counselors can also examine a number of properties that have directly measurable attributes. These properties also vary and thus are variables, and they can be physically seen and directly measured. There is a long list of such properties, including eye color with attributes such as brown, black, green, gray, hazel, blue, etc. Some properties can be scaled to measure various dimensions, including height, serum oxygen levels, weight, annual income, diastolic blood pressure, galvanic skin response, age in months, subcutaneous fat, gender, and hearing acuity. Just because properties are physical entities does not imply they are more accurately measured than psychological constructs. Whenever people employ measurement instruments, there is always some degree of loss of accuracy and precision.

VARIABLES

Not all variables are created equal; they can be expressed at one of four levels of mathematical precision. Variables may be measured as continuous numbers, or they may exist as a series of descriptive names. In the mental health professions, constructs all exhibit a degree of measurable variation. The **precision** of measurement scales used to assess constructs and properties has been organized as a variable with five levels. These are **nominal scales,** two types of **ordinal scales, interval scales,** and **ratio scales.** This sequence of nominal, ordinal, interval, and ratio represents a hierarchy from the least mathematically precise (nominal scales) to the most precise (ratio scales).

Nominal Scales

When the data from a variable exist in categorical form, and those categories cannot be placed into a logical order, then the variable is labeled **nominal.** As the word implies, these categories exist as names only; therefore, the only possible arithmetic operation one can perform with them is counting. For example, clients being counseled as part of their treatment program for an anxiety disorder could include, among many others, those with social phobia, posttraumatic stress disorder, acute stress disorder, trichotillomania, and general anxiety disorder. These mental disorders cannot be sequenced into any meaningful hierarchical scale; they can only be listed. If an arrangement is needed, these mental disorders could be sequenced in alphabetical order or by the nonmathematical reference code number assigned in the American Psychiatric Association's *Diagnostic and Statistical Manual* (2012).

Two Types of Ordinal Scales

Theoretical constructs are not always expressed by measures composed of equal-size units. This restricts the type of arithmetic operations available to the researcher. For example, "faculty rank" is a construct (instructor, assistant, associate, and professor), but it does not qualify for mathematical and statistical operations.

Ordinal Scale Type I

The commissioned corps of officers in the US Army includes the ranks of second lieutenant, first lieutenant, captain, major, lieutenant colonel, colonel, brigadier general, major general, lieutenant general, full general (4 stars), and general of the army.[2] We know these ranks exist in an order, but they are not describable with a mathematical system involving arithmetic operations. To test this, note that if you subtract a first lieutenant from a brigadier general, you don't get a colonel plus change. It is clear that this rank system, like many others (e.g., church hierarchy, NCAA competition level, hospital nurse rank[3]) is an ordered sequence, not a numerical system. For that reason, these types of data are referred to as ordinal data.

CASE IN POINT 7.1

Ordinal scores (positions) from various scales can be compared. Consider the military, where the Army rank of warrant officer is equivalent to the Marine Corps rank master gunner, which is equivalent to the Navy rank of master chief. People holding one of these ranks will have a similar level of responsibility, pay, and authority as their peers in the other branches. Another example is found in academic ranks. In American higher education, the highest academic rank is professor, as it is in British universities. The academic ranks in American universities in descending order are professor, associate professor, assistant professor, instructor, and lecturer. The equivalent academic ranks in the traditional universities of Great Britain are professor, reader, senior lecturer, principal lecturer, and research fellow.

A commonly applied ordinal in educational settings is the level reached in school, or years of education completed. Most graduate students would be hard-pressed to give a meaningful answer to the question "How many years of schooling have you finished?" This is because this variable is an ordinal. Once again, remember our little test: A person with a master of science degree in counseling is not the equivalent of a person with an associate's degree added to another person with a bachelor's degree. The step between a person who completes 3 years of college then drops out and a person who attains a bachelor's degree is a very big step. It is clearly larger than the 1-year step between being a sophomore dropout and being a junior dropout.

Counting college years of education is made difficult by another problem. Does a student who takes 5 years to earn a bachelor's degree have more education than a person who completes the degree in the standard 4 years? Does a doctoral student who takes 6 years to write a dissertation have 5 more years of college than a peer who only took 1 year to write the dissertation?

Ordinal Scale Type II

Professors assign letter grades at the end of every semester or term. Each of these grades is an ordinal datum. We know they are ordinals because they do not have a continuous numerical basis (are not real numbers). The distance between a grade of A and B is not the same in all classrooms, and it is not necessarily the same as the distance between a D and a C in the same professor's classroom. In the pure world of mathematics, letter grades are ordinals, and it is not correct to perform most arithmetic operations with them. Yet we ignore the rules and do what we want with letter grades. We create grade averages and honorific ranks (*cum laude,* meaning with praise; *magna cum laude,* meaning with great praise; *summa cum laude,* meaning with highest praise, *egregia cum laude,* with outstanding praise). We even admit graduate students into our most prestigious programs of advanced study based in part on these averaged ordinal scores.

This violation of the rules of arithmetic also occurs whenever an attitude scale is averaged. For example, when clients evaluate the therapy they received at an outpatient clinic by responding to an opinion-measuring instrument, many clinic administrators feel compelled to average the data. Those data are typically collected as the clients complete their course of treatment and are discharged. The ordinal score from each client's evaluation for each item is added to create a total "opinion score." This total is combined with the total scores provided by other clients of the therapist to determine an average value. As an example, the following frequently appears on questionnaires used for client evaluations of therapist quality:

Overall, how well do you feel your counselor understood your feelings and needs?

5. Always very well attuned to my feelings and needs
4. Usually well attuned to my feelings and needs
3. Not well in touch with my feelings and needs during about half of the sessions
2. Typically not focused on me and my feelings
1. Completely self-absorbed and unaware of my feelings and needs

Clearly, when a number of these items about a therapist are averaged together, he or she would want to have an "average" score that is high, near the value of 5.0. This would indicate that he or she is viewed by clients as being well attuned to their needs. That designation implies a precision level that is someplace between the precision of interval and ordinal scales (Knapp, 1990). This in-between level of precision is designated *ordinal II*.

Another special case is that of a **dichotomy.** There are variables that exist as natural dichotomies (e.g., gender), which cannot be ordered in some form of hierarchy. Thus, such variables are **binary,**[4] making it possible to use them in mathematical analysis. Binary variables can also be artificially created by dividing a ratio or ordinal data set into two parts. One way to do this is to divide the variable at its center point or **median.** Another binary variable could be defined as those who pass the state board's license examination and those who are required to retake it.

Interval Scales

Another class of measurements for variables describes equal interval data without true zero values. These measurement scores are generally described as being **standard scores** or **standardized scores.** They are determined by charting or comparing individual scores against a comparison group. The comparison group is also known as the **norm group** or **normative comparison group.** Such scores can be expressed in a number of formats that share the common property of having scales with equal unit sizes. Having equal-sized units makes it possible to use standard arithmetic and statistical procedures with the measurements. These interval-scaled variables include the Educational Testing Service (ETS)™ scores across the range of tests that corporation publishes, **IQ scores,** scores from developmental inventories and batteries, scores on the Examination for the Professional Practice of Psychology, and scores from the National Board for Certified Counselors' National Counselor Examination.

Ratio Scales

Some variables are measurable using a continuous mathematical scale known as a ratio scale. The ratio scale implies that the units of measurement are "real numbers." This makes the use of fractions and decimals appropriate, as there is a true zero point in the number system. These measures include most physical dimensions, such as the volume of water needed by an outpatient clinic each day, weight of the physical therapist in kilograms, relative humidity of the air in office areas, or range of tones heard by a client expressed in Hz at an intensity of 35 dB. Thus, using a ratio scale, it is possible to determine a patient's body mass index, height in centimeters, oral cavity temperature in degrees Celsius, and beta amyloid peptide count from spinal fluid (Alzheimer's test).

CASE IN POINT 7.2

In a major breakdown of campus communication on November 18, 2011, a dozen nonviolent student protesters at the University of California–Davis were targeted by campus police, who used military-grade pepper spray to force them to disperse. Police-grade pepper spray, like other forms of pepper, can be measured and discussed using a ratio scale of measurement, the Scoville heat unit (SHU).

In 1912 a chemist and pharmacist, Wilbur L. Scoville, employed by Park-Davis Laboratories, devised a measurement scale for the piquancy of peppers and foods containing peppers (Tabasco Inc., 2011). This measure for pepper pungency now has limits defined accurately by modern food chemistry. The pure essence of "hotness or piquancy" is the chemical capsaicin. The oily capsaicin spray interacts with nerve receptors in the mucus linings of the human mouth, eyes, sinuses, lungs and bronchial tubes, nose, ears, and gastrointestinal track as capsacinoid, a chemical that produces a powerful burning sensation.

Photo 7.1 Students and Police at University of California–Davis, November 18, 2011

Wayne Tilcock/*The Davis Enterprise*

Capsacinoid that is 100% pure has a Scoville scale value of 16 million SHU, 15 times greater than the hottest known natural pepper. The military-grade pepper spray used by some law enforcement officers to incapacitate suspects has an SHU index score of 5.3 million. The pain produced by this military-grade pepper spray can be overwhelming and cause temporary blindness, asthma, and a panic reaction. There is anecdotal evidence that it has caused the death of asthmatics by anaphylactic shock.

NOIR

To better remember these scales, the French word *noir,* meaning the color black, provides the handy acronym **NOIR.** When arranged from least to most precise, the types of measurement scales are the following: **N**ominal, **O**rdinal (I & II), **I**nterval, and **R**atio. Examples of these four types are seen in Table 7.1

Variables in Research

Practitioner-researchers engage in studies to determine the relationship among variables and find out how some variables influence others. In qualitative research, many of the variables that are of consequence in understanding individuals are revealed by the research itself. In quantitative studies, the researcher manipulates one or more variables to determine the relationship among known variables.

Independent Variables

In quantitative studies, the variable that the researcher analyzes as the likely change agent for one or more other variables is described as the independent variable. The independent variable is considered the engine that brings about changes and/or modifications in an individual's pattern of growth, learning, or behavior.

The independent variable may be an **active independent variable,** meaning that the researcher manipulates the condition or value of the independent variable. For example, independent variables are manipulated when the researcher decides how many hours

Table 7.1 A Collection of Examples of Each of the Four Types of Measures

Nominal	*Ordinal*	*Interval*	*Ratio*
Blood group	Medical pain scales	Fahrenheit scale	Kelvin scale[5]
Handedness	Movie ratings	WAIS III IQ scores	Dollars in the bank
Least liked food	Opinion surveys	ASVAB scores	Vote count
Brand of automobile	Horse race results	Developmental score	Birth weight
Gender	Social class	GRE scores	Chronological age
Breed of dog	Restaurant ratings	Personality profile	Blood pressure
Psychiatric disorder	Pass/Fail cut score	License test score	MMPI subtotals
Myers-Briggs type	Recommendations	Millon BHI scores	Beck anxiety scores
Illegal drug choice	APGAR scores	Serum creatinine	Chronological age

of counseling a client will receive, whether it is to be individual or group counseling, or whether the counseling program will include significant others in the client's life. Active independent variables can be pharmacological or environmental or be based on behavioral rewards.

The variable could also be an **assigned independent variable,** meaning that the practitioner-researcher has no control over how the variable appears for each subject. For example, the subject's gender may be identified as an independent variable, as it is an important factor in many aspects of each individual's persona. However, the practitioner-researcher has no control over which of the participants in a study are male and which are female. There are numerous other assigned independent variables, for example, age, ethnicity, education level, religiosity, number of siblings and/or birth order, number of parents in the home during childhood, immigration status, and many others.

Applications from the Literature

Examples of Independent Variables from the Literature

Example A

Note: In this study, the impact of four independent variables was studied with the dependent variable of forgiveness. Three of the independent variables were based on demographic characteristics of the subjects, and the fourth independent variable was based on scores derived from a compilation of three measures of personality.

Sample: Two hundred and twenty (220) residents in Mafikeng municipal area participated in this study. Of the 220 participants, 88 (40%) were males and 132 (60%) were females. Age of participants ranged from 18–68 years with mean age of 28.8 years (SD = 11.93). Marital status showed 150 (69.1%) were single, 55 (25.3%) are married, and 12 (5.6%) were separated, divorced or widowed.

Design: The study is a cross-sectional study of two hundred and twenty participants. The independent variables were gender, age, religion, education, marital status and education levels (demographic variables) and three personality factors. The dependent variable is willingness to forgive.

Source: Idemudia, E. S., & Mahri, S. (2011). Can gender, religion, education, age and personality predict willingness to forgive? *Gender and Behaviour, 9*(1), 3765–3781.

Example B

Note: In this study, the author examined whether increased exposure to 3,4-methylenedioxymethamphetamine (MDMA; "Ecstasy") causes decreased tissue stores of serotonin and results in behavioral effects when there are massive releases and subsequent depletions of brain serotonin. Such waves of serotonin in the chemistry of the brain of ecstasy users were thought to have a powerful effect on the user's emotional intelligence (EI). Ecstasy use was the independent variable, and psychological affect levels were elements of the dependent variable.

A description and examples of the use of analysis of covariance (ANCOVA) can be found on the student website for this book.

(Continued)

(Continued)

Design: The independent variable (IV) was drug use (non-user, cannabis only user and polydrug user). The dependent variable (DV) was Emotional Intelligence (EI). The basis of any drug-related differences in EI was explored through analysis of covariance. . . . Drug-related differences in the measures of psychological affect were also explored with drug use as the IV and the measures of psychological affect (anxiety, arousal, depression/hedonic tone) as Dependent Variables. The source of any drug-related differences in the psychological affect measures will be further explored using ANCOVA with the EI and the parenting measures as covariates.

Source: Craig, L., Fisk, J. E., Montgomery, C., Murphy, P. N., & Wareing, M. (2010). Is emotional intelligence impaired in ecstasy-polydrug users? *Journal of Psychopharmacology, 24*(2), 221–231. doi: 10. /0269881108095713

Dependent Variables

In research efforts, the variable that changes in response to the independent variable(s) is known as the dependent variable. Any empirical or qualitative study may have one or more dependent variables. In a qualitative study, the researcher's efforts may be directed toward identifying the possible independent variables of importance in influencing the dependent variable(s). In empirical research efforts, the practitioner-researcher works to determine the extent to which previously identified independent variables (both active and assigned) influence the dependent variable(s).

It is good to keep in mind that an independent variable from one study may be used as the dependent variable in a future study, and vice versa. For example, in a study of charitable contributions, the dependent variable could be the amount of time and or money donated, while the independent variable could be the emotional intelligence level of those participating in the research.

Applications from the Literature

Examples of Dependent Variables from the Literature

Example A

Note: In this research, the dependent variable was the children's level of self-esteem. The principal independent variable was parental alcoholism. The authors introduced other family and home dimensions to learn if they moderate the impact of parental alcoholism.

Family Environment: We developed a measure of family environment based on existing scales, including the Family Environment Scale and the FACES III. Our measure consists of two dimensions: Parental disregard and family stressors. Participants were asked to retrospectively report on perceptions of their family environment. The parental disregard subscale comprised six items (e.g., 'One or both of my parents was not around to take care of me') and the stressors subscale comprised six items (e.g., 'One or both of my parents had trouble keeping a job').

Parental Alcoholism, Family Environment, and Self-esteem: The primary goal of this investigation was to determine if family environment mediates the relationship between the independent variable, parental alcoholism, and the dependent variable self-esteem. Hierarchical regression analysis was performed to test the effects of parental alcoholism on parental disregard, while controlling for participant gender, age, and socio-economic status (SES). The overall model was significant and explained 26% of the variance in parental disregard. Control variables (i.e., participant gender, age, and family SES) were not significantly related to disregard. Perceptions of paternal alcoholism had a moderately large effect, and maternal alcoholism had a small effect on parental disregard.

Source: Rangarajan, S., & Kelly, L. (2006). Family communication, family environment, and the impact of parental alcoholism on offspring self-esteem. *Journal of Social and Personal Relationships, 23*(4), 655–671. doi: 10.1177/0265407506065990

Example B

Note: In this study of over 100 outpatients receiving therapy for serious and persistent mental disorders, the dependent variable involved dimensions of communication ability. The independent variables hypothesized to be related to those dimensions of the dependent variable are listed on the following table.

Independent variable	***Dependent variable***
Sociodemographic and clinical: visiting psychiatric outpatient services, living in flat vs house, diagnosis of mood disorder or neurosis vs belonging to the heterogeneous group *Health-related:* self rated health, quality of life, self esteem, sense of coherence, mastery	Availability of social integration
Sociodemographic and clinical: age, living alone, living in flat vs house *Health-related:* self rated health, quality of life, sense of coherence, mastery	Availability of attachment
Sociodemographic and clinical: age, years of contact with psychiatric services *Health-related:* quality of life, self-esteem, sense of coherence, mastery	Adequacy of social integration

(Continued)

(Continued)

Independent variable	*Dependent variable*
Sociodemographic and clinical: age, living in flat vs house *Health-related:* self-rated health, quality of life, self esteem, sense of coherence, mastery	Adequacy of attachment
Sociodemographic and clinical: visiting psychiatric outpatient services, living alone, living with one's children, living in flat vs house, having high school vs college education *Health-related:* self-rated health, quality of life, self-esteem, sense of coherence, mastery	Total ISSI score

Source: Eklund, M., & Hansson, L. (2007). Social network among people with persistent mental illness: Associations with sociodemographic, clinical, and health-related factors. *International Journal of Social Psychiatry, 53*(4), 293–305. doi: 10.1177/0020764006074540

Predictor and Criterion Variables

In studies that are simply relational in design, such as those studying the concordance or correlation among variables, the variable assumed to be the determining issue or factor may be referred to as the **predictor variable.** (This type of research is discussed in Chapter 8.) Likewise, the dependent variable in relationship studies is frequently given the name **criterion variable.**

Applications from the Literature

Examples of Criterion and Predictive Variables

Example A: Predictors

Note: In this study of one graduate program's training clinic, clients who did and did not finish a counseling intervention were studied, and several predictors of completion were identified. The researchers found lower-income clients were "more difficult" individuals with whom to work but that higher-income clients with healthier Global Assessment of Functioning scores were more likely to drop out of counseling as they felt "improved" by the counseling they had completed.

Results: The goal of this study was to establish which among a set of potential predictor variables would be useful in discriminating among the three groups of clients based on the time and manner at which the client-counselor relationship was ended: intake dropouts, therapy dropouts, and completers. Of the 380 client cases examined, 61 (16.1%) were categorized by their counselors as intake dropouts, 218 (57.4%) as therapy dropouts, and 101 (26.6%) as completers. (Percentages do not equal 100% because of rounding.) These results were consistent with those obtained in other training clinics. As expected, completers ($M = 10.26$, $SD = 9.68$) on average attended more counseling sessions than did therapy dropouts ($M = 5.14$, $SD = 4.76$). A number of potential predictor variables were investigated: age, education (seven levels, ranging from eighth grade or less to beyond a master's degree), annual family income (six levels, ranging from under $10,000 to $30,000 and higher), number of children in the client's living unit, number of the client's presenting problems, perceived client difficulty (low, medium, high), Global Assessment of Functioning Scale score, prior treatment (never, more than a year ago, less than a year ago), gender, employment status (four levels: not applicable, currently employed, unemployed less than 6 months, and unemployed more than 6 months), case urgency (crisis, noncrisis), and referral source (self, other).

Source: Lampropoulos, G. K., Schneider, M. K., & Spengler, P. M. (2009). Predictors of early termination in a university counseling training center. *Journal of Counseling and Development, 87*(1), 36–46.

Example B: Criterion Variables

Note: In this study of how counselor training changes a student's understanding of multicultural factors in the counseling relationship, the criterion measures are subtests of the self-report instrument Multicultural Counseling Inventory (MCI).

Results: Criterion variables, MCI Multicultural Knowledge subscale, MCI Multicultural Awareness subscale, and MCI Multicultural Skills subscale (i.e., those not found to have significant program effects), were conducted using student mean scores as the unit of analysis ($N = 516$). Significant differences for student ethnicity were found on the MCI Multicultural Awareness subscale scores. . . .

Tests comparing MCI Multicultural Awareness subscale scores across all ethnic groups revealed that African American students and Hispanic students scored significantly higher than Caucasian students. To control for these effects on MCI Multicultural Awareness subscale scores, . . . dummy variables for six ethnic groups (Caucasian, African American, Asian, Asian American, Latino(a) or Hispanic, and other) were entered, with the Caucasian variable coded zero to function as the constant. Finally, significant gender effects were found on the MCI Multicultural Skills subscale scores. Male participants' scores on the MCI Multicultural Skills subscale criterion variable were significantly higher than female participants' scores.

We examined the models to identify predictors of student multicultural competencies. First, we observed whether the overall prediction of the multicultural criterion was significant. If it was, then we looked for predictor groups, namely, program cultural ambience, instructional strategies, and clinical training that added significantly to the prediction when prior variables were controlled.

Source: Dickson, G. L., & Jepsen, D. A. (2007). Multicultural training experiences as predictors of multicultural competencies: Students' perspectives. *Counselor Education and Supervision, 47*(2), 76–95.

Parametric Versus Nonparametric

The term **parametric** is often applied when describing ratio and interval data. It means that the variables are real numbers that can have the full range of mathematical operations performed on them. The term applied to ordinal and nominal data is nonparametric. The exception to this generalization is **ordinal type II data,** which are not quite interval but are treated as such. While a large number of nonparametric procedures for data analysis can be used with variables of all four levels of precision, the procedures reserved for parametric variables tend to have greater **statistical power**[6] (Cohen, 1988; Siegel & Castellan, 1988).

Level of Precision

When scientific instruments are used to measure a ratio variable, it is very possible that the resulting score or measurement will be highly precise. High precision of measurement can be seen when medicine is measured and distributed in milligrams. Such scientific measurements are normally ratio scales. It is always possible to reduce the level of precision for a variable, but the reverse is not usually possible. Thus, it is possible to convert a ratio scale such as annual family income into an ordinal scale, such as impoverished, working poor, middle income level, upper income level, wealthy, and über wealthy, or even into a binomial, such as top 1% and other 99%. However, an ordinal scale such as movie-rating scores (e.g., 1 to 4 stars) cannot be changed into a ratio scale.

CASE IN POINT 7.3

A good stopwatch can be used to determine the exact amount of time that an individual needs to solve complex cognitive (performance) tasks on measures such as the Wechsler Adult Intelligence Test–IV. These measures of latency of correct response are used to award bonus points for the problem's solution. This is a highly precise method for scoring. If the test's items were only scored as a pass/fail dichotomy, the resulting scale score would be less nuanced.

CENTRAL TENDENCY

A practitioner-researcher can report data collected in a report listing every individual's score on some variable. That exhaustive approach is detailed but difficult to interpret. For that reason, data are customarily described in summary form. Commonly employed summary statistics include the **average.**[7] In science, there are three forms of average collectively called measures of **central tendency.**

Mode

The **mode** is the category that has the greatest number of cases. This statistic is appropriate for variables of all four (NOIR) levels of precision. There is no way to identify a center point among nominal data, and the lack of any type of sequence in the categories prevents the determination of a central category. However, mode can be determined. For example, *At grand rounds it was recently reported that there were more counseling psychologists in attendance than members of any other of the hospital's departments.* This statement presents the mode value or score (counseling psychologists) in a data set (those attending grand rounds) that is measured using nominal data (departments).

The mode can also be found for ratio, interval, and ordinal data. For example, *Surveys have found that most social workers have a master's degree.* Here the modal category (master's degree) of a variable that is ordinal (education level) has been identified for a data set (social workers). A report of a mode for ratio data can take the following form: *An analysis of all group homes in the state indicates that the largest single expense is the salary of the professional employees.* In this case, the mode (largest expense category) is used to summarize ratio data (facility finances).

The center (average) of ordinal data can be calculated for ordinal, interval, and ratio data but not for nominal data. This average of ordinals is referred to as the median and is the case that lies in the middle of an ordered sequence of data. In other words, it is the halfway point in the data, with half the cases above it and half below. Thus, the median is also the 50th **percentile.** It is important to note that ratio and interval data can be lowered in precision and become ordinal data for which a median can be calculated. In the set of numbers 9, 11, 15, 20, 17, 10, 14, the median (Mdn) is equal to 14. If the data set had two values of 14, the Mdn would still be 14. But if the data set were 9, 11, 15, 20, 17, 10, 14, 13, the Mdn would lie between the value of 13 and 14. When the distributions of ratio or interval data are symmetrical and there are no unusual cases, the two statistical values, the mean ($\bar{X}$ or M) and the median (Mdn) will approximate the same numerical value.

Then there is the special case of ordinal type II data. These ordinals are summarized as though they were equal-interval data, which they are not. As noted above, the use of opinion scales and **grade point averages (GPAs)** are classic examples of this misapplication of the parametric statistics for central tendency.

Mean ($\bar{X}$ or M)

The general public's use of the vernacular word *average* almost always refers to what a measurement specialist would describe as the "**mean** of the data." A statistician would explain the determination of a mean as

$$M \text{ or } \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

In this equation, the symbol $\bar{X}$ is used to represent the arithmetic average or mean of the data. In many journals, the letter M is used as the symbol for the mean because the symbol $\bar{X}$ is difficult to typeset. The uppercase Greek letter sigma Σ instructs us to add all cases from the first individual to the final, or case number n. The term X_i represents the score from any individual or case (sometimes called replicate, subject or participant). The lowercase letter n represents the number of subjects or participants in the data set. The equation can then be read as, "The mean equals the sum of all cases divided by the number of scores (cases) in the set of data."

The mean can only be calculated for ratio and interval data. It is the arithmetic center point in the data and can be viewed as a balance point or fulcrum for the data. Obviously, it is never possible to determine the mean of most ordinal and nominal data because addition and division are arithmetic operations not available with nominal and ordinal data.

NORMAL DISTRIBUTION

The distribution of scores drawn **randomly** from a large population will assume the shape of the **Gaussian normal curve.** This curve has been applied in millions of research studies in the social sciences as a means of interpreting individual scores and making sense of data that has been collected. The normal curve has even been applied to assist parole officers in interpreting the nature of threats posed by sex offenders they supervise in the community.

Applications from the Literature

Basic Interpretation of Data Matching a Normal Curve

Note: This is an interim report from the Dynamic Supervision Project. The authors developed a dynamic measurement plan to estimate the likelihood that a court-adjudicated parolee originally convicted for a sex offense will become a recidivist. In this study, there were eventually 1,000 offenders originally convicted for sexual offenses involving children or nonconsenting adults. The authors used the term *static risk score* as a one-time measure of the risk of recidivism.

Results: This analysis shows approximately 26 percent of offenders scoring in the low static risk category (9 percent projected recidivism risk over 10 years), 42 percent of offenders scoring in the low-moderate static risk category (13.5 percent projected recidivism risk over 10 years), 22 percent of offenders scoring in the moderate-high static risk category (34.5 percent projected recidivism risk over 10 years) and 10 percent of offenders scoring in the high static risk category (45 percent projected recidivism risk over 10 years). The scores of the 500 stable assessments . . . form a normal curve with about 42 percent of offenders falling into the low stable risk category, about 46 percent . . . falling into the moderate stable risk category and about 12 percent . . . falling into the high stable risk category.

Source: Harris, A., & Hanson, R. K. (2003). The dynamic supervision project: Improving the community supervision of sex offenders. *Corrections Today, 65*(5), 60–64.

The Bell Curve

Graduates preparing to sit for their state's licensing examination have been reported as asking "Is the test graded on a curve?" Few really know what they are asking (for more about the normal curve, see Chapter 9).

Applications from the Literature

Frequently Asked Question About the Licensing Examination of the Association of Social Work Boards (ASWB)

Note: The license examination for professional social workers, like the exams for professional counselors and psychologists, is graded on a pass/fail basis against an absolute standard for performance set by the various states.

Q. Is the exam graded on a "curve" for the whole country or the area you are in?

A. The exam isn't graded on a "curve."

The grading "curve" we are used to from school is a process by which the instructor moves the grading scale up or down, depending on the performance of the group being tested. This does not happen with the ASWB examinations.[8]

Source: Assocation of Social Work Boards. Retrieved from http://www.aswb.org/SWLE/faqs.asp#Curve

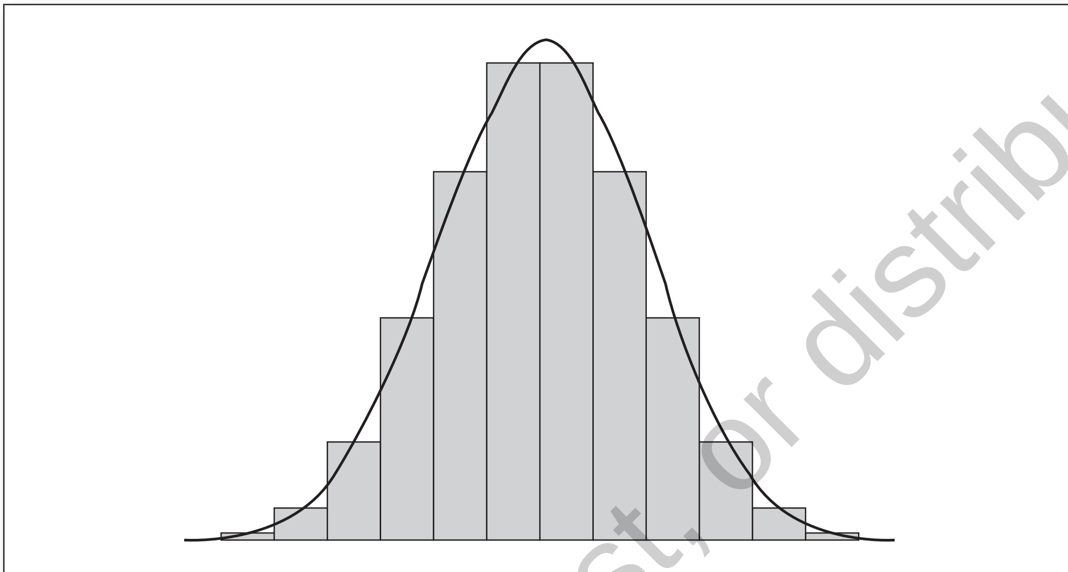
Nonetheless, all major licensing examinations use a curve in the scoring process. For students to understand the "curve" that is used in scoring licensing examinations, they must be familiar with the **normal distribution of errors** first defined 330 years ago. In 1686, the second Lucasian Professor of the Natural Philosophy of Mathematics at Cambridge University, Isaac Newton, published a book that changed Western science. It described his new calculus and applied it to understanding gravitation and planetary motion.[9]

In Newton's later life, one of the mathematicians in his orbit was Abraham de Moivre, a Huguenot, who escaped from France when King Louis XIV purged all Protestant intellectuals. This Huguenot émigré was an important mathematician in his own right. de Moivre studied the properties of the binomial distribution and, with the help of the calculus of Newton, was able to mathematically define what is now recognized as the **normal distribution** in 1738 (de Moivre, 1738/1985).[10]

In the early 19th century (1809), Carl Friedrich Gauss applied this mathematical model to the astronomical research dealing with asteroid orbits. It is Gauss whose name is associated with the development of this powerful mathematical tool (Wright, 2008). Today, most graduate students have never heard of either de Moivre or Gauss and refer to the statistical distribution not by its originator but by its shape, the **bell curve.**

The curve depicts a randomly occurring, continuous probability distribution. It can be used to approximate a number of different distributions in nature (Anton, Kolman, & Averbach, 1988). For example, the normal curve can approximate the distribution of the weight of the population of adults, the GRE scores (**raw score** form) of all college seniors applying for admission to graduate school, the length of blades of grass measured in centimeters growing wild in a meadow, or the IQ of the population of clients seeking vocational guidance and employment counseling.

Figure 7.1 Distributed Binomial Scores Overlaid by a Gaussian Normal Curve



Source: Wright, R. J. (2010). *Multifaceted assessment for early childhood education*, p. 92. Reprinted with permission from Sage.

Central to the use of all standardized measures, including national licensing examinations in the mental health professions, is the assumption that raw scores from the population of tested individuals are distributed as a Gaussian normal curve. When data are plotted from a statistical test, the "goodness of fit" test can be used to determine whether the new distribution is highly similar to the Gaussian normal distribution. (Goodness of fit is described in Chapter 11.) There are several Internet solutions for making this judgment:

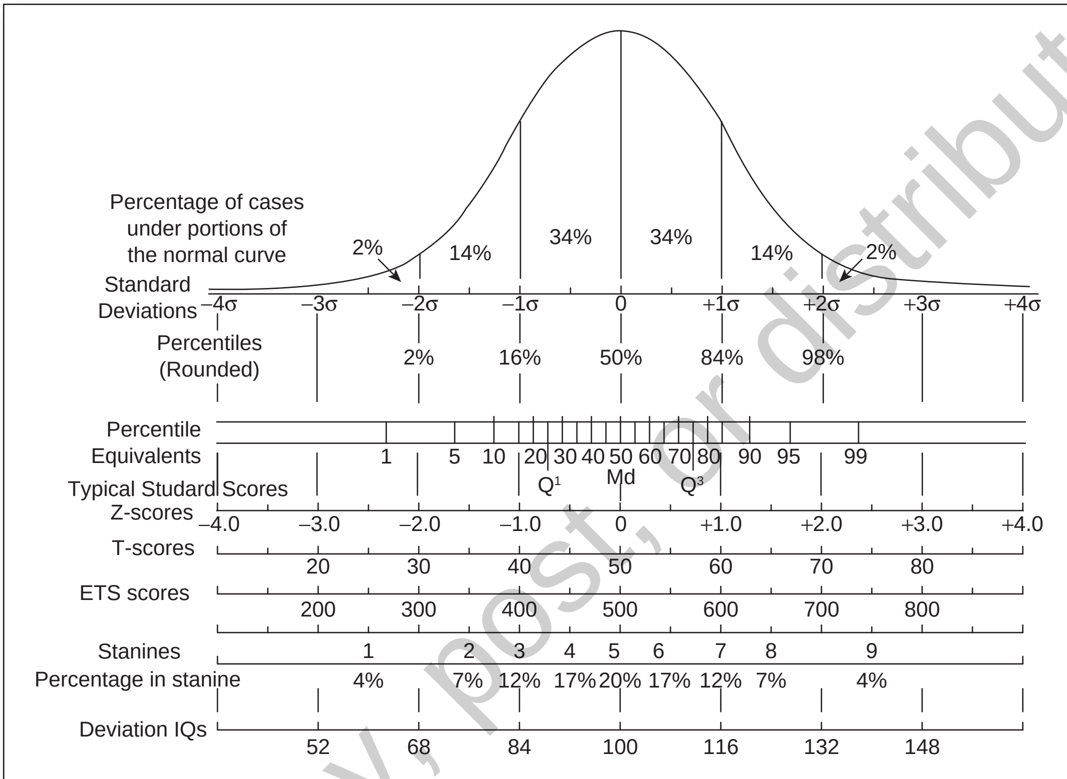
From Laura Schultz at Rowan University, http://users.rowan.edu/ ~ schultzl/TI/chi-square.pdf

From Richard Lowry at Vassar College, http://faculty.vassar.edu/lowry/csfit.html

From GraphPad Software, http://www.graphpad.com/quickcalcs/chisquared1.cfm

This calculation can also be done by other statistical software available on many universities' servers. Programs being used today on campuses across the country include, among others, those from ASReml, BMDP, EViews, Maple, Mathematica, Minitab, NMathStats, SAS, SPSS, STATA, and SYSTAT. The technology support office of most universities provides brief inservice courses for individuals wishing to learn to use the local system.

Figure 7.2 An Elaborated Gaussian Normal Distribution



Source: Wright, R. J. (2010). *Multifaceted assessment for early childhood education,* p. 52. Reprinted with permission from Sage.

Skew

There are several ways the normal symmetry of a distribution can be lost. One occurs when there is an excess of extreme scores, that is, either too many high or too many low scores. The excess of unusual scores results in a separation of the mean and the median of ratio or interval data being examined. A surplus of unusually high or low scores will cause the mean to be drawn in the direction of the unusual cluster. This excess of scores at either the high or low end creates a **skew** in the data. The median tends to be much less volatile and only exhibits a very small movement when there is an unusual cluster of either high or low scores. Table 7.2 presents GPA data from a counseling study conducted with college freshmen at a small, private university.

In this hypothetical experiment, the participants were 38 freshmen students who had self-referred to the campus counseling center to learn how to improve their study habits

Table 7.2 GPA Data From Experimental and Control Groups

	Experimental Group	*Control Group*
	2.66	0.16
	2.70	3.20
	2.75	3.21
	1.76	0.75
	3.76	2.78
	2.78	3.78
	3.80	2.80
	2.78	3.80
	1.82	2.82
	2.83	2.80
	2.85	2.85
	2.86	3.86
	2.78	2.80
	2.89	2.80
	2.90	2.90
	3.92	3.92
	1.94	2.94
	2.96	3.96
	2.97	0.07
Number (n) =	19	19
Mean ($\bar{X}$) =	2.83	2.59
Mode =	2.78	2.80
Median (Mdn) =	2.83	2.82
Skew =	0.343	–1.11*
SE-skew = 0.52		

*Significant level of skewness[11]

and achievement in college. These students all achieved less than a GPA of 2.00 during the first semester. In the spring term, half (n = 19) were randomly assigned to a small class teaching study skills, test-taking strategies, and critical reading at the college level. The **experimental group** (n = 19) all received individual coaching by a counseling intern for 3 hours a week. The coaching covered similar topics as those covered in the classroom (control or comparison group).

The outcome of this study could be measured in terms of the students' GPA for just the second semester (spring term). Data from the study presented in Table 7.2 show a large mean ($\overline{X}$) difference but little difference in the median scores. This indicates a large skewness toward the low side of the distribution within the control group's data. This can be confirmed by visually examining the data from the control group. In that data set, 3 members of the classroom-taught group (control group) ended the spring term with a GPA lower than 1.00 (GPA ≤ 1.00).

This resulting separation of the mean and median of this hypothetical data provides evidence of significant skewness in the data. When the distortion is caused by unusually low scores, the skewness has a negative value, and when there is an excess of unusually high scores, the skewness is referred to as positive skewness. The mean will be greater than the median when the data have a positive skewness. Theoretical distributions depicting these types of skewness are presented in Figure 7.3.

In this age of accountability, practitioners are all too frequently required to justify their interventions and therapeutic choices. In this regard, it is important to note that a few disgruntled clients can introduce **negative skewness** into the data and thereby reduce the mean score for a program of an individual therapist. In the example study described in Table 7.2, there is a very high probability that the scores exhibit a significant level of skewness. The computer-calculated value for the skewness of these data is $Sk = -1.11$, $p > .01$.

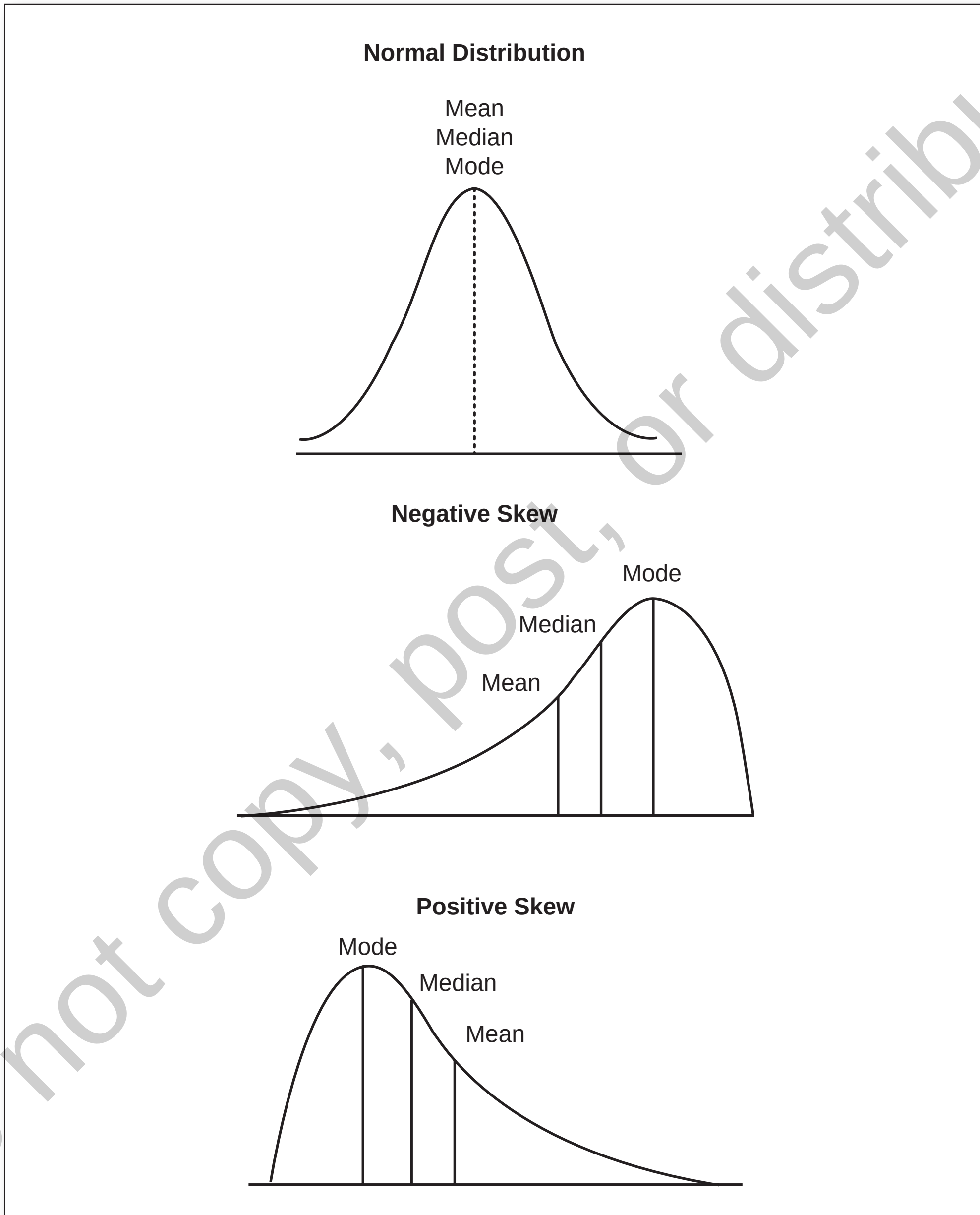
CASE IN POINT 7.4

The nursing, social work, medicine, dentistry, nutrition, and clinical psychology staff of a large urban children's hospital and child guidance clinic went out on strike a few years ago. By the 4th week, both sides were at a serious impasse, and children and their families were suffering. During a press conference called by the public relations office, the press was informed that the clinic staff was being greedy.

The point was made that the average professional earned a salary of over $85,000 per year. When benefits were included in the total package, the average professional staff member received over $125,000 in annual compensation. At a local church, the representatives of the bargaining unit held a press conference a day later. The union reported that the average full-time professional earned just under $58,000 a year. Also, when the benefits were added to the mix, total compensation was only $76,500.

The fact is that both parties reported the same data and both told the truth. The data about pay were badly skewed. Medical staff, board-certified neuropsychologists, and senior nurse administrators were paid at a much higher level than were the larger groups of professionals (counselors, floor nurses, nutritional specialists, social workers, and professional nurse assistants). This excess of high salaries resulted in a positive skew in the salary data. At that time, the vast majority of professional employees were paid lower salaries and had modest benefits.

Figure 7.3 Curves With Negative and Positive Skewness



Source: Wright, R. J. (2010). *Multifaceted assessment for early childhood education,* p. 95. Reprinted with permission from Sage.

DATA PRESENTATION

The goal of data presentation is to make it simple for others to visualize and understand the data you wish to describe. The practitioner-researcher has a number of approaches available to employ when presenting his or her data. The precision of the variable being described is one factor to keep in mind when deciding which approach to employ.

Nominal Data

Nominal data are organized in name-only categories and do not exist in any hierarchical sequence. Therefore, the presentation method must not imply a logical sequence or hierarchical ordering of the categories. Three methods for depicting nominal data commonly appear in the scientific literature, official reports, and even in mass media. These are the pie graph, bar graph, and pictograph.

Pie Graphs (Charts)

This form of graphic presentation uses a circle format and therefore implies no beginning or ending point. The lack of a high-to-low value implication makes a pie chart appropriate for nominal data presentations. It is also appropriate for presenting data measured at more precise levels.

Table 7.3 Presenting Descriptive Data

NOIR Classification	*Graph Formats*	*Appropriate Central Tendency Statistics*
Ratio	pie graph, bar graph, pictograph, histogram, frequency polygon, continuous curves and ogives	mean median mode
Interval	pie graph, bar graph, pictograph, histogram, frequency polygon, continuous curves and ogives	mean median mode
Ordinal II	pie graph, bar graph, pictograph, histogram, frequency polygon	mean median mode
Ordinal I	pie graph, bar graph, pictograph, histogram, frequency polygon	median mode
Nominal	pie graph, bar graph, pictograph	mode

Nominal data can be presented on a pie graph by first assigning each case (data point) to one of the nominal categories. Next, the number of cases (data points) in each category is converted into a percentage of all the cases in the data set. The final step is the assignment of portions of the "data pie." The pie is divided proportionally according to the percent of the total data set represented by members in each category. This is done by plotting the percent assigned to each nominal category on a circle (pie graph). There are 360° making up any circle, so a nominal category including 25% of the cases would be assigned 25% of the circle (data-pie), and 25% of the 360° of a circle is 90°. Figures 7.4 and 7.5 show some examples of pie graphs.

Pie graphs are widely employed in presenting budget data and other statistical administrative information. The ease of reading and understanding pie charts makes them the graph of choice for many financial officers.

There are a number of free programs for constructing various graphs, including pie graphs. Following are three of these:

National Center for Educational Statistics, http://nces.ed.gov/nceskids/createagraph/default.aspx

National Council of Teachers of Mathematics, http://illuminations.nctm.org/ActivityDetail.aspx?ID=204

MrNussbaum.com, http://mrnussbaum.com/coolgraphing/

In addition to stand-alone software packages for graphing, major statistical packages used extensively by faculty and graduate students in counselor education programs also provide high-quality graphing systems.

ASReml, VSN International, http://www.vsni.co.uk/software/asreml/

BMDP, http://www.statistical-solutions-software.com/products-page/bmdp-statistical-software/

EViews v.7.2, http://www.eviews.com/

IBM SPSS (also see Appendix A), http://www.spss.com/software/statistics/text-analytics-for-surveys/index.htm?tab=1

Maple v.16.0, http://www.maplesoft.com/products/maple/

Microsoft Excel, http://office.microsoft.com/en-us/excel/

Minitab v.16, http://www.minitab.com/en-US/products/minitab/default.aspx

NMath Stats, v.3.5, http://www.centerspace.net/products/nmath-stats/

SAS Institute, http://www.sas.com/ads/google_stats/?gclid=CKfxm-L9iqgCFUta2godcQWECQ

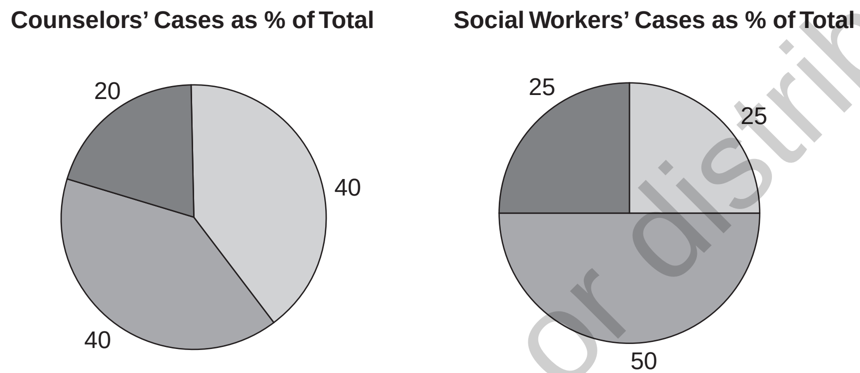
STATA v.12.0, http://www.stata.com/stata12/

SYSTAT v.13, Cranes Software International, http://www.systat.com/

Wolfram Mathematica v.8, http://www.wolfram.com/mathematica/

Figure 7.4 Two Pie Graphs With Nominal Data

(Right) A graph of the types of cases seen by the social workers of a family guidance clinic; (Left) a graph of the types of cases seen by the clinic's professional counselors.



Medium gray = Drug and/or alcohol abuse.

Light gray = Abusive relationships within the family.

Dark gray = Affect disorders.

Source: Developed by author.

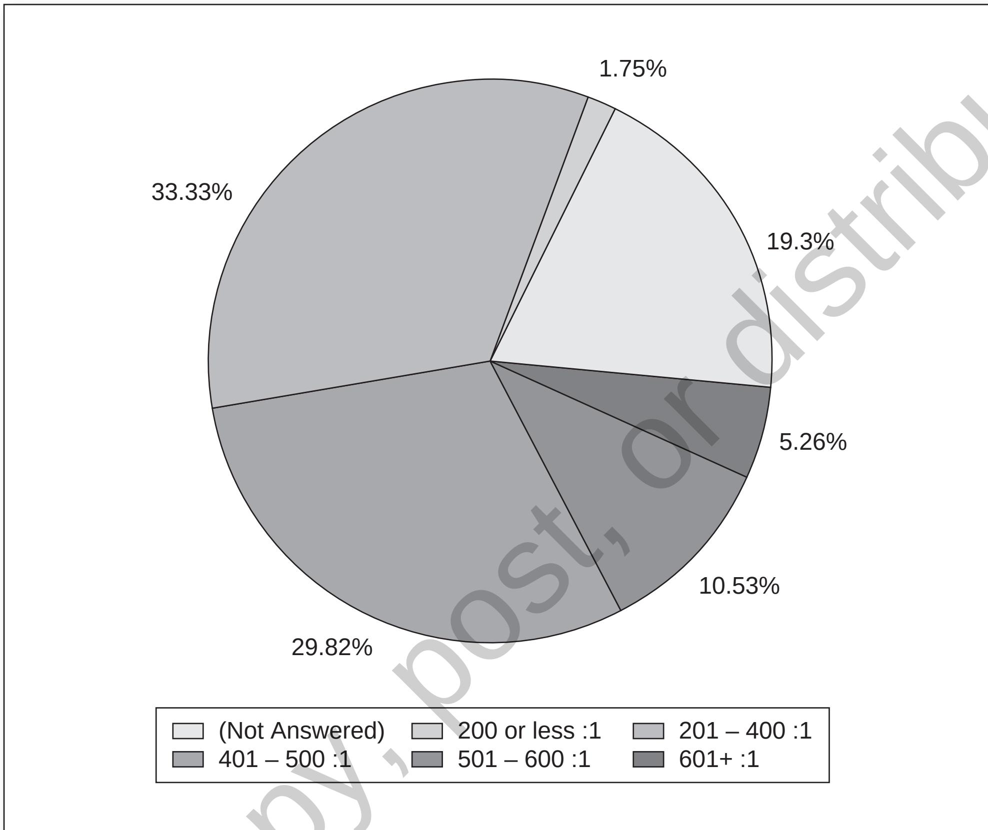
Bar Graphs

Like pie graphs, **bar graphs** are appropriate to use when presenting a set of data found at any level of measurement precision, including nominal data. Generally, bar graphs are plotted using two axes meeting at a right angle, the **ordinate** (y-axis) and the **abscissa** (x-axis). The categories of the nominal variable are then listed along the abscissa, and the number of cases at each nominal category is listed on the ordinate. Bars are drawn to a height (distance from the abscissa) representing the number of cases. This is also described as the **frequency** (f). This height can be read against the ordinate to see how many cases occur at any point on the abscissa. Figures 7.6–7.8 show examples of bar graphs.

Bar graphs can also be used to depict variables measured using a more precise scale. Figure 7.7 presents Medicare costs in dollars (ratio scale) from 2001 and projected through until 2017 (time is a ratio scale).

The practitioner-researcher may elect to use a format for the bar graph that is set at a 90° rotation from those above. The same rules apply with this and other formats for bar graphs.

Figure 7.5 Simulated Example of a Pie Graph from Published Data

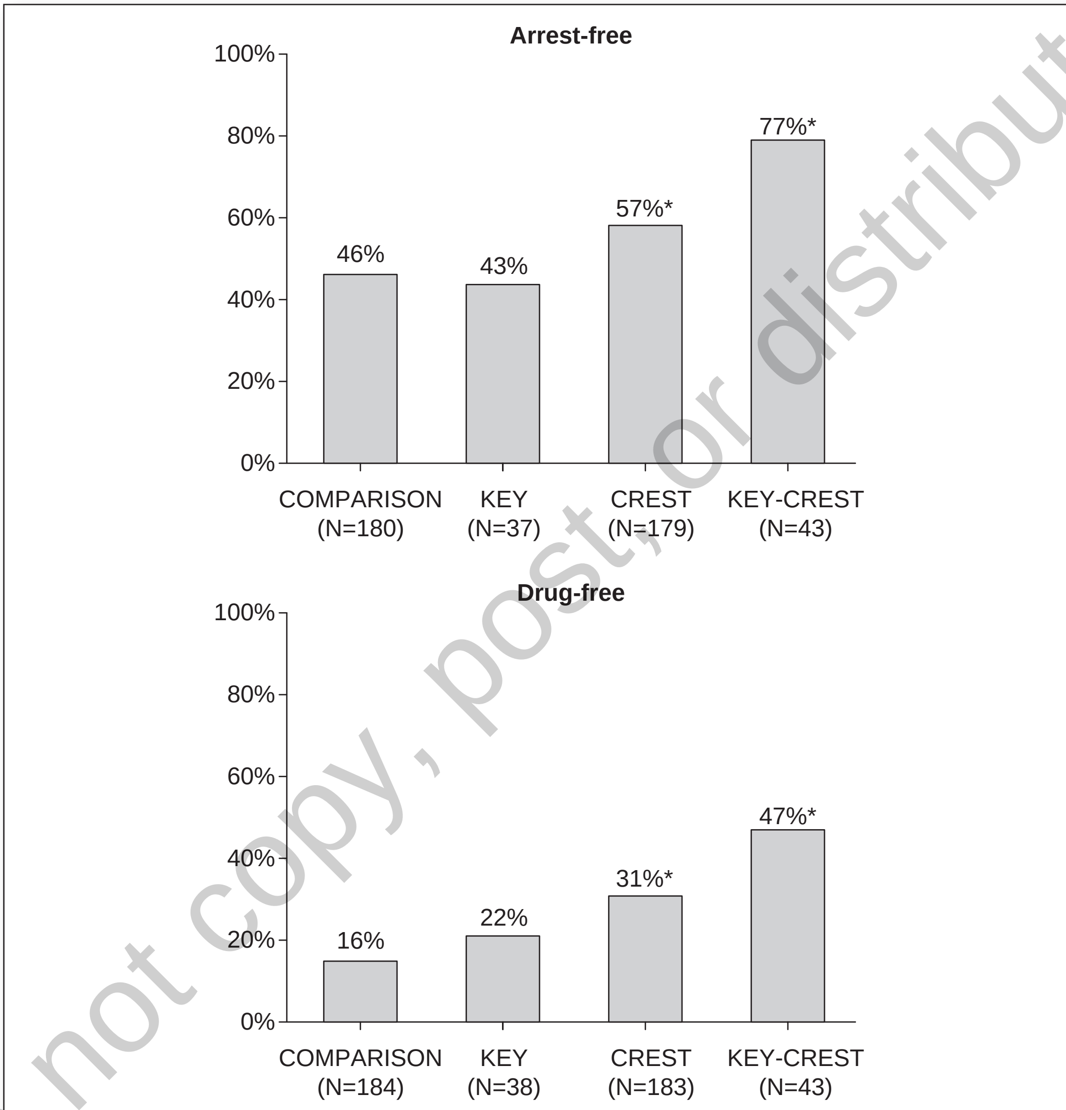


Much Ado About Nothing: Bar graphs can be distorted to reflect the bias of the researcher reporting the data. One of the key ways to do this is to not start the graph at the zero point. By starting the scale at a higher point, a small difference between groups can be magnified to appear much larger than it actually is. Another way to exaggerate findings is to report the log transformation of the original data. Figure 7.9 shows an example of an unexaggerated and exaggerated graph.

Pictographs

Data presentations that would typically be shown as a bar graph or pie graph may also be drawn as a **pictograph.** Pictographs are more dynamic and intriguing for readers than are the more traditional graphical formats. Figure 7.10 is a pictograph. These picture-based graphics involve using line drawings, cartoons, or other artwork to design easy-to-interpret charts. They are used to explain data to children in children's magazines, and they are

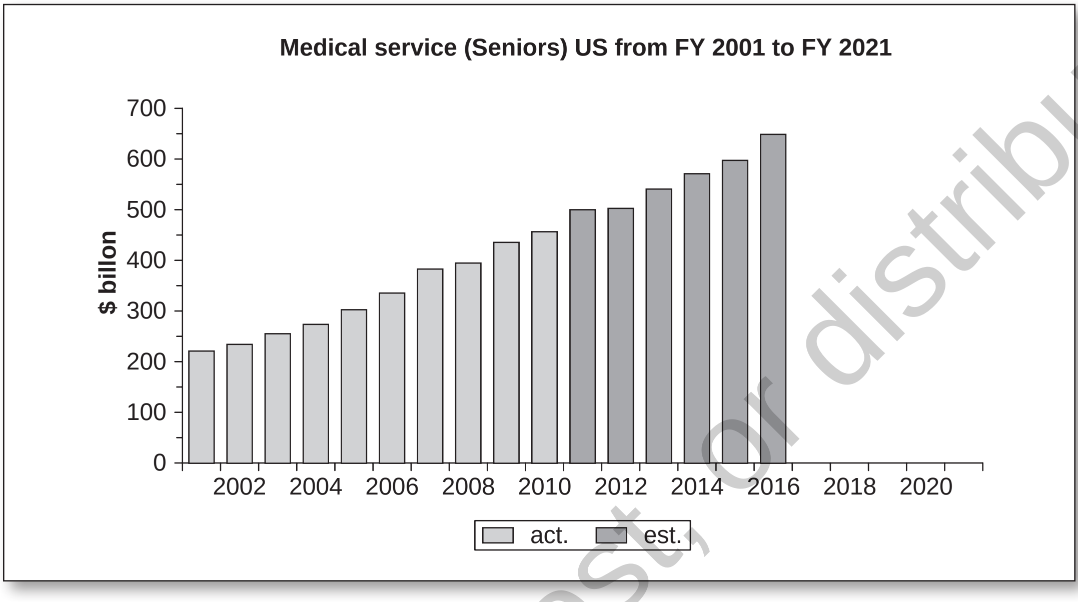
Figure 7.6 Bar Graph of Results From an Intervention Program for Released Drug Offenders



Source: Martin, S. S., Butzin, C. A., Saum, C. A., & Inciardi, J. A. (1999). Three-year outcomes of therapeutic community treatment for drug-involved offenders in Delaware: From prison to work release to aftercare. *The Prison Journal, 79*(3), 294–320. doi: 10.1177/0032885599079003002

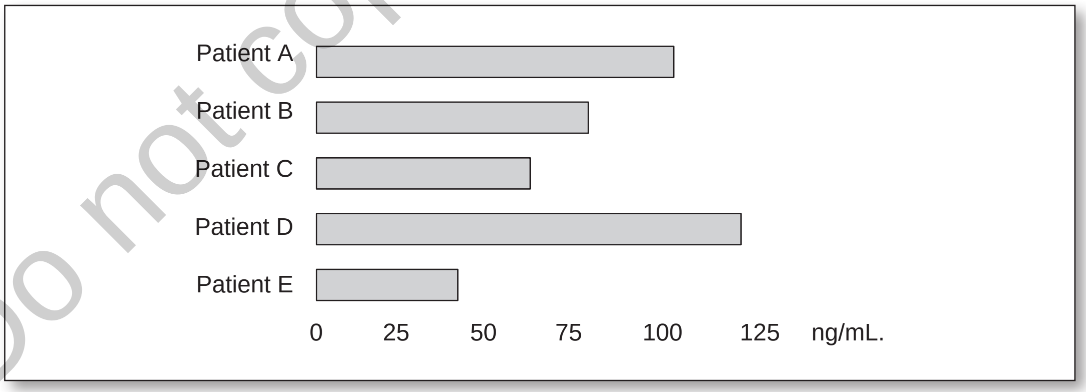
Note: This study of three models for providing treatment for addiction to illegal drugs was conducted in Delaware. The KEY program was conducted within the prison environment using a therapeutic community. The CREST program involved the members of the therapeutic community in a work-release program. The KEY-CREST was a combination of both approaches. The asterisks indicate that the difference between this average and that of the comparison group was beyond a chance occurrence.

Figure 7.7 Medicare Costs Projected Through 2021



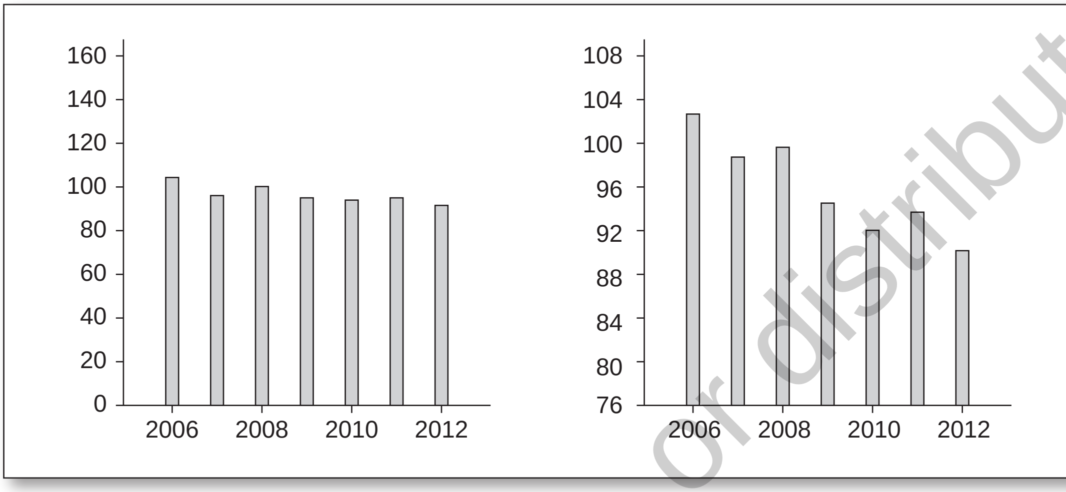
Source: http://www.usgovernmentspending.com/downchart_gs.php?year=2001_2021&view=1&expand=&units=b&fy=fy12&chart=12-fed&bar=1&stack=1&size=l&title=&state=US&color=c&local=s

Figure 7.8 Serum Methadone Levels in Five Patients 1 Week Into a Combined Therapy Treatment Program (Counseling With Methadone)



Source: Developed by author.

Figure 7.9 Bar Graph With a Zero Point and Bar Graph Uncorrected Without a Zero Point



Source: Developed by author.

Note: On the left is a graph showing the 7-year history of staphylococcus infections among clients in a residential drug rehabilitation center for adolescents. The same data are also depicted on the right bar graph. Which graph do you think the center's administrators shared with state child welfare inspectors?

Infections including staphylococcus and HIV are not uncommon among users of illegal drugs. Employees, including therapists, need to pay attention to sanitation when in close physical contact with rehabilitation inpatients (Hughes & Gray, 2009).

employed by mass media designed for adults who are on the go. This format for graphic display of data is inappropriate for presentations that require precise data interpretation.

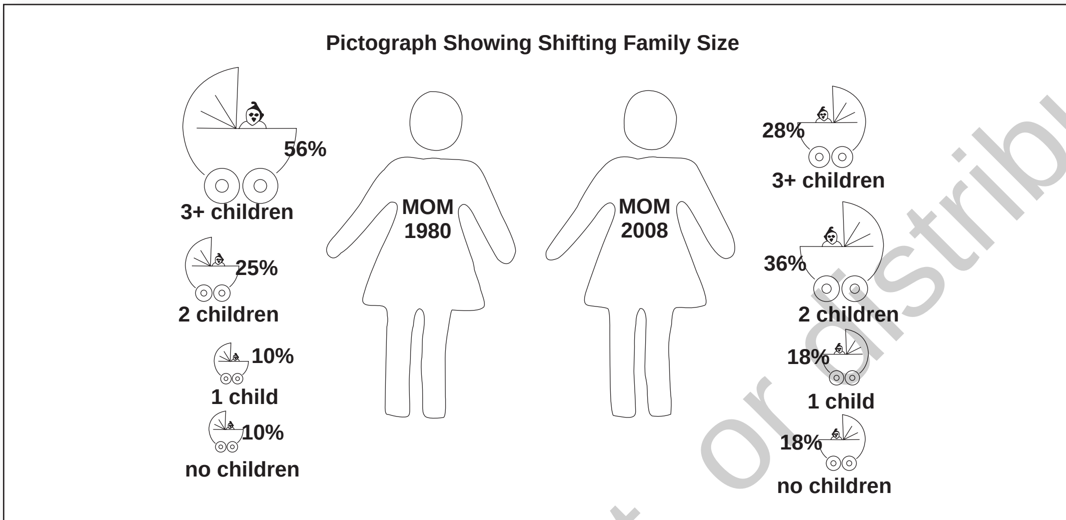
Ordinal Data Presentation

Ordinal data can also be represented by pictographs, bar graphs, and pie graphs, but these data may be depicted by other formats as well.

Histograms and Frequency Polygons

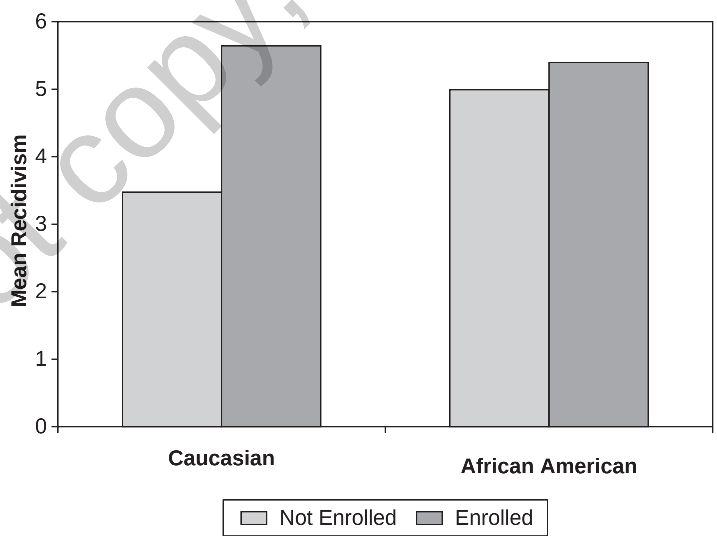
The alternative formats include **histograms** and **frequency polygons** (line graphs). These graphic display systems are inappropriate for displaying nominal data but ideal for presenting ordinal data. This reflects the fact that they present information in a hierarchical sequence. What histograms and frequency polygons do well is present ordinal and interval data. A histogram appears to be much like a bar graph but for the fact that the columns are all contiguous. Figure 7.11 shows an example of a histogram.

Figure 7.10 Pictograph Showing Shifting Family Size

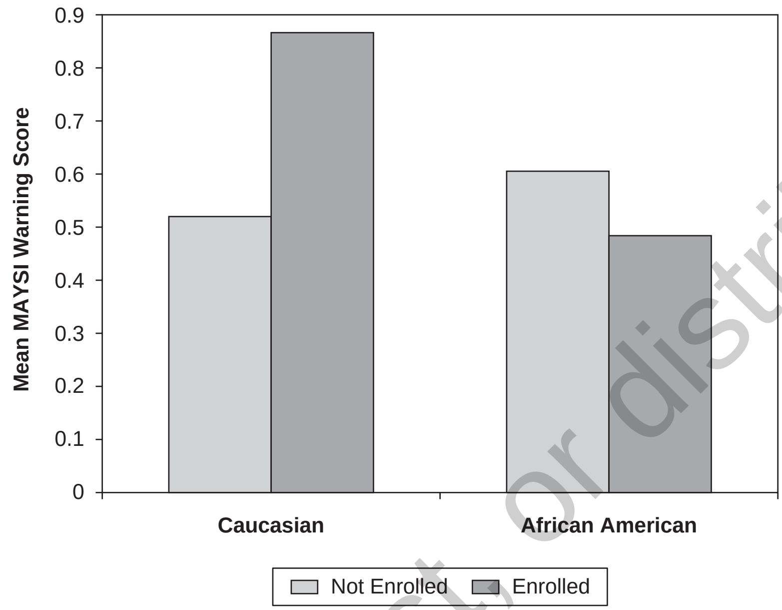


Source: Drawn by author based on data from US Census Bureau.

Figure 7.11 Four Histograms From a Study of Recidivism Among Juvenile Offenders



Mean Recidivism × Race (African American vs. Caucasian) Interaction



Mean Massachusetts Assessment of Youth Screening Inventory (MAYSI) Warning Score × Race (African American vs. Caucasian) Interaction

Source: Lopez-Williams, A., Stoep, A. V., & Stewart, D. G. (2006). Predictors of mental health service enrollment among juvenile offenders. *Youth Violence and Juvenile Justice, 4*(3), 266–280. doi: 10.1177/1541204006290159

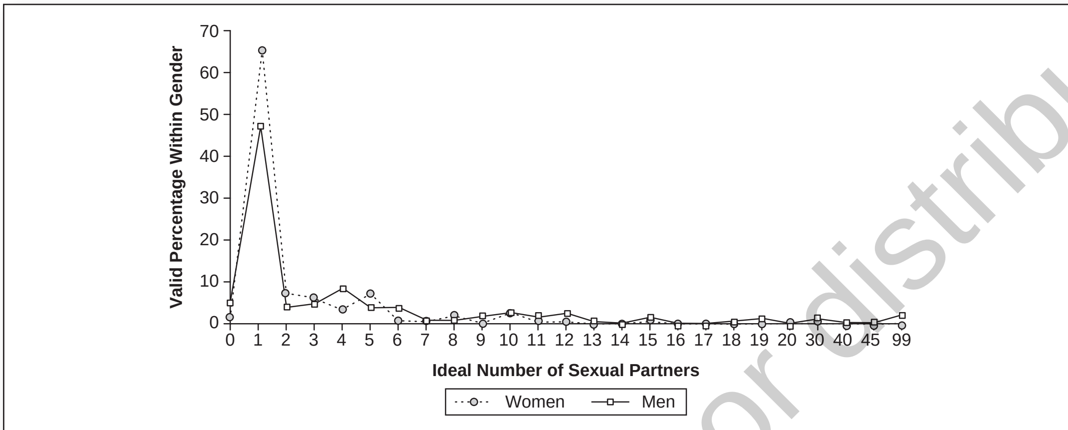
Note: In this study, juvenile detainees between the ages of 13 and 17 years in a correctional facility were assessed using a number of demographic and psychological measures. One research question the authors hoped to resolve was which factors from the youthful detainees' backgrounds were predictive of whether they were enrolled in a counseling-based mental health program at the time of adjudication.

A line graph has no columns but uses point references off the abscissa (horizontal axis), representing frequency of occurrence. The abscissa lists the ordinal groups in sequence described by the line graph. A continuous line is drawn, connecting the points of reference for frequencies of occurrence. The rules for histograms and frequency polygons (line graphs) are similar to those for bar graphs. One difference is that the variable used to group the data must be provided in a logical hierarchical sequence. Figure 7.12 shows an example of a frequency polygon.

Curves and Ogives

When data have equal-size units that are measured as **real numbers,** they can be presented by a smooth, curved line.[12] One form of such a curve, a **continuous distribution**

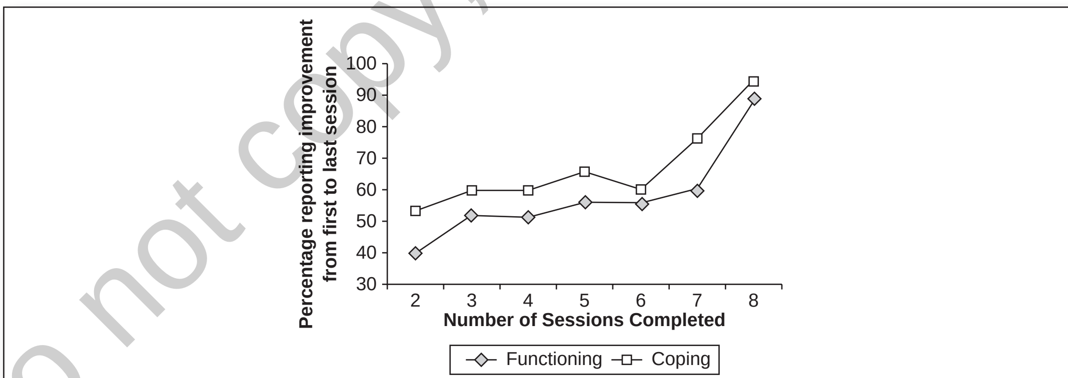
Figure 7.12 Overlaid Frequency Polygons With Skewed Data of Undergraduate Students' Ideal Number of Sex Partners



Source: Pedersen, W. C., Miller, L. C., Putcha-Bhagavatula, A. D., & Yang, Y. (2002). Evolved sex differences in the number of partners desired? The long and the short of it. *Psychological Science, 13*(2), 157–161. doi: 10.1111/1467-9280.00428

Note: This depicts the ideal number of sex partners envisioned by 266 undergraduates over the next 30 years of their lives. These data are badly skewed (positive skew), as most students prefer having only one lifelong sex partner.

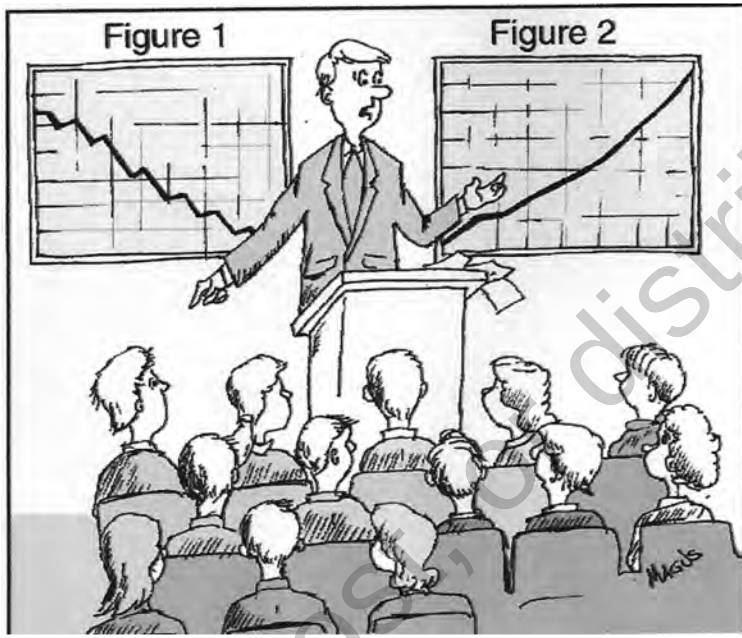
Figure 7.13 Frequency Polygon Depicting Outcomes From a Solution-Focused Brief Counseling Program



Source: Fischer, R. L. (2004). Assessing client change in individual and family counseling. *Research on Social Work Practice, 14*(2), 102–111. doi: 10. 1177/1049731503257868

Note: This study was conducted over a 2-year period with almost 4,000 subjects and 40 counselors providing a program of solution-focused brief counseling. The maximum number of sessions was eight, and the clients were assessed for their perceived improvement in coping with problems and interpersonal functioning. This frequency polygon indicates that eight may be an ideal number of sessions to provide.

Cartoon 7.1 Fun With Statistics



Cartoon by Merv Magus.

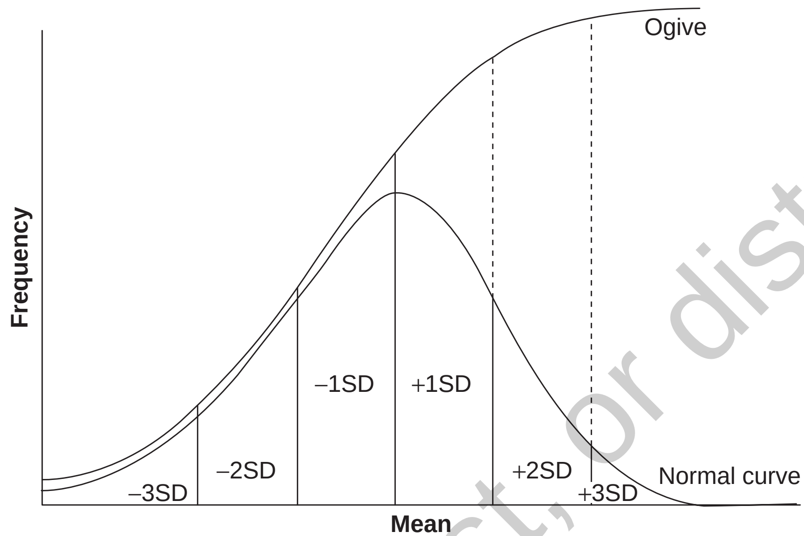
"Chart #1 shows the client satisfaction survey data for our clinic last year, and Chart #2 shows how our program director was able to reinterpret them."

curve, moves from low to high scores along the abscissa (horizontal axis), and the frequency is depicted by the height as measured along the ordinate (vertical axis) of the curve at any point along the abscissa. The other form of the smooth line curve is the **ogive** (pronounced o-jive). This curve is a cumulative statement about the data (see Figure 7.14). Starting with the lowest score on the graph, the frequency of each increasing score is added and plotted as a curve.[13]

Smooth line curves may assume the characteristically symmetric shape of the Gaussian normal curve or any of a number of other shapes, depending on the nature of the scores that are being plotted (see Figure 7.15). The ogive is a variation of the bell curve. By plotting the continuous accumulation of scores along the ogive, the practitioner-researcher can see where any sharp change occurs in the graph's inflection (slope). These points of inflection provide researchers with information about the research participants' average point on one variable when they suddenly have made new insights or when a breakthrough has occurred.

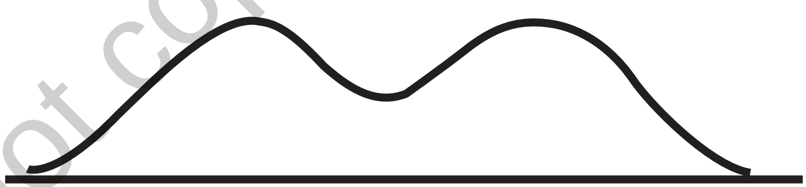
An ogive of this distribution has a slow change in the slope of the curve and starts at a higher level than a truly normal distribution.

Figure 7.14 Ogive and Normal Curve Plotting the Same Data Set



Source: Wright, R. J. (2008). *Educational assessment: Tests and measurements in the age of accountability,* p. 237. Reprinted with permission from Sage.

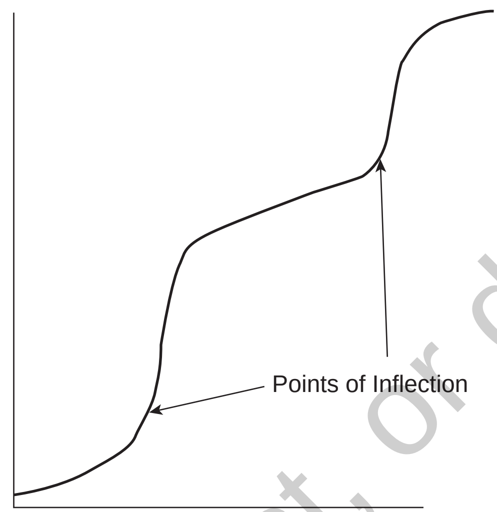
Figure 7.15 Variations in Smooth Line Graphs



Bimodal curve: Smooth curve depicting the performance of a measure with two different centers of maximum population density. **Bimodal** curves have two modes and may indicate to the researcher that there are two underlying distributions of subjects. This may occur when diverse groups (samples) are combined and measured on a single defining variable. The ogive plotting these data will have two points of inflection, one for each mode.

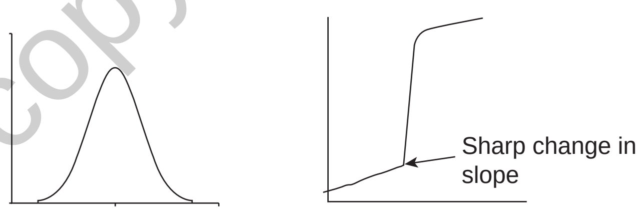
Source: Developed by author.

Figure 7.16 Bimodal Ogive



Source: Developed by author.

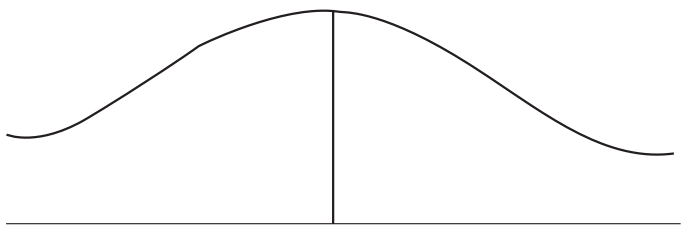
Figure 7.17 Leptokurtic Distribution Leptokurtic Ogive



This peaked or **leptokurtic** smooth curve depicts data with little individual score variation reflecting how many scores are near the mean. This may occur when a point is reached in a pharmacology intervention where larger concentrations of the serum medication no longer bring about noticeable improvement for the patient. The ogive for this type of distribution will have one sharp inflection with an abrupt change in slope functions.

Source: Developed by author.

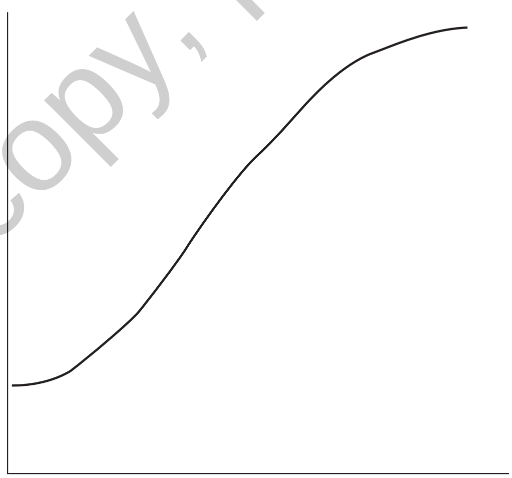
Figure 7.18 Platykurtic Distribution



A flat or **platykurtic** smooth curve depicts data with a high degree of variation, with many individuals' scores spread away from the mean. This distribution is sometimes called the **fat tail** distribution. It can occur when there is excess variation in the data. For example, this distribution can occur with data from discourse analysis. When a researcher reads transcripts of interviews searching for and recording the expression of certain empathy versus self-centered concepts, the distribution may be platykurtic, indicating that some participants are very empathetic while others are very self-centered.

Source: Developed by author.

Figure 7.19 Platykurtic Ogive



Source: Developed by author.

VARIABILITY AND VARIANCE

Dispersion

For a variable to exist, it must occur in varying amounts in different individuals. Therefore, when scores from all individuals are graphed, there will be a spread of score values. Some will be above the median and some below. In the case of smooth curved data, some individuals will score on either side of the mean and median. The amount of this spread of scores is **dispersion,** and it is described in a number of ways.

Nominal Data

Dispersion for nominal data can only be expressed by counting the categories used to group the data. For example, "During the homecoming parade on campus this year, floats representing all eight social fraternities and seven sororities took part."

Ordinal Data

Ordinal data can be shown as the **range** of scores. This can be expressed as a number representing the distance between the first and last score in the set. "Students of our program had GRE Verbal scores ranging 21 points, from a low of 140 to a high of 161." Or, "The faculty committee included one member from each of four levels of rank, from instructor to full professor."

Ordinal data ranges can be greatly influenced by just one outlying case. In our example above, one graduate student with a 169 on the Verbal section of the GRE would change the range of scores from 21 to 29. To prevent that type of distortion, practitioner-researchers usually report the range of scores between the middle two **quartiles,** or the **interquartile range.** This is the range of scores for the middle 50% of the cases. The scores must first be arranged from lowest to highest (sequenced hierarchically). Then the 25th percentile can be found as the score below which 25% of all scores are located. Likewise, the 75th percentile is the case below which there are 75% of all cases. The score range between these two is the interquartile range.

Finding the Interquartile Range

- First arrange the student-reported hours spent by 19 graduate students completing graduate reading assignments each week in a sequence from low to highest.

  2, 5, 6, 6, 8, 9, 9, 12, 15, 18, 18, 19, 25, 27, 30, 31, 35, 45, 70

- Next find the median and highlight it.

  2, 5, 6, 6, 8, 9, 9, 12, 15, **18**, 18, 19, 25, 27, 30, 31, 35, 45, 70

(Continued)

(Continued)

- Identify the 25th and 75th percentiles and underline the values between them.

  2, 5, 6, 6, **8**, 9, 9, 12, 15, **18**, 18, 19, 25, 27, **30**, 31, 35, 45, 70

- Subtract to find the interquartile range.

  30 – 8 = 22

The interquartile range of time spent by graduate students on assigned reading at home each week is 22 hours.

Ratio and Interval Data

Interquartile range values can be used with interval and ratio data (see the example of GRE scores above), or the degree of score dispersion can be expressed as a statistic. That statistic is named **variance** and can be defined as the average of all squared differences between individual scores and the mean. Thus, to find variance, it is first necessary to find the data's mean (average) and, in turn, subtract the mean from each individual's score $X_i - \bar{X}$. This will produce a series of both positive values from scores greater than the mean, and it will produce equivalent negative values from scores that are less than the mean. The negative and positive values will cancel each other out. In other words, the variation is zero:

$$\sum_{i=1}^{n}\left(X_i - \bar{X}\right) = 0$$

The solution to this conundrum is to make all values positive once they are subtracted from the mean. This is accomplished by multiplying each difference score by itself (squaring it) before adding the squared difference scores together.[14] Negative values when squared become positive. This produces the sum of squared differences from the mean:

$$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \text{Total of summed squared deviations}$$

The last step in finding variance is to divide the squared sum of squared differences by the size of the sample, or the n number of cases included in the distribution of scores being described. For technical reasons, we subtract 1 from n when calculating variance for a sample.

$$\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1} = \text{variance } (s^2, sd^2, \sigma^2)$$ [15]

The symbols used to represent variance are sd^2, s^2, and, in the case of an entire population, the lowercase Greek letter sigma squared (σ^2). In each case, the symbol is a squared value to account for the step taken in calculating the value of variance for the data.[16] There are several free Internet calculators that will find descriptive statistics, including standard deviations (described below). For example, see the following:

http://www.ajdesigner.com/phpstatistics/standard_deviation_population.php

http://www.easycalculation.com/statistics/standard-deviation.php

http://www.mathsisfun.com/data/standard-deviation-calculator.html

http://www.csgnetwork.com/stddeviationcalc.html

In addition, the statistical software maintained on most universities' servers has this capability. Those software packages were listed earlier in this chapter, and an example of IBM-SPSS is given in Appendix A.

The actual values of variance can only be positive or zero. If everyone scores exactly the mean, there is no dispersion, and the value of the variance is zero. All other values for variance must be positive. The problem is that the statistic for variance is not easily interpreted, as it is influenced by the size of the measurement units used to find it. Some variance values may be very high because the variable uses large numbers and reflects considerable individual differences. For example, the level of the enzyme aspartate aminotransferase (AST) is a measure of human liver function. It can be found through standard vein-puncture collection and a serum assay. Table 7.4 presents AST levels for 15 recovering alcoholics enrolled in a 12-step program.

Table 7.4 Serum Levels of Aspartate Aminotransferase

Patient	*Age*	*Serum AST*
Ann	26	60
Bill	42	48
Carol	19	52
Doug	55	27
Erin	22	52
Frank	66	130
Grace	53	67

(Continued)

Table 7.4 (Continued)

Patient	*Age*	*Serum AST*
Harry	36	70
Ida	62	44
Jim	51	120
Kim	38	96
Liam	25	112
Missy	47	86
Neil	20	32
Oprah	58	9
Range	47	121
Interquartile Range	30	52
Median	42	60
Mean	41.3	67.0
Variance	259.4	1256.6
Standard Deviation	16.1	35.4

Variance and Standard Scores

Researchers have a number of available methods for reporting measurement scores based on the framework of variance and the normal curve.

Standard Deviation

For most measurement applications, variance is presented in the form of **standard deviation.** This statistic is the square root of variance.

$$\sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}} = \text{Standard deviation}\,(s, SD, \sigma)$$

The statistic standard deviation is sometimes denoted as *s* or *SD* and, in the case of an entire population, as the lowercase Greek letter sigma σ. The utility of the standard deviation statistic is drawn from its relationship to the Gaussian normal curve. This relationship can be seen in Figure 7.2. When data are collected from a **random sample** of a normal population, a distribution of those scores will approximate the normal curve. The approximate proportion of scores that will lie between the mean and 1 standard deviation above the mean (+1*s*) is a constant 34.13%.

The normal distribution is symmetric, so another 34.13% of all the scores in the sample will lie between the mean and 1 standard deviation below the mean (–1*s*). Thus, just over 68% of all scores lie between –1*s* and +1*s*. By expanding this to almost ±2*s*, it is possible to account for 95% of all scores within a normal sample. This represents 47.5% above and 47.5% below the mean.

It is this consistency of the normal distribution that is the foundation of large-scale score reports. The distribution may be of scores that are based on large numbers (e.g., the NBCC Licensing Examination has a possible 160-point range [0–160]), or small numbers (e.g., undergraduate grade point average has a 4-point range [0.0–4.0]), but the relationship of the standard deviation, *s*, to the distribution will be constant.

z-Scores

Scores from any normal sample can be converted into units of standard deviation. This process makes it possible to compare one set of test scores to the scores from another test or measure. For example, if the undergraduate GPA of an applicant to graduate school is set side by side with the applicant's GRE score, the two seem incompatible. However, both the GPA and GRE can be converted into standard deviation (standard score) format and then compared.

Table 7.5 Mean, Median, and Mode for Each of the Three Groups of Subjects

Sample and Demographic Characteristics for Emerging Adult, Early Adult, and Middle Adult Women			
Variable	*Emerging adult women: 18–25 years old*	*Early adult women: 26–39 years old*	*Middle adult women: 40–65 years old*
n	318	238	245
Average age (years)			
M (SD)	19.47 (1.90)	32.63 (4.06)	51.38 (7.07)
Median	19.00	33.00	52.00
Mode	18.00	36.00	54.00
Ethnic identification: *n* (%)			
White/European American	256 (80.5)	198 (83.1)	206 (84.1)
Black/African American	22 (6.9)	14 (5.9)	14 (5.7)
Asian American	19 (6.0)	8 (3.3)	8 (3.2)

(Continued)

Table 7.5 (Continued)

Variable	*Emerging adult women: 18–25 years old*	*Early adult women: 26–39 years old*	*Middle adult women: 40–65 years old*
Latina or Hispanic	9 (2.8)	4 (1.7)	6(2.4)
Native American	1 (0.3)	4 (1.7)	2 (0.8)
Multiracial	3 (1.0)	3 (1.3)	2 (0.8)
International	2 (0.6)	3 (1.3)	5 (2.0)
Did not report	3 (0.9)	2 (0.8)	1 (0.4)
Socioeconomic identification: *n* (%)			
Working class	39 (12.3)	30 (12.6)	28 (11.4)
Middle class	252 (79.2)	184 (77.3)	194 (79.2)
Upper class	23 (7.2)	24 (10.1)	19 (7.8)
Did not report	4 (1.3)	0 (0.0)	4 (1.6)
Relationship status: %			
Single	89.6	28.5	17.1
Partnered	2.5	7.1	4.5
Married	1.3	59.7	62.9
Divorced or separated	0.3	3.4	9.8
Widowed	0.0	0.0	3.7
Did not report	6.3	1.2	2.0

Source: Augustus-Horvath, C. L., & Tylka, T. L. (2011). The acceptance model of intuitive eating: A comparison of women in emerging adulthood, early adulthood, and middle adulthood. *Journal of Counseling Psychology, 58*(1), 110–125. doi: 10.1037/a0022129

Note: The authors reported the mean, median, and mode for each of the three groups of subjects in the study. Because age is a ratio variable, they also provided the standard deviation for age within each of the three groups. The variables ethnicity and relationship status were also reported. Because those variables are nominal, only a simple percentage of the total group membership was reported for the different levels. The variable socioeconomic status is an ordinal and sequenced in the presentation, but it too is only reported as a percentage of the membership of the group.

The symbol of this standard score is z. The score z expresses any raw score in terms of its location in the distribution using units of standard deviation. If an individual score and the sample mean are exactly the same, the **z-score** will be 0.0. If the score is equal to 1 standard deviation above the mean, the z-score will be +1.0. To find any z-score, it is necessary to know the original (raw) score, sample mean of the raw scores, and the sample's raw score standard deviation. The z-score can be found as follows:

$$z = \frac{X - \bar{X}}{s}$$

In this equation, the value of X is any individual's score. By subtracting the sample mean and dividing that difference by the standard deviation, the score is changed into a z-score (standard score).

As an example, if an applicant to graduate school has a GPA of 3.25 from a sample of applicants with a mean of $\bar{X} = 2.50$ and a standard deviation of 0.50, that student would have a z-score of +1.50. This indicates that this student has a college grade point average that is equivalent to one and a half standard deviation units above the mean of that year's applicant pool. Here is the calculation:

$$z = \frac{3.25 - 2.50}{0.50}$$
$$z = 1.5$$

Percentiles

Percentiles may represent an ordinal transformation of data. The concept behind percentiles is that each percentile represents 1/100th of the data. The first percentile includes all those data points arranged from the lowest score to the point where 1% of the data are included. The **first quartile** is the 25th percentile, the point that cuts off the lowest 25% of the scores in the data set. The second **quintile** is the point that cuts off the lowest 40% of the scores from the data set. The median is the 50th percentile, and the seventh **decile** is a score that cuts off the lowest 70% of the data.

Percentile equivalents are reported for most standardized tests. These scores are not simply from a basic count of cases in an ordinal sequence but rather are based on comparisons to a data set that was collected previously. The consistency of the normal curve makes it possible to interpolate z-scores into percentiles. For example, the score point in the center of the Gaussian normal curve ($z = 0.0$) is the 50th percentile. That point has half of all scores below and above it. The Gaussian normal curve starts and ends at infinity. Because of that, there can never be a zero percentile point with these transformed scores. Likewise, there can never be a score equal to the 100th percentile.

For example, the NBCC examination has 160 questions each year that are used to determine a license candidate's status. However, the NBCC examination presents each test taker with 200 items. The extra 40 are not identified as experimental but are integrated into the examination and later used to build next year's normal curve against which the next group of license candidates will be compared. A small number of the extra 40 items are used year after year to demonstrate the stability of the examination over time. These items are known as **anchor items.** Not all candidates taking this year's test see the same 40 items. Thus, it is possible for the NBCC examiners to find a distribution of all of next year's 160 items.

The created comparison group is known as the norm group or normative reference group. Thus, a raw score on this year's test is compared to how candidates last year did on the 160 items, and a percentile score can be found. In Table 7.6, z-scores are presented, and areas under the normal curve are given. These can be easily read as percentiles.

Table 7.6 Areas (Percentiles) Under the Normal Curve by *z*-Score

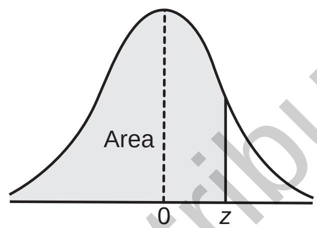


Table 1 The normal curve

(a) Area under the normal curve

Z	*0.00*	*0.01*	*0.02*	*0.03*	*0.04*	*0.05*	*0.06*	*0.07*	*0.08*	*0.09*
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0124	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0085	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0,1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	1.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0,3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0,3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

100 STATISTICAL TESTS

Z	*0.00*	*0.01*	*0.02*	*0.03*	*0.04*	*0.05*	*0.06*	*0.07*	*0.08*	*0.09*
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.56.36	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.662S	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0,9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0 9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	1.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Source: Kanji, G. K. (1999). *100 statistical tests,* pp. 159–160. Reprinted with permission from Sage.

Note: Researchers do not need to use these tables, as the same information is provided online and by statistical software. The actual table is depicted here to illustrate the concept.

From the information in Table 7.6, it can be seen that an undergraduate student who has a GPA equal to a *z*-score of 1.50 would be in the 93rd percentile of a normally distributed set of grade point averages.[17] Likewise, a graduate student who has a score on the NBCC Licensing Examination equal to a *z*-score of –0.7 (the negative sign means below the mean) would be in about the 24th percentile.

Psychologists and counselors are aware of the definition of a cognitively gifted person as having a mental ability score that is 2 or more standard deviations above the mean IQ value of 100. For most tests of cognitive ability, this level equates to an IQ above 129. That IQ score can also be expressed as a *z*-score. To convert the IQ score to a *z*-score, the psychologist needs to know that the published mean of IQ tests is 100 and the standard deviation is 15. The math then would be

$$z = \frac{130 - 100}{15}$$
$$z = 2.0$$

On a table of *z*-scores, it can be seen that a *z*-score of 2.0 occurs around the 97.7th percentile. In other words, only a little over 2% of a normal population (1 out of 50) qualifies as being mentally gifted according to this definition.

In 2007, a total of 25,693 students took the 205-question GRE Subject Test in psychology. The mean score was 600, and the standard deviation was 100 (ETS, 2009). Using Table 7.6, it can be seen that a student scoring 450 on this examination is in approximately the 6th percentile with 94% of all graduates scoring better ($z = -1.5$). A score of 660 ($z = +1.6$) would be at about the 72nd percentile, and a score of 780 ($z = +1.8$) would be in the 96th percentile.[18]

Commonly Reported Standard Scores Based on the Normal Curve

Raw score. A raw score is the total number of actual items on the measure that the test taker answered correctly. To better understand what a raw score indicates, the measure's report should indicate the total number of questions.

Normative percentile. A method of finding any individual's test score percentile is to compare that score to scores of the normative comparison group (i.e., norm group). If a score were at the center point of the norm group, it would be at the 50th percentile, with half of the norm group's scores above and half below. That 50th percentile is also the median of all scores. If a person did better than three quarters of the individuals of the norm group, his or her score would equate to the 75th percentile. Test publishers like ETS report a standard score and the equivalent percentile to the test taker.

Standard scores or scaled scores. There are numerous formats for standard scores (*SS*), sometimes referred to as **scaled scores.** They all are derived mathematically from a huge normative distribution of scores. Under normal conditions, a large distribution of test scores will follow clear mathematical laws. We know that 34% of all scores will be above and 34% below the mean score (arithmetic average score) by a unit of measurement called a standard deviation. Plus or minus 1 standard deviation includes the 68% of all scores closest to the mean (average). A standard score is based on standard deviation units above (positive values) and below (negative values) the mean. To make interpretation and comparisons easier, these scores are mathematically modified to make them all positive and give them recognizable values. One commonly employed standard score is used to describe an outcome of the GRE. Instead of values above and below the mean, GRE scores are mathematically designed to have a mean of about 150 on the core tests. Each standard deviation is set at a value of about 9. Thus, a quantitative GRE score of 163 is 1.5 standard deviations over the mean. Because 34% of all verbal GRE scores are between the mean (GRE = 150) and 1 standard deviation over the mean (GRE = 159), and the average is also the median (50th percentile), adding the two shows that a GRE of 159 is near the 85th percentile (50 + 34).

Undergraduate college admission tests are also standardized scores. The ACT has a national mean (average) of about 21 and a standard deviation of 5. Thus an ACT score of 11 (2 standard deviations below the mean) is at the 2nd percentile. An ACT score of 18.5 (0.5 standard deviations below the mean) is at the 31st percentile, and a score of 28.5 (1.5 standard deviations above the mean) is at the 93rd percentile.

***z*-score.** This score is a real number that has a practical range of about 6, from about −3 to +3, with a zero value indicating the mean of the data. It is an expression of how far any score is from the mean as measured in standard deviation units.

Stanine. The norm group can be divided into nine parts, with the middle seven parts (stanine 2 to 8) each being 0.5 standard deviation wide (see Figure 7.2). This statistic is a core tool in many vocational placement tests, including those used by the American armed services. Because of the characteristic bell-shaped form of a normal distribution of scores, the central three stanines, 4, 5, and 6, encompass 55% of all cases. Stanines 1, 2, and 3 include the lowest 22.5% of all scores, and the top 22.5% of all scores are in stanines 7, 8, and 9. The top stanine, 9, includes the highest 4% of scores, while the lowest stanine, 1, includes the lowest 4% of scores.

Normal curve equivalent (NCE). This is a standard score that has a mean of 50. One standard deviation above the mean (84th percentile) is an NCE score of 71. An NCE score of 29 is 1 standard deviation below the mean, and an NCE score of 10 is the 3rd percentile of the normative distribution of scores.[19]

SUMMARY

Professionals in mental health deal with psychological constructs. Attributes of clients as well as constructs are all variables. Variables can be described by scales, and some scales are more precise than others. The most precise, ratio and interval scales, are described as parametric measures, while most ordinal variables and all nominal variables are nonparametric. Parametric variables are typically presented as continuous line curves. Ordinal variables are generally best depicted using a histogram or frequency polygon. Nominal variables can be presented as pie graphs and bar graphs.

The form of central tendency for data is limited by the precision of the scale used to measure it. Means are reserved for parametric data, the median is used with ordinal data, and the center of nominal data can only be expressed as the mode.

Knowing the mean of parametric data makes it possible for the practitioner-researcher to determine the variance of a data set and its standard deviation. It also is possible to determine whether the symmetry of the normal curve is violated by skewing within the data set. The Gaussian normal curve has many applications, including creating standard scores and finding percentile equivalents to any individual's score on a measure using a parametric scale.

DISCUSSION QUESTIONS

1. In your professional practice, what measures, questionnaires, and other background variables do you plan to collect from prospective clients during intake?
2. On the Internet, search a politically charged topic such as US federal funding for mental health care and examine charts and graphs used to argue various points of view. Bring some of them to class and discuss any errors or distortions you may have found.
3. Develop a list of 10 distributions that are likely to show a high degree of skewness. Try to identify 5 distributions that are likely to show positive skew as well as 5 distributions that are likely to show negative skew.
4. Visit the university's library and read the *Chronicle of Higher Education* issue that lists all university administrator salaries (usually one of the November issues of this weekly paper). Pick a category of administrator and calculate the mean and standard deviation for that position's salary across all universities included in a category of your choice (e.g., *I* for doctoral institutions).

NOTES

1. Students with an interest in these statistical concepts are invited to use the student website for this textbook where there are more computational examples of the various concepts.
2. Congress created the rank "general of the Army" in 1944. It is signified by five stars. Only five senior commanders of the US Army in World War II were ever promoted to this rank.
3. Hospital-based nurses are ranked based on education, experience, and certification tests as nurse technician (NT), certified nurse assistant (CNA), licensed practical nurse (LPN), registered nurse (RN), charge nurse, and supervising nurse.
4. Binary variables can be assigned values of 0 or 1.

5. The Kelvin scale of temperature begins at a point equal to a Celsius scale temperature of −273.15°. At this point, there is no thermal energy (heat) in the phenomenon being measured. This quality of no thermal energy is known as the absolute zero point. At this point, all matter settles into its absolute lowest energy state, and even all subatomic particles stop moving as matter assumes the Bose-Einstein condensate state and produces super fluids as electrons bond (Donley et al., 2001).

6. Statistical power describes the sensitivity of a statistical analysis to detect meaningful differences between groups of subjects on a variable.

7. The word *average* in vernacular English is normally considered as being what a testing expert would call the mean, or arithmetic center, of the set of scores.

8. "Because different test takers receive different sets of questions, . . . ASWB does have to account for differences in the difficulty levels of individual items on different versions of the tests. When a candidate completes an examination, the testing software calculates a raw score—the actual number of questions . . . answered correctly. Because raw scores can be affected by the difficulty of individual items on a particular version ("form") of an examination, these variations are accounted for through an equating process. Equating adjusts the number of items [the test taker needs] to answer correctly up or down depending on the difficulty levels on a particular form (version) of the examination. Through equating, the passing raw score is adjusted for each examination so that fewer correct items are needed to pass a more difficult form of the test (and more correct answers are needed to pass an easier form of the test). Making these statistical adjustments ensures that the overall ability that needs to be demonstrated remains the same from test form to test form. In other words, nobody receives an advantage or disadvantage because of the version of test they receive. This is why ASWB cannot identify an unchanging number of correctly answered items needed to pass the examinations" (from http://www.aswb.org/SWLE/faqs.asp#Curve).

9. Newton was recognized by the government of Charles II as Great Britain's leading scientist and mathematician. Later, in 1705, he was knighted by Queen Anne. Huguenots, including de Moivre, were French Protestants (Calvinists) who were driven out of France in a religious diaspora in the early 18th century. Prior to that forced exile, thousands of Huguenots were massacred by Catholic zealots in the streets of numerous French cities.

10. *Binomial* refers to a probability experiment with only two possible outcomes (e.g., heads vs. tails; yes vs. no; pass vs. fail). A binomial distribution is a chart of those outcomes. As the number of probability experiments increases, the distribution will approximate a bell-shaped curve.

11. The mean for skewness is zero. When the measured level of skewness is equal to or greater than 1.96 times the standard error of skewness above or below zero in the data set, the data can be described as having significant skewness.

12. Real numbers have all the properties required for arithmetic operations. They can be divided and reported as decimals, thus making it possible to interpolate between measurement points. This allows a smooth line curve to be drawn.

13. The term *ogive* was first used by pre-Islamic designers and later by medieval architects to describe the graceful S-shaped arch supports used to hold the massive weight of vaulted stone roofs.

14. When a negative number is squared, it becomes a positive value. I was taught this in eighth grade and have never questioned it since. Likewise, it is not possible to find the square root of a negative number on the number line. The result is an imaginary number, i.

15. The logical divider (n) for the number of cases is replaced by ($n - 1$). This is known as the Bessel correction. It corrects the overestimation in size that occurs when dealing with samples (Warner, 2013). The value n is used when the calculation is based on an entire population (e.g., US Census data), and ($n - 1$) is used with calculations for samples of data. This same principle is seen with sample data where a vector of residuals from original observations will lose a degree of freedom ($n - 1$) when they are recombined to sum to zero. Zero summation around the mean is an axiom of statistics.

16. This is a messy-looking process and can appear to be intimidating to individuals without experience with college mathematics. The good news is that the statistical software used on university campuses today will find this value for any data set entered into the system.

17. Grade inflation on most campuses makes it unlikely to ever find a normal, nonskewed distribution of grade point averages.

18. In the summer of 2011, the scoring system used by the GRE General Test was changed. The GRE Subject Test in psychology (as of 2013) is still unchanged.

19. The least useful of the various standardized test score reporting systems is the **grade equivalent score** (GES). This score is read as a grade level expressed in years of 10 months' length. A GES of 10.6 is February of 10th grade. These scores are ordinals, yet there is a tendency for policy makers to try to find the mean of a set of grade equivalent scores.

The most problematic concern is that most school counselors, and virtually all parents, think that children who are enrolled in the fifth grade and are reported to have a reading GES of 9.1 can read ninth-grade books. This is absolutely wrong. What those data indicate is that the child in fifth grade did as well on the fifth-grade reading test as would the average child in the ninth grade if that ninth grader took the fifth-grade test.