

---

## Chapter 1. Principles of Measurement

### Part 1

---

Here is where we stand at this very first step in the enterprise. Off in the distance is the vision of an elaborate and potentially very useful structure that we want to build. But before we can begin building it, we must first sort out the raw materials, which are the logical and conceptual counterparts of bricks and boards, mortar and nails, trowels and hammers. In the present chapter we begin by sorting out the most essential raw materials of all, which are those that pertain to the elemental process of measurement. At first glance, much of what is presented in this chapter is likely to seem more or less familiar, and some of it might even appear rather obvious and common-sensical. If you are just beginning your study of the subject, however, please bear with it and read the chapter carefully through without skipping or skimming, for there are subtleties and complexities here that have very far-reaching implications.

The watchword for this chapter will be the somewhat ugly but very expressive term **GIGO**, borrowed from the language of computer engineers and programmers. GIGO, which is an acronym for "Garbage In, Garbage Out," is the computer mavin's update of the old adage "You can't make a silk purse out of a sow's ear." Computers are marvelous devices. They can crunch numbers and process information with astonishing speed and accuracy. But no matter how fast or full of multi-megabyte RAM a computer might be, what it gives you by way of output will only be as sensible as what you give it by way of input. If nonsensical garbage is what you put in, then nonsensical garbage is also all that will come out. It will, of course, be very thoroughly processed nonsense. It might even seem elegant and profound. But it is nonsense all the same.

The point is equally applicable to the study of statistics. Statistical methods are basically instruments for processing information. The information that they process is numerical in nature and derives from one or another of several forms of measurement. The various statistical procedures that we will eventually be examining make differing assumptions about the particular kinds of measures you are feeding into them. Feed them the kind of information that they assume they are getting, and the results they give you in return will provide a firm basis for drawing rational conclusions. Feed them information that violates their assumptions, and they will still patiently process it and crank out a result. But that result will be nonsense, and so too will be any conclusion you might draw from it. It might, of course, be very impressive and elegant-looking nonsense, with all the trappings of scientific profundity. But it will be nonsense all the same. Garbage In, Garbage Out. This is why it is so very important that you begin your study of the subject with a clear understanding that there are several different forms of measurement, each with its

own set of properties, strengths, and limitations.

---

## ¶ Fundamental Forms of Measurement: A Preliminary Overview

At the level of fundamentals, all of the forms of measurement that you have ever encountered or are ever likely to encounter come down to a process of either counting, ordering, or sorting. For example:

**Counting.** (i) Take a metal tape measure to your desk, count off the number of inches or centimeters it takes to get from the left edge to the right edge, and you are in effect measuring the desk with respect to the general characteristic of width. (ii) Step dry and naked onto a bathroom scale, count off the number of pounds or kilograms that the instrument registers in result of that action, and you are in effect measuring your body with respect to the general characteristic of weight. ~ ~ Procedures of this general type, based on counting-off the units on some standard measurement scale, constitute what is by far the most commonly understood referent of the term "measurement." The other two types of procedures, based on ordering and sorting, require a somewhat broader understanding of the concept.

**Ordering.** The meaning of the term "ordering" in this context is not "placing an order" but rather "arranging in order." (i) Make a list of the things you recurrently worry about—money, grades, family, friends, personal attractiveness, the degradation of the environment, the meaning of life, or whatever they might be—arranging them in the rank order of the strength of the hold they have upon your mind (strongest, second strongest, third strongest, and so on), and you are in effect measuring each, in relation to all the others, with respect to the general dimension of personal worrisomeness. (ii) A variation on this theme would be to rate each of the worrisome items on, for example, a 5-point scale, with "1" and "5" respectively representing what you judge to be the lowest and highest degrees of worrisomeness, and "2" through "4" representing the gradations of worrisomeness that range between these two extremes.

**Sorting.** The very simplest form of measurement—so simple, indeed, that your first impulse will probably be to resist seeing it as being any kind of "measurement" at all—involves sorting items into categories. Do, however, try to resist that impulse, for this really is a kind of measurement, and a very useful one at that. (i) Sort the individuals who have applied for admission to a certain college into the two categories *female* and *male*, and you are in effect measuring each applicant with respect to the general characteristic of gender. (ii) Sort them into the three categories *admitted*, *not admitted*, and *wait-listed*, and you are in effect measuring each with respect to the general characteristic of admission status. (iii) Perform both of these sortings concurrently, and you are in effect measuring each student with respect to  $2 \times 3 = 6$  possible categories of cross-classification: *female and admitted*, *male and admitted*, *female and not admitted*, *male and not admitted*, and so on.

These three forms are not just different versions of measurement. They are different *types*. Like the bricks, boards, nails, and mortar that might be used in the

construction of a house, they are made of different stuff, with quite different properties and quite different uses to which they can properly be put. It is exceedingly important for the student of statistics to understand the properties of these three forms of measurement, because these properties are among the main determinants not only of the kinds of mathematical and statistical procedures that can be legitimately applied to a set of measures, but also of the kinds of conclusions that can be meaningfully drawn from the application. Before we get into the details of this point, however, we must first make a brief detour to lay out a distinction and some terminology that applies to all three forms of measurement.

### • Variables and Variates

In statistical parlance, the general property that is being measured through one or another of the three processes that we have outlined is spoken of as a **variable**, and any particular measured instance of that property is spoken of as a **variate**:

Thus, for the first example of measurement by counting, in which we take a tape measure to the width of a desk, the variable is desk width and the variate would be the measured width of this, that, or the other particular desk. For the second counting example, in which we step on a bathroom scale and record the number of pounds or kilograms, the variable is body weight, and the variate would be the measured weight of this, that, or the other particular body.

For the example in which you rank-order your recurrent worries, the variable is the relative degree of worrisomeness, as assessed by you subjectively, and each item in the list of worries, according to whether it is ranked as first, second, third, and so on, would be a variate instance of that variable. If you were instead to rate your worries on a 5-point scale, the variable would still be relative degree of worrisomeness, as assessed by you subjectively, though the variates would now be the rating-scale values ("1", "2", "3", etc.) that you assign to each of the particular items on the list.

For the first example of measurement by sorting, the variable is gender, and each applicant sorted as female or male would represent a variate instance of that variable. For the second sorting example, the variable is admission status and each applicant sorted into one or another of the three categories would represent a variate instance of that variable. In the third sorting example there are two variables concurrently, gender and admission status, and each applicant cross-classified as female and admitted, male and admitted, and so on, would represent a **bivariate** instance of both variables together.

The term "variable" implies that the results of the measurement process are capable of *varying* from one time to another or from one item to another. Thus, the categorical measurement of gender among a mixed group of human subjects will vary from one subject to another between the two possible outcomes, *female* and *male*; rank-order measurements of worrisomeness will vary not only according to the specific items that are being ranked, but also according to who is doing the ranking; and measurements of desk width will of course vary from one desk to another. The

opposite of a variable is a **constant**. The classical example of a constant is the value of pi, which is the ratio of the circumference of a circle to its diameter. Precisely measure this ratio with respect to any circle at all (so long as you remain within a Euclidean universe), and it will come out to an absolutely unvarying value of 3.14159... .

If this detour on variables and variates has seemed chiefly a matter of belaboring the meanings of words, please bear in mind that the terminology of a technical subject is no mere decoration. Its function is to provide a way of saying things compactly and efficiently, and the sooner you achieve a mastery of it, the sooner your mind will be free to ascend to higher levels. At any rate, we now conclude the detour and return to the task at hand, which is to describe the properties of the three forms of measurement and examine the implications of these properties. We begin with the process that involves counting, since that is the form of measurement with which you are probably already most familiar. It is also the form of measurement that permits the most powerful forms of mathematical and statistical analysis.

---

### **¶Measurement by Counting: Standard Scalar Measurement, Equal Interval Scales and Ratio Scales**

If you wish to measure the width of your desk, you take a tape measure to it and count off the number of inches or centimeters. If you wish to measure the outdoor temperature at the present moment, you take a thermometer outdoors and count off the number of degrees Fahrenheit or degrees Celsius. If you are sitting in a classroom and wish to measure the number of students in the room, you count them. We will speak of this type of measurement as **standard scalar measurement**, since each individual instance of it results in a numerical value that refers to a point on some particular standard measurement *scale* — inches, centimeters, degrees Fahrenheit, degrees Celsius, pints, liters, bushels, grams, ounces, light years, volts, ohms, and so on. If you count up the number of students in a classroom as one, two, three, four, etc., and run out of students to count after reaching 29, this numerical value of 29 represents a point on the most general and familiar standard measurement scale of all, the scale of cardinal numbers.

When measurement involves simply counting out the number of a set of items or events according to the series of cardinal numbers—one, two, three, four, etc.—the scale of measurement is spoken of as an **absolute scale**. All other commonly recognized measurement scales are *relative* in the sense that they are designed to measure not the absolute number of items or events but rather the magnitude of some particular attribute—length, width, weight, temperature, velocity, electrical potential, etc.—relative to the units of some particular scale that has been designed, or has evolved, for taking the measure of that attribute.

Standard scalar forms of measurement can also be sorted out according to whether they are interconvertible. The desk at which I am presently sitting is 60 inches wide when measured with a yard stick and 152.4 centimeters wide when measured with a meter stick. The actual width of the desk is of course the same in both cases. Inches and centimeters are simply two different and interconvertible ways of measuring it. Each inch is equal to 2.54 centimeters, and each centimeter is equal to 0.3937 inches. Similarly, the outdoor temperature on a pleasant June morning is 61.5 degrees Fahrenheit and 16.4 degrees Celsius. The actual degree of warmness is the same in both cases; it is simply being measured by two different scales. In general, any two scales of measurement that measure the same general property (length, warmness, heaviness, volume, velocity, etc.) and can be systematically translated back and forth into each other's terms are said to be **commensurate**; otherwise they are **incommensurate**. Thus, inches and centimeters are commensurate scales of measurement, as are degrees Fahrenheit and degrees Celsius; whereas inches (or centimeters) and degrees Fahrenheit (or Celsius) are incommensurate scales of measurement.

### • **Equal Interval Scales**

Suppose, now, that we were actually to measure the width of a desk in inches, the current outdoor temperature in degrees Fahrenheit, and the number of students in a classroom by counting them off as one, two, three, four, and so on. Examine these various measures closely and you will see there is one general property, apart from the aspect of counting, that all three have in common. It is that the scales of measurement to which they refer— inches of width, degrees Fahrenheit of temperature, and number of students—are all delineated by **equal intervals** between their successive units of measurement. Thus, the one-inch interval that separates 23 inches from 24 inches is precisely the same size as the one-inch interval that separates 24 from 25 inches, 30 from 31 inches, or 43 from 44 inches. By extension, the three-inch interval that separates 20 from 23 inches is precisely the same size as any other three-inch interval along the scale; further, any three-inch interval is precisely three times as large as any one-inch interval, but only half as large as any six-inch interval; and so on. The same is true for degrees Fahrenheit and student headcount, as well as for virtually any other commonly recognized scale of measurement that comes down at the level of fundamentals to a process of

counting—centimeters, degrees Celsius, volts, ohms, miles per hour, miles per gallon, quarts, bushels, pecks, acres, dollars, square feet, and cubic centimeters, to mention but a few. Measurement scales of this general type are spoken of as **Equal Interval Scales**.

If a measurement scale possesses this property of equal intervals, it is then possible and meaningful to take two or more measures from that scale and perform the simple arithmetic operations of addition and subtraction. For example:

$$5 \text{ inches} + 3 \text{ inches} = 8 \text{ inches}$$

$$45.3^{\circ}\text{F} - 5.0^{\circ}\text{F} = 40.3^{\circ}\text{F}$$

$$5 \text{ students} - 3 \text{ students} + 10 \text{ students} = 12 \text{ students}$$

Because they allow these primary arithmetic operations of addition and subtraction, equal interval scales also permit the calculation of various secondary measures that describe the aggregate properties of what we will be speaking of later as *samples* and *distributions*. An example of such an aggregate measure, with which you are certainly already familiar, is the simple arithmetic average. Suppose that the high temperature on three successive winter days is  $31.7^{\circ}\text{F}$ ,  $36.4^{\circ}\text{F}$ , and  $29.0^{\circ}\text{F}$ . Given that the temperature scale of degrees Fahrenheit has equal intervals, we could then meaningfully conclude that the average high temperature for those three days is

$$\frac{31.7^{\circ}\text{F} + 36.4^{\circ}\text{F} + 29.0^{\circ}\text{F}}{3 \text{ days}} = 32.37^{\circ}\text{F}$$

These, however, would not be legitimate or meaningful operations for measures based on scales that have **unequal** intervals, such as the decibel scale of sound intensity or the Richter scale of earthquake intensity. On the Richter scale, for example, each succeeding unit represents a 10-fold increase in intensity. Thus, a quake that registers 5 on the scale is 10 times stronger than one that registers 4; a quake that registers 6 is 10 times stronger than one that registers 5, thus 100 times stronger than one that registers 4, 1,000 times stronger than one that registers 3; and so on. It would therefore make no sense at all to say that three earthquakes measuring 3, 4, and 8 on the scale have an average magnitude of 5. It is of course certainly possible to calculate the raw average of the numbers 3, 4, and 8, and that average will certainly end up as  $(3+4+8)/3 = 5$ . But it would not mean anything. In order to calculate a rigorously meaningful average for any set of numerical values, it is essential that the numerical values be based on a measurement scale that has equal intervals.

But for now, back to inches, degrees Fahrenheit, and cardinal numbers. Examine these three standard measurement scales more closely and you will find there is one further property possessed by inches and degrees Fahrenheit that is not shared by the headcount of students, and another property possessed by inches and student headcount that is not shared by degrees Fahrenheit. Although these properties will perhaps seem rather obvious, they have implications that are non-obvious and quite far-reaching.

## • Continuous Scales versus Discrete Scales

For inches and degrees Fahrenheit versus student headcount, the distinction is between **continuous scales** of measurement versus **discrete scales** of measurement or, alternatively, between **continuous variables** versus **discrete variables**. If you are counting up the number of students in a classroom it is possible to have, for example, either 16 or 17 students, but it would be utter nonsense to end up with a count of 16.75 students, for there is simply no such thing as three-quarters of a student. The same is true when you count up the number of any other set of items that come only in discrete, indivisible units. In cases of this sort the variable (which is the number of items in the set) is discrete, and so too is the integer scale of cardinal numbers—one, two, three, four, etc.—by which the variable is measured. If you are measuring the width of an object in inches, on the other hand, the possible outcomes are not restricted to integer values such as 12 inches, 13 inches, 14 inches, and so on. It is entirely possible to come out with a measure of width that falls somewhere between two successive integer values: for example, 12.47 inches. The same is true for degrees Fahrenheit, pints, grams, seconds, cycles per second, microfarads, and any number of other delineations that would normally be spoken of as a scale of measurement. The common property of such continuous scales is that their units of measurement are in principle infinitely divisible, so that any particular measure taken on them could potentially be drawn out to as many other decimal places as one might care to take it. In reality, of course, there are often practical limits on the precision of the measurement process. With an ordinary household thermometer, for example, it would not be practically possible to measure the current outdoor temperature as  $61.0034709102^{\circ}\text{F}$ . Nonetheless, the temperature of any particular object or location *could* potentially have this or any other multi-decimal value, if only we had a sufficiently precise instrument for measuring it.

## • Ratio Scales versus Non-Ratio Scales

For inches and student headcount versus degrees Fahrenheit, the distinction is between measurement scales for which the point designated as zero represents an **absolute zero** of the quantity that is being measured, versus measurement scales whose designated zero is only an arbitrary point that happens to be called "zero." Thus, zero inches represents an absolute absence of width. Zero students represents an absolute absence of students. Zero degrees Fahrenheit, however, does not represent an absolute absence of warmth, for it is obviously possible to have measures of temperature that fall *below* zero on the Fahrenheit scale. The same is true of the Celsius temperature scale, whose arbitrary zero merely marks the point at which water freezes under normal barometric pressure. Zero degrees on the Kelvin scale of temperature measurement, on the other hand, *does* mark an absolute absence of temperature, which is to say, an absolute absence of warmth and an absolute absence of the molecular motion from which the quality of warmth derives. In brief, zero degrees Kelvin, which corresponds to about  $-273^{\circ}\text{C}$  and  $-460^{\circ}\text{F}$ , is as cold as it gets, and there is nothing colder.

So if you are ever out in a temperature of  $-18^{\circ}$  Fahrenheit, you can comfort yourself

with the thought that it is still about  $442^{\circ}$  Fahrenheit warmer than it could be. On the other hand, you will never under any circumstances find a temperature of  $-18^{\circ}$  Kelvin, or even  $-0.000,018^{\circ}$  Kelvin, even if you go to the farthest reaches of interstellar space. Scales of measurement that have an absolute zero point can yield negative values, such as  $-18^{\circ}$  Kelvin, only in two special kinds of cases, both involving the concept of *polarity* or *directionality*. [[SideTrip](#)]

Scales of measurement that have both equal intervals *and* absolute zero points are spoken of as **ratio scales**, for the simple reason that they permit the meaningful calculation of ratios. If you find, for example, that object A is 5 inches wide and object B is 15 inches wide, it is legitimate and meaningful to conclude that object B is three times as wide as object A, or alternatively, that object A is only one-third as wide as object B. Similarly, it makes sense to say that 15 students are three times as many as 5 students, and that 5 students are only one-third as many as 15 students. If the high temperatures on two successive winter days are  $5^{\circ}\text{F}$  and  $15^{\circ}\text{F}$ , on the other hand, it makes no sense at all to conclude that the second day is three times as warm as the first—because the zero point from which  $5^{\circ}\text{F}$  and  $15^{\circ}\text{F}$  are starting out is only an arbitrary marker on a scale that potentially extends all the way down to about  $-460^{\circ}\text{F}$ . In order to make such ratio judgments concerning temperatures we would have to use a scale, such as the Kelvin scale, whose zero point does mark an absolute zero level of temperature.

---

### **¶Compound Measures on Equal Interval Scales**

There are two desks in my office. The width of the larger is 60 inches, and the width of the smaller is 48.5 inches. The sum of these two widths is 108.5 inches, their average is 54.25 inches, and the difference between them is either 11.5 inches or  $-11.5$  inches, depending on whether we subtract the smaller number from the larger or the larger from the smaller. Each of these three quantitative facts is also a *measure*, specifically, a compound measure based on two component measures. In science and in human affairs in general, you will find many compound measures of these general types, all owing their existence to the fact that equal interval scales readily permit addition, subtraction, and the calculation of averages. A familiar family of examples includes that whole range of measures that involve the relational word "per"—miles per gallon, miles per hour, dollars per hour, cycles per second, calories per gram, bushels per acre, and so on. Here are some general precepts for determining the scale properties that pertain to such compound measures.

**Sums.** The general principle is that the sum of a set of equal interval measures will have the same scale properties as the component measures on which the sum is based. Thus, the sum of a set of measures expressed in inches is itself a measure of inches, and it will have the same scale properties as the original scale of inches (continuous, equal interval, and ratio). The sum of a set of measures of temperature expressed in degrees Fahrenheit is itself a measure of degrees Fahrenheit, and it will have the same scale properties as the original scale of degrees Fahrenheit (continuous, equal interval, and non-ratio). If you count up the number of students in each of three sections of a statistics course, each count is a measure on the scale of cardinal numbers. The sum of the counts will also be a measure on the scale of cardinal numbers, and it will have the same scale properties (discrete, equal interval, and ratio).

**Averages** The same principle holds for the average of a set of equal interval measures, with one exception. Thus, the average of a set of equal interval ratio scale measures (e.g., inches) will have the properties of a ratio scale, and the average of a set of equal interval non-ratio measures (e.g., degrees Fahrenheit) will have the properties of a non-ratio scale. The exception is that any average of two or more equal interval measures will have the properties of a *continuous* scale, even though the original component measures themselves might be discrete. Thus, if three sections of a statistics course have 27, 31, and 30 students, respectively, the average number of students in these sections is 29.33..., that is, 29 and one-third "students per class." This of course does not mean that some hapless student is divided into three parts; it is simply the way the arithmetic of the situation happens to work out.

**Differences** The principle also holds for the differences between equal interval measures, but again with one exception. Thus, the difference between two ratio measures will also be a ratio measure; the difference between two discrete measures will also be a discrete measure; and the difference between two continuous measures will also be a continuous measure. The exception here is that the difference between two equal interval measures will have the properties of a *ratio* scale, even if the original component measures belong to a non-ratio scale. The reason for this is that it is always possible to end up with a difference between two equal interval measures of absolutely zero, even though the two measures themselves belong to a scale, such as degrees Fahrenheit of temperature, that does not have an absolute zero point. Thus, if you were to measure the temperature at three locations as  $A = 40^\circ\text{F}$ ,  $B = 50^\circ\text{F}$ , and  $C = 60^\circ\text{F}$ , it would of course make no sense to say that B is  $50/40 = 1.25$  times as great as A, nor that C is  $60/50 = 1.2$  times as great as B. However, it would make perfectly good sense to say that the *difference* between C and A ( $60 - 40 = 20$ ) is twice as large as the difference between B and A ( $50 - 40 = 10$ ).

## • A Scorecard for Equal Interval Scales of Measurement

As shown in [Table 1.1](#), this roster of players is not nearly so complex as it might at first seem. For equal interval scales of measurement there are basically two dimensions of classification—ratio versus non-ratio and discrete versus continuous—which

yield a total of four possible cross-classifications:

non-ratio and discrete;  
ratio and discrete;  
non-ratio and continuous; and  
ratio and continuous

And it is really only the last three that you need to keep track of, as it is difficult to imagine any practically useful scales of measurement that could fall into the first cross-classification category of 'non-ratio and discrete.' This is because any scale delineated by discrete equal intervals is basically a scale for counting the number of discrete, indivisible units of some particular type, be they students in a classroom or hydrogen atoms in a cubic meter of interstellar space, and in any such enumeration it is always a logical possibility to end up with a count of absolutely zero. In any event, bear in mind that the more far-reaching distinction is the one between ratio scales and non-ratio scales, which depends upon whether the scale in question does or does not have an absolute zero point. All equal interval scales, whether ratio or non-ratio, permit the mathematical operations of addition and subtraction, as well as the calculation of averages. But statements of *relative magnitude*, such as "A is three times as large as B," or "B is one-third as large as A," are permitted only by equal interval scales that have an absolute zero point.

---

### End of Chapter 1, Part 1.

[Return to Top of Part 1](#)

[Go to Chapter 1, Part 2](#)

[Home](#)

Click this link only if the present page does not appear in a frameset headed by the logo **Concepts and Applications of Inferential Statistics**

---