

Before presenting my central examples of baseball and life, I offer two cases to illustrate my contention that our culture encodes a strong bias either to neglect or ignore variation. We tend to focus instead on measures of central tendency, and as a result we make some terrible mistakes, often with considerable practical import.

• • •

• 4 •

*Case One: A Personal Story*

**Where any measure of central tendency  
acts as a harmful abstraction,  
and variation stands out  
as the only meaningful reality**

In 1982, at age forty, I was diagnosed with abdominal mesothelioma, a rare and “invariably fatal” form of cancer (to cite all official judgments at the time). I was treated and cured by courageous doctors using an experimental method that can now save some patients who discover the disease in an early stage.

The cancer survivors’ movement has spawned an enormous literature of personal testimony and self-help. I value these books, and learned much from them during my own ordeal. Yet, although I am a writer by trade, and although no experience could possibly be more intense than a long

fight against a painful and supposedly incurable disease, I have never felt any urge or need to describe my personal experiences in prose. Instead, as an intensely private person, I view such an enterprise with horror. In all the years then and since, I have been moved to write only one short article about this cardinal portion of my life.

I accept and try to follow the important moral imperative that blessings must be returned with efforts of potential use to others. I am therefore enormously grateful that this article has been of value to people, and that so many readers have requested copies for themselves, or for a friend with cancer. But I did not write my article either from compulsion (as a personal testimony) or from obligation (to the moral requirement cited above). I wrote my piece, *The Median Is Not the Message*, from a different sort of intellectual need. I believe that the fallacy of reified variation—or failure to consider the “full house” of all cases—plunges us into serious error again and again: my battle against cancer had begun with a fine example of practical benefits to be gained by avoiding such an error, and I could not resist an urge to share the yarn.

We have come a long way from the bad old days, when cancer diagnoses were scrupulously hidden from most patients—both for the lamentable reason that many doctors regarded deception as a preferred pathway for maintaining control, and on the compassionate (if misguided) assumption that most people could not tolerate a word that conveyed ultimate horror and a sentence of death. But we cannot overcome obstacles with ignorance: consider what Franklin D. Roosevelt could have contributed to our understanding of disability if he had not hidden his paralysis with such cunning care, but had announced instead that he did not govern with his legs.

American doctors, particularly in intellectual centers like Boston, now follow what I regard as the best strategy for this most difficult subject: any information, no matter how brutal, will be given upon request (as compassionately and diplomatically as possible, of course); if you don't want to know, don't ask. My own doctor made only one departure from this sensible rule—and I forgave her immediately as soon as I faced the context. At our first meeting, after my initial surgery, I asked her what I could read to learn more about mesothelioma (for I had never heard of the disease). She replied that the literature contained nothing worth pursuing. But try-

ing to keep an intellectual from books is about as effective as that old saw about ordering someone not to think about a rhinoceros. As soon as I could walk, I staggered over to the medical school library and punched *mesothelioma* into the computer search program. Half an hour later, surrounded by the latest articles, I understood why my doctor had erred on the side of limited information.

All the literature contained the same brutal message: mesothelioma is incurable, with a median mortality of eight months following diagnosis. A hot topic of late, expressed most notably in Bernie Siegel's best-selling books, has emphasized the role of positive attitude in combating such serious diseases as cancer. From the depths of my skeptical and rationalist soul, I ask the Lord to protect me from California touchie-feeliedom. I must, nonetheless, express my concurrence with Siegel's important theme, though I hasten to express two vital caveats. First, I harbor no mystical notions about the potential value of mental calm and tenacity. We do not know the reasons, but I am confident that explanations will fall within the purview of scientific accessibility (and will probably center on how the biochemistry of thought and emotion feed back upon the immune system). Second, we must stand resolutely against an unintended cruelty of the “positive attitude” movement—insidious slippage into a rhetoric of blame for those who cannot overcome their personal despair and call up positivity from some internal depth. We build our personalities laboriously and through many years, and we cannot order fundamental changes just because we might value their utility: no button reading “positive attitude” protrudes from our hearts, and no finger can coerce positivity into immediate action by a single and painless pressing. How dare we blame someone for the long-standing constitution of their tendencies and temperament if, in an uninvited and unwelcome episode of life, another persona might have coped better? If a man dies of cancer in fear and despair, then cry for his pain and celebrate his life. The other man, who fought like hell and laughed to the end, but also died, may have had an easier time in his final months, but took his leave with no more humanity.

My own reaction to reading this chillingly pessimistic literature taught me something that I had suspected, but had not understood for certain about myself (for we cannot really know until circumstances compel an ultimate test): I do have a sanguine temperament and a positive attitude.

I confess that I did sit stunned for a few minutes, but my next reaction was a broad smile as understanding dawned: "Oh, so that's why she told me not to read any of the literature!" (My doctor later apologized, explaining that she had erred on the side of caution because she didn't yet know me. She said that if she had been able to gauge my reaction better, she would have photocopied all the reprints and brought them to my bedside the next day.)

My initial burst of positivity amounted to little more than an emotional gut reaction—and would have endured for only a short time, had I not been able to bolster the feeling with a genuine reason for optimism based upon better analysis of papers that seemed so brutally pessimistic. (If I had read deeply and concluded that I must inevitably die eight months hence, I doubt that any internal state could have conquered grief.) I was able to make such an analysis because my statistical training, and my knowledge of natural history, had taught me to treat variation as a basic reality, and to be wary of averages—which are, after all, abstract measures applicable to no single person, and often largely irrelevant to individual cases. In other words, the theme of this book—"full house," or the need to focus upon *variation within entire systems*, and not always upon abstract measures of average or central tendency—provided substantial solace in my time of greatest need. Let no one ever say that knowledge and learning are frivolous baubles of academic sterility, and that only feelings can serve us in times of personal stress.

I started to think about the data, and the crucial verdict of "eight months' median mortality" as soon as my brain started functioning again after the initial shock. And I followed my training as an evolutionary biologist. Just what does "eight months median mortality" signify? Here we encounter the philosophical error and dilemma that motivated this book. Most people view averages as basic reality and variation as a device for calculating a meaningful measure of central tendency. In this Platonic world, "eight months' median mortality" can only signify: "I will most probably be dead in eight months"—about the most chilling diagnosis anyone could ever read.

But we make a serious mistake if we view a measure of central tendency as the most likely outcome for any single individual—though most of us commit this error all the time. Central tendency is an abstraction, vari-

ation the reality. We must first ask what "median" mortality signifies. A median is the third major measure of central tendency. (I discussed the other two in the last chapter—the mean, or average obtained by adding all the values and dividing by the number of cases; and the mode, or most common value.) The median, as etymology proclaims, is the halfway point in a graded array of values. In any population, half the individuals will be below the median, and half above. If, say, in a group of five children, one has a penny, one a dime, one a quarter, one a dollar, and one ten dollars, then the kid with the quarter is the median, since two have more money and two less. (Note that means and medians are not equal in this case. The mean wealth of \$2.27—the total cash of \$11.36 divided by five—lies between the fourth and fifth kids, for the tycoon with ten bucks overbalances all the paupers.) We favor medians in such cases, when extension at one end of the variation drags the mean so far in that direction. For mortality in mesothelioma and other diseases, we generally favor the median as a measure of central tendency because we want to know the halfway point in a series of similar outcomes graded in time. A higher mean might seem misleading in the case of mesothelioma because one or two people living a long time (the analog of the kid with ten bucks) might drag the mean to the right and convey a false impression that most people with the disease will live for more than eight months—whereas the median correctly informs us that half the afflicted population dies within eight months of diagnosis.

We now come to the crux of practice: I am not a measure of central tendency, either mean or median. I am one single human being with mesothelioma, and I want a best assessment of *my own* chances—for I have personal decisions to make, and my business cannot be dictated by abstract averages. I need to place myself in the most probable region of the variation based upon particulars of my own case; I must not simply assume that my personal fate will correspond to some measure of central tendency.

I then had the key insight that proved so life-affirming at such a crucial moment. I started to think about the variation and reasoned that the distribution of deaths must be strongly "right skewed" in statistical parlance—that is, asymmetrically extended around a chosen measure of central tendency, with a much wider spread to the right than to the left. After all, there just isn't much room between the absolute minimum value of

zero (dropping dead at the moment of diagnosis) and the median value of eight months. Half the variation must be scrunched up into this left half of the curve (see Figure 4) between the minimum and the median. But the right half may, in principle, extend out forever, or at least into extreme old age. (Statisticians refer to the ends of such distribution as “tails”—so I am saying that the left tail abuts a wall at zero survivorship, while the right tail has no necessary limit but the maximal human life span.)

I needed, above all, to know the form and expanse of variation, and my most probable position within the spread. I realized that all factors favored a potential location on the right tail—I was young, rarin’ to fight the bastard, located in a city offering the best possible medical treatment, blessed with a supportive family, and lucky that my disease had been discovered relatively early in its course. I was therefore far more interested in the right tail (my probable residence) than in any measure of central tendency (an abstraction with no special relevance to my case). What, then, could possibly be more uplifting than an inference that the spread of variation would be strongly right skewed? I then checked the data and confirmed my supposition: the variation was markedly right skewed, with a few people living a long time. I saw no reason why I shouldn’t be able to reside among these inhabitants of the right tail.

This insight gave me no guarantee of normal longevity, but at least I had obtained that most precious of all gifts at a crucial moment: the prospect of substantial time—to think, to plan, and to fight. I would not immediately have to follow Isaiah’s injunction to King Hezekiah: “Set thine house in order: for thou shalt die, and not live.” I had made a good statistical inference about the importance of variation and the limited utility of averages, and I had been able to confirm this suspicion with actual data. I had used knowledge and gained succor. (This story boasts an even more favorable outcome. I was destined for the right tail anyway, but an experimental treatment worked and has now probably removed the disease entirely. Old distributions offer no predictions for new situations. I trust that I am now headed for the right tail of a new distribution based on this successful treatment: death at a ripe old age in two high figures—maybe even three low ones.)

I present this tale not only for the pleasure of retelling a crucial yarn about my life, but because it encapsulates all the principles that form the

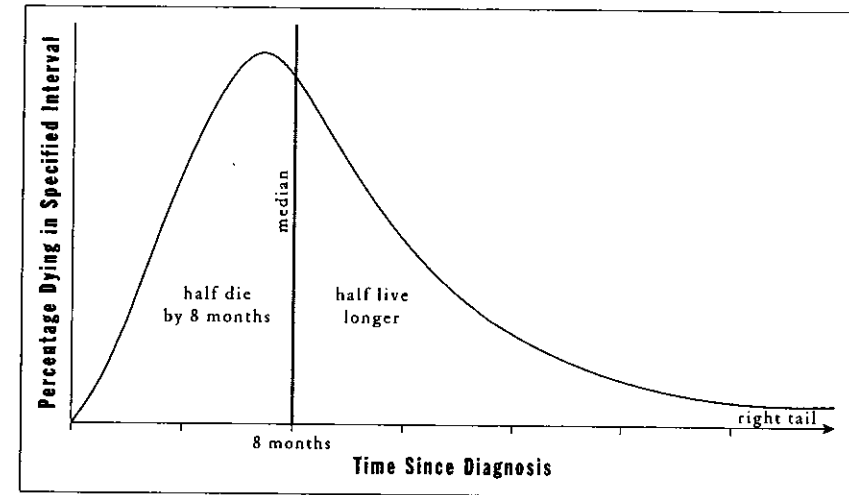


FIGURE 4

A right-skewed distribution for time of death for an illness with a median mortality of eight months. Each individual must be considered as a separate entity and the entire distribution cannot be characterized by its median value.

core of this book. First of all, my story illustrates the importance of *variation within whole systems* as an ultimate reality—and the limited utility (and abstract nature) of averages. Moreover, in a didactic sense for this book, my story embodies the three terms and concepts that form the conceptual apparatus for all the examples to follow. Let me try, then, to present these principles in a formal way, and in a context that will not seem too dry or forbidding.

**THE SKEW OF A DISTRIBUTION.** If we decide to treat variation as a principal reality, then we must consider the standard terms and pictures for portraying populations and their spread. We all know the conventional icon, called a frequency distribution, with the horizontal axis scaled as a graded series for the measure under consideration (height, weight, age, survivorship in disease, batting average, anatomical complexity, etc.), and the vertical axis scaled for the number of individuals in each interval of horizontal values (those weighing between ten and twenty pounds, between twenty and thirty, etc.; those between ten and fifteen years of age, between fifteen and twenty, etc.). Frequency distributions may be symmetrical—that is, with an identical shape and number on either side of the central

tendency. The ubiquitous and idealized “normal distribution” or “bell curve” of current notoriety (Figure 5) is defined as symmetrical in this manner. We have all seen normal curves so often that we have been subtly led to treat natural systems as though they longed to display this ideal form. But most actual populations are not so simple or tidy. (Systems with purely random variation around a mean value will be symmetrical—as variation falls with equal probability on either side of the mean, with any single case more likely to lie close to the mean than far away. Runs of heads or tails in coin tossing, for example, form normal distributions. We regard the normal distribution as canonical because we tend to view systems as having idealized “correct” values, with random variation on either side—another consequence of lingering Platonism. But nature does not match our expectations very often.)

Actual distributions are often asymmetrical, or skewed. In a skewed distribution, as illustrated by my personal story, variation stretches out farther on one side than the other—called either “right” or “left” skewed de-

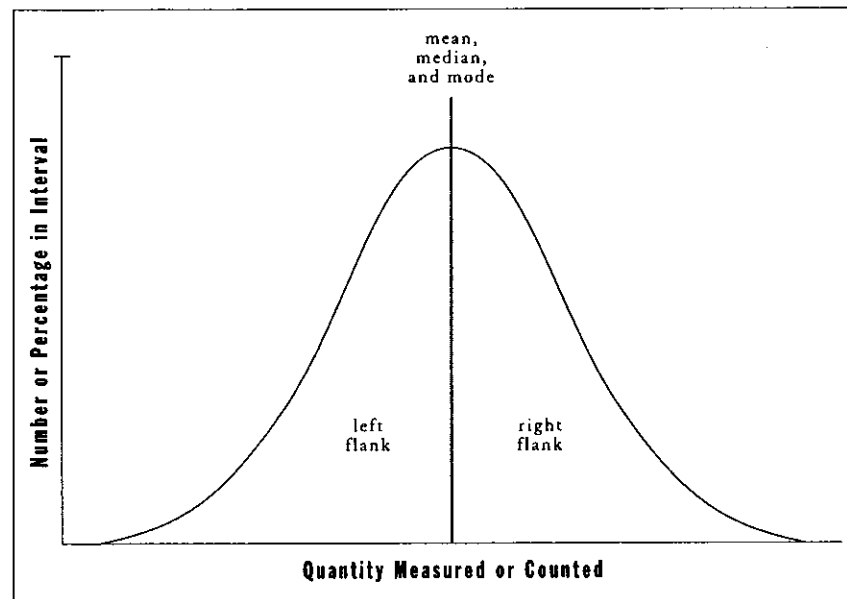


FIGURE 5

An idealized bell curve or normal frequency distribution, showing that all measures of central tendency (mean, median, and mode) coincide.

pending on the direction of elongation (Figure 6). The reasons for skewing are often fascinating and full of insight about the nature of systems—for skewing measures departure from randomness. Since this book treats the nature of variation, and the reasons for changes in spread through time, skewing becomes an important principle in all my examples.

MEASURES OF CENTRAL TENDENCY AND THEIR MEANING. I have discussed the three standard measures of central tendency, or “average” value—the mean (or conventional average calculated by adding all values and dividing by the number of cases), the median (or halfway point), and the mode (or most common value). In symmetrical distributions, all three measures coincide—for the center is, simultaneously, the most common

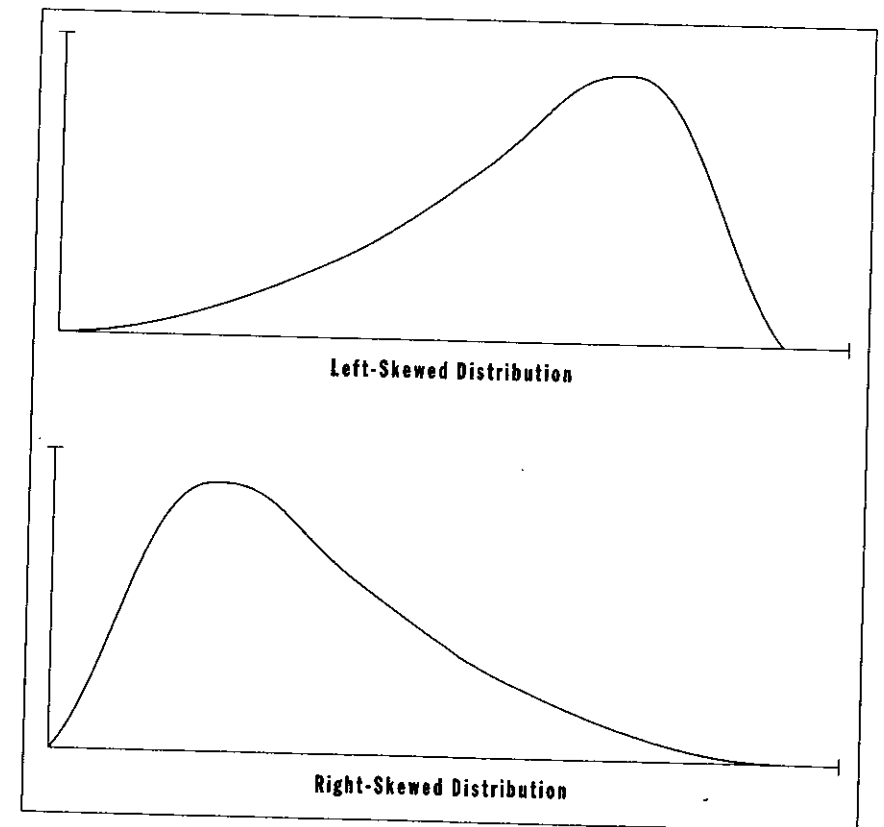


FIGURE 6

Left- and right-skewed distributions.

value, the halfway point (with equal numbers of cases on either side), and the mean. This coincidence, I suspect, has led most of us to ignore the vital differences among these measures, for we view “normal curves” as, well, normal—and regard skewed distributions (if we grasp the principle at all) as peculiar and probably rare. But measures of central tendency differ in skewed distributions—and a major source of employment for economic and political “spin doctors” lies in knowing which measure to choose as the best propaganda for the honchos who hired your gun.

I have already shown how the higher mean and lower mode of a right-skewed distribution in incomes can be so exploited (see page 37). In general, when a distribution is prominently skewed, mean values will be pulled most strongly in the direction of skew, medians less, and modes not at all. Thus, in right-skewed distributions, means generally have higher values than medians, and medians higher than modes. Figure 7 should make these relationships clear. If we start with a symmetrical distribution (with equal mean, median, and mode), and then pull the variation to form a right-skewed distribution, the mean will change most in the direction of skew—for one new millionaire on the right tail can balance hundreds of indigent people on the left tail. The median changes less, for a single pauper will now compensate the millionaire when we are only counting noses on either side of a central tendency. (The median might not move at all if only the wealth, and not the number, of people increases on the right side of the distribution. But if the number of wealthy people at the right tail increases as well, then the median will also shift to the right—but not so far as the mean.) The mode, meanwhile, may well stay put and not vary at all, as mean and median grow in an increasingly right-skewed distribution. Twenty thousand per year may remain the most common income, even while the number of wealthy people constantly increases.

“WALLS,” OR LIMITS TO THE SPREAD OF VARIATION. As a major reason for skew, variation is often limited in the extent of potential spread in one direction (but much freer to expand in the other). The reasons for such limits may be trivial or logical—as in my cancer story where a person can’t die of mesothelioma before he gets the disease, and zero time between onset and death therefore becomes an irreducible minimum. The reasons may also be subtle and more interesting—as in the examples of batting averages and life’s history to be presented in Parts Three and Four of this book.

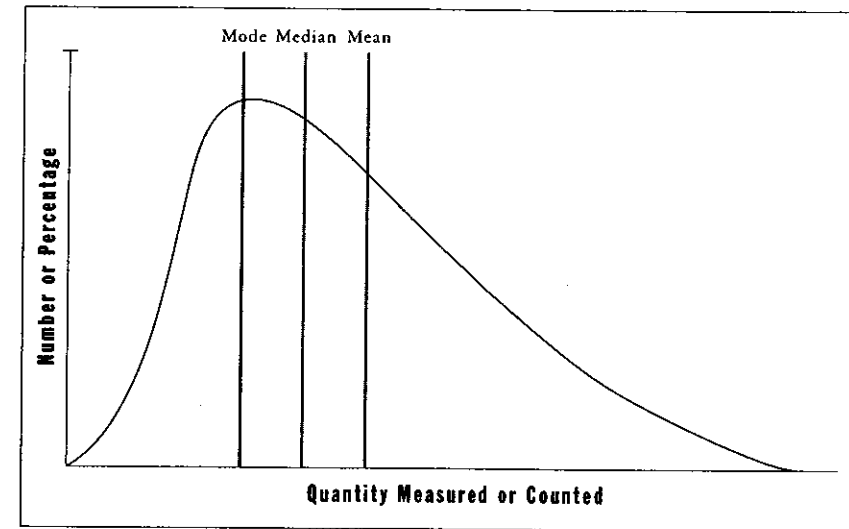


FIGURE 7

In a right-skewed distribution, measures of central tendency do not coincide. The median lies to the right of the mode, and the mean lies to the right of both other measures.

In either case, such limits often produce skewed distributions, because variation can expand in only one direction—you can’t die of mesothelioma before you get it, but you can live for years and years after a diagnosis. With an eight-month median mortality, and a rigid lower limit at time zero, how could the distribution of deaths be anything but strongly right skewed?

Throughout this book, I shall refer to such limits upon the spread of variation as “walls”—either “right walls” or “left walls” depending upon their position. Left walls induce right-skewed distributions (for variation is only free to expand away from a wall); right walls provoke left-skewed distributions. The left wall of my cancer story leads to a right-skewed distribution of deaths.

(I have considered the cultural bias involved in the largely arbitrary designation of right as the direction for higher values, left for lower—though, depending upon the example, lower may be judged better, as in distributions for weight in our diet-conscious society. I suppose that we fall into this bias for two reasons, one insidious and the other benign. Prejudice against our left-handed minority—an old and probably universal feature of human cultures, I fear—must set the major reason. Jesus sits *ad*

## FULL HOUSE

*dextram patris*, at the right hand of the father. Right, etymologically, is dextrous—and “law” is *droit* in French and *Recht* in German, both meaning right. Left is both sinister and gauche. For the benign reason, we read from left to right and therefore conceptualize growth and increase in this direction. Were I writing this book in Israel, which also has a right-handed majority, I might think of left walls as directions of increase. Were I writing in Japan, I might be talking of top and bottom walls. So be it.)

Readers need to grasp only these three nontaxing concepts about the nature of variation in order to render all the examples of this book fully digestible—right and left walls as limits to the spread of variation; right- and left-skewed distributions arising as results of these limits; and differences among means, medians, and modes as measures of central tendency.

.5.

### *Case Two: Life's Little Joke*

**Genuine changes in central tendency are meaningful, but our failure to consider variation has led to a backwards interpretation: the evolution of horses**

**T**he most erroneous stories are those we think we know best—and therefore never scrutinize or question. Ask anyone to name the most familiar of all evolutionary series and you will almost surely receive, as an answer: horses, of course. The phyletic racecourse from small, many-toed protohorses with the charming name *eohippus*, to a big, single-toed Clydesdale hauling the Budweiser truck, or Man O' War thundering down the stretch, must be the most pervasive of all evolutionary icons. Does any major museum not have a linear series of cases against a long wall, or