Menu

Table of Contents

It is not uncommon for Chambers of Commerce to report the mean annual temperature without reporting variability. A city with an "ideal" annual temperature of, say, 74° may go below zero in winter and above 100° in summer. The term "ideal" becomes a misleading statistical abstraction.

Sir Francis Galton Discovers Variability

There are two important uses for measures of variability: (1) describing the spread of distributions, and (2) describing an individual's standing within a group independent of the units of measurement. z-scores could tell us that a person's height is 1 standard deviation above the mean and his weight 0.5 standard deviations above the mean without mentioning inches or pounds. These ideas may seem obvious and indispensable now, but this was not always so.

While the concept of average originated at some uncertain period in antiquity, the idea of a variability measure has a much shorter history. During the first part of the nineteenth century, mathematicians described dispersion by probable error and the interquartile deviation.[1] The science of statistics was well underway at the time, but statisticians were slow to appreciate the utility of these mathematical curiosities. The idea of using variability indices, as we use them here, originated primarily with the English biologist, anthropologist, and psychologist Sir Francis Galton (1822-1911).

Galton's overriding interest was in uncovering the laws of inheritance; to this end he spent an active lifetime collecting and analyzing massive numbers of measurements on all sorts of phenomena. Sweet peas, race horses, dogs, and all classes of Englishmen were brought to his laboratory and measured in dozens of ways. When available statistical methods of his day would not suffice, he invented his own. The concepts of correlation and linear regression were developed by Galton as means of describing the relation between characteristics of parents and their offspring. He also pioneered the use of the normal curve in analyzing frequency distributions of biological and psychological characteristics. All these contributions, in turn, would not have been possible had he not brought variability into the laboratory.

> It is hard to believe now, but the idea of describing variability in groups was quite foreign to Galton's contemporaries. In 1890 he wrote:
> It seems to be a great loss of opportunity when, after observations have been laboriously collected, and been subsequently discussed in order to obtain mean values from them, that the small amount of extra trouble is not taken, which would determine other values whereby to express the variety of all the individuals in those groups.. . . . There are numerous problems of special interest to anthropologists that deal solely with variety.
>
> There can be little doubt that most persons fail to have adequate conceptions of the orderliness of variability, and think it is useless to pay scientific attention to variety, as being in their view, a subject wholly beyond the powers of definition. They forget that what is confessedly undefined in the individual may be definite in the group, and that uncertainty as regards the one is in no way incompatible with statistical assurance as regards the other . . . .
>
> Greater interest is attached to individuals who occupy positions towards the middle of a marshalled series than to those who stand about its middle. An average man is morally and intellectually an uninteresting being.[2]

Just as important, moreover, was his realization that distributions of different variables can be compared to one another if the observations are specified in units of variability. Galton's term for measures of variability was "statistical units," whose "office", he said,

> is to make the variabilities of totally different classes, such as horses, men, mice, plants, proficiency in classics, etc., comparable on equal terms. The statistical unit of each series is derived from the series itself.[3]

He no doubt recognized that his discovery of a use for measures of variability was of major consequence, and he described the moment of insight with a scene that brings to mind Newton's apple tree:

> As these lines are being written, the circumstances under which I first clearly grasped the important generalization that the laws of Heredity were solely concerned with deviations expressed in statistical units, are vividly recalled to my memory. It was in the grounds of Naworth Castle, where an invitation had been given to ramble freely. A temporary shower drove me to seek refuge in a reddish recess in the rock by the side of the pathway. There the idea flashed across me, and I forgot everything else for a moment in my great delight.[4]

Applications of Galton's "important generalization" go far beyond z-scores and t-scores. Without the ability to measure variability, and to use that index to specify an individual's standing in a group, we would have to end statistics at this point. Correlation, regression, and most inferential statistics require a means of specifying an individual's place in a group independently of the original units of measurement.

[1]The probable error is very similar to our modern standard deviation; the standard deviation was brought into general use by Karl Pearson in the 1890's.
[2]From Galton's Anthropometric Laboratory, Notes and Memoirs, No. 1, quoted in Karl Pearson, The Life, Letters and Labours of Francis Galton (Cambridge: Cambridge University Press, 1924), II, 384-385.
[3]Sir Francis Galton, Memories of My Life, 3d ed. (London: Methuen, 1909), p. 298.
[4]Ibid., p. 300.

Understanding and Using Statistics-Basic Concepts, Marty J. Schmidt, pp. 116-117.

The Normal Curve