

Statistics at Square One

III—Standard deviation

T D V SWINSCOW

British Medical Journal, 1976, 1, 1393-1394

In addition to knowing the mean value of a series of measurements it is often informative to have some idea of their range about the mean. For example, the measurements of the urinary concentration of lead that Dr Green obtained for 15 children (Part I) ranged from 0.1 to 3.2 $\mu\text{mol}/24\text{ h}$, with a mean of 1.5. When he extended his study to 140 children the range was from 0.1 to 4.2, with a mean of 2.18 $\mu\text{mol}/24\text{ h}$.

The range is an important measurement, sometimes overlooked in the rush to calculate a standard deviation. The figures at the top or the bottom of it may be exceptionally important, for they denote the findings furthest removed from the generality. However, they do not give much indication of the spread of the observations about the mean. This is where the standard deviation comes in.

The theoretical basis of the standard deviation is complex and need not trouble the ordinary user of it except in one particular. This is that, whether the calculation is done on the whole "population" of data or on a sample drawn from it, the population itself should at least approximately fall into a so-called "normal" (or Gaussian) distribution. When it does so the standard deviation provides a useful basis for interpreting the data in terms of probability. If the population is not normally distributed, the standard deviation cannot be used in this way. (Some discussion of "populations" and "samples" will appear later.) When expressed graphically it appears symmetrically bell shaped, with most observations clustered round the mean and fewest scattered in the tails on each side of the curve. Many biological characteristics conform to it closely enough for it to be commonly used—for example, heights of adult men and women, blood pressures in a healthy population, random errors in many types of laboratory measurements and of biochemical data. Fig 3.1 shows a normal curve calculated from the diastolic blood pressures of 500 men, mean 80 mm Hg, standard deviation 10 mm Hg. The ranges representing $\pm 1\text{ SD}$, $\pm 2\text{ SD}$, and $\pm 3\text{ SD}$ about the mean are marked.

The reason why the standard deviation is such a useful measure of the scatter of the observations is this: if the observations follow a "normal" distribution, a range covered by one standard deviation above the mean and one standard deviation below it ($\bar{x} \pm 1\text{ SD}$) includes about 68% of the observations, a range of 2 standard deviations above and 2 below ($\bar{x} \pm 2\text{ SD}$) about 95% of the observations, and of 3 standard deviations above and 3 below ($\bar{x} \pm 3\text{ SD}$) about 99.73% of the observations. Consequently if we know the mean and standard deviation of a set of observations, we can obtain some useful information by simple arithmetic. By putting 1, 2, or 3 standard deviations above and below the mean we can estimate the ranges that would be

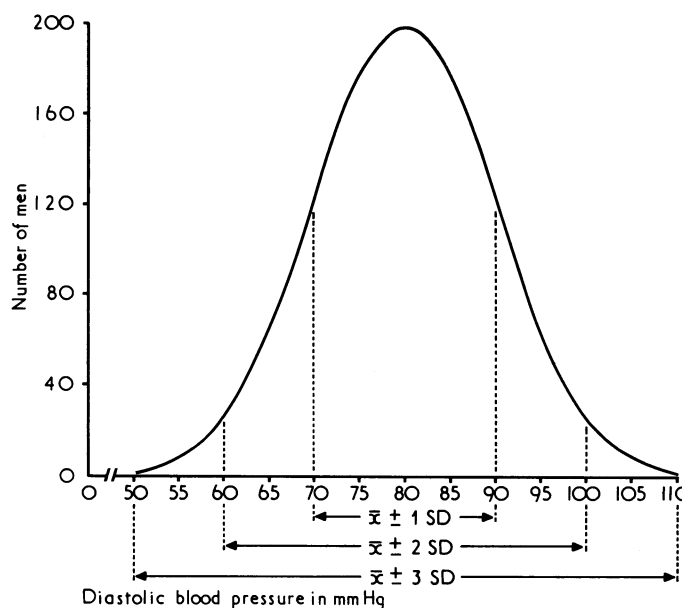


FIG 3.1—Normal curve calculated from diastolic blood pressures of 500 men, mean 80 mm Hg, standard deviation 10 mm Hg.

expected to include about 68%, 95%, and 99.7% of the observations.

Calculation of the standard deviation

The standard deviation is a summary measure of the differences of each observation from the mean. If the differences themselves were added up, the positive would exactly balance the negative and so their sum would be 0. Consequently the squares of the differences are added. The sum of the squares is then divided by the number of observations minus one to give the mean of the squares, and the square root is taken to bring the measurements back to what we started with. (The division by the number of observations *minus one* instead of the number of observations itself to obtain the mean square is because "degrees of freedom" must be used. In these circumstances they are one less than the total. The theoretical justification for this need not trouble the user in practice.)

This procedure is now illustrated from table 3.1 with the 15 readings obtained by Dr Green in his preliminary study of urinary lead concentrations. The readings are set out in col (1). In col (2) is recorded the difference between each reading and the mean. The sum of the differences is 0. In col (3) the differences are squared, and the sum of those squares is at the bottom of the column.

The sum of the squares of the differences (or deviations) from

the mean, 9.96, is now divided by the total number of observations minus one, to give the variance. Thus, the variance = $\frac{\sum(x-\bar{x})^2}{n-1}$.

Finally, the square root of the variance provides the standard deviation:

$$SD = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

This procedure illustrates the structure of the standard deviation, but in practice it is calculated less laboriously. Finding all the deviations of the observations from the mean, as in table 3.1, col (2), can be bypassed.

TABLE 3.1—Calculation of standard deviation

(1) Lead concentration, $\mu\text{mol}/24\text{ h}$ x	(2) Differences from mean $x-\bar{x}$	(3) Differences squared $(x-\bar{x})^2$	(4) Observations in col (1) squared x^2
0.1	-1.4	1.96	0.01
0.4	-1.1	1.21	0.16
0.6	-0.9	0.81	0.36
0.8	-0.7	0.49	0.64
1.1	-0.4	0.16	1.21
1.2	-0.3	0.09	1.44
1.3	-0.2	0.04	1.69
1.5	0	0	2.25
1.7	0.2	0.04	2.89
1.9	0.4	0.16	3.61
1.9	0.4	0.16	3.61
2.0	0.5	0.25	4.00
2.2	0.7	0.49	4.84
2.6	1.1	1.21	6.76
3.2	1.7	2.89	10.24
Total 22.5	0	9.96	43.71

$n = 15, \bar{x} = 1.5$.

If the number of observations is not too many—say, up to 100—they are set out in one or more columns as shown in table 3.1, col (1). Their sum is put at the bottom and their mean calculated as before. Each observation is then squared, as shown in table 3.1, col (4), and the sum of the squares set down (= 43.71). We now square the sum of the observations (foot of col (1)), divide that by the number of observations, and subtract the result from the sum of the squares of the observations (foot of col (4)). Dr Green's figures are treated as follows:

$$43.71 - \frac{22.5^2}{15}. \text{ This equals } 9.96.$$

The figure thus obtained is the same as the 9.96 at the foot of the table 3.1, col (3). The reason for this is that

$$\sum x^2 - \frac{(\sum x)^2}{n} = \sum (x - \bar{x})^2.$$

We now find the variance by dividing 9.96 by 14 (which is $n - 1$), and so obtain 0.714. The square root of this is the standard deviation, 0.84.

The procedure may be summarised as follows:

- Tabulate the observed figures in a column and add them $\sum x$
 - Square this total $(\sum x)^2$
 - Divide by the number of observations $\frac{(\sum x)^2}{n}$ (1)
 - Square each observed figure x^2
 - Add the squares $\sum x^2$ (2)
 - Subtract (1) from (2) $\sum x^2 - \frac{(\sum x)^2}{n}$
 - Divide by the number of observations minus one $\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$
 - Take the square root $\sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$
- This is the standard deviation.

A calculator with a memory and keys for squares and square roots makes light work of this procedure. All that need be set on paper is the column of figures. Each is squared in turn on the calculator, and the squares are accumulated in the memory. The sum of the figures is then added up on the calculator, squared, and divided by the number of observations. This total is subtracted from the sum of squares in the memory. The resulting difference is extracted from the memory on to the display screen, and divided by the number of observations minus one. The square root then gives the standard deviation.

Exercise 3. Dr Green obtained a further series of lead concentrations in urine as follows: 0.2, 0.8, 3.1, 1.9, 0.3, 1.8, 1.7, 1.5, 3.4, 2.0, 2.1, 0.6, 1.9, 2.8, 2.0, 0.7 $\mu\text{mol}/24\text{ h}$. What is their mean and standard deviation?
Answer: Mean = 1.675, SD = 0.96.

How does the food value of beef protein compare with that of soya bean protein?

The food (biological) value of protein depends on its content of the essential amino-acids, and the extent to which these are present in the proportion required by man. The protein with the highest biological value is egg protein, and this is given an empirical score of 100. Beef then has a score of about 80, and the best soya protein a score almost as high at about 70. But the process of preparing the dry powder may decrease the value of soya to as low as 30. In practice, these values for protein from individual foods usually do not mean very much, since other foods in the diet will contain protein that supplements the essential amino-acids—for example, the limiting amino-acids in both beef protein and soya protein are the sulphur-containing amino-acids cystine and methionine, which are plentifully supplied by the proteins in bread; on the other hand, bread is short of lysine, but this is plentiful in both beef and soya.

Do the newer automatic sphygmomanometers have any advantages over the traditional ones?

The ideal sphygmomanometer should be cheap, light, robust, and accurate; it should also eliminate or reduce the well-known tendency

for observer error and bias. The traditional mercury sphygmomanometer is cheap, relatively light, and very robust; it is as accurate as the newer instruments,¹ but measurements made with it are exposed to observer error and bias.² This is of little consequence in clinical management as the decision to treat or to modify treatment in hypertension is based on much larger changes of blood pressure. For clinical practice, therefore, the standard sphygmomanometer is perfectly adequate. For the epidemiologist or for a clinical trial it is not. Comparison of blood pressure in two groups of people may show small, highly significant differences within a range in which bias could account for the difference. There is no alternative but to use either double-blind recording with a standard sphygmomanometer or, preferably, double-blind recording with a special sphygmomanometer capable of eliminating or reducing observer bias.³ Such machines operate on mechanical or electronic principles¹⁻³; some produce a written print-out of blood pressure and some take regular measurements automatically. These may be useful for serial observations in anaesthesia, intensive care units, or medical research. Portable machines recording intra-arterial pressure through an indwelling catheter can record blood pressure continuously for prolonged periods.⁴

¹ Labarthe, D R, Hawkins, C M, and Remington, R D, *American Journal of Cardiology*, 1973, 32, 546.
² Rose, G A, Holland, W W, and Crowley, E A, *Lancet*, 1964, 1, 337.
³ Garrow, J S, *Lancet*, 1963, 2, 1205.
⁴ Littler, W A, et al, *British Medical Journal*, 1972, 3, 76.