PRINCIPLES OF MEDICAL STATISTICS

IV.—THE VARIABILITY OF OBSERVATIONS

When a series of observations has been tabulated—i.e., put in the form of a frequency distribution—the next step is the calculation from the series of certain values which may be used as descriptions of it, and enable comparisons to be made between one frequency distribution and another. The value most commonly employed is the *mean*, or arithmetic average, and frequently the worker's analysis comes to an end with this calculation. But taken alone the mean is of very limited value, for it can give no information regarding the variability with which the observations are scattered around itself, and that variability (or lack of variability) is an important characteristic of the frequency distribution. As an example Table II shows the frequency distributions of age at death from two causes of death amongst women. The mean, or average, age at death does *not* differ greatly between the two, being 3·72 years for deaths registered as due to diseases of the Fallopian tube and 35·2 years for those attributed to abortion. But both the table and the diagram based upon it (Fig. 5) show that the difference in the variability, or scatter, of the observations round their respective means is very considerable. With diseases of the Fallopian tube the deaths in the year's records are spread over the age-groups 0–5 to 70–75, while deaths from abortion range only between 20–25 and 45–50. As a further description of the frequency distribution we clearly need a measure of its degree of variability round the average. A measure commonly employed in medical (and other) papers is the *range*, as quoted above—i.e., the distance between the smallest and greatest observations.

Table II

Frequency distribution of deaths of women in England and Wales, 1934, *from* (1) *diseases of the Fallopian tube, and* (2) *abortion not returned as septic, according to age.*

Age in years.	Diseases of the Fallopian tube.	Abortion not returned as septic.
0–	1	—
5–	—	—
10–	1	—
15–	7	—
20–	12	6
25–	35	21
30–	42	22
35–	33	19
40–	24	26
45–	27	5
50–	10	—
55–	6	—
60–	5	—
65–	1	—
70–75	2	—
Total ..	206	99

Though this measure is often of considerable interest, it is not very satisfactory as a description of the general variability, since it is based upon only the two extreme observations and ignores the distribution of all the observations within those limits—e.g., the remainder may be evenly spread out over the distance between the mean and the outlying values or be bunched very closely round the mean, leaving the outlying values as very rare events. In publishing observations it is certainly insufficient

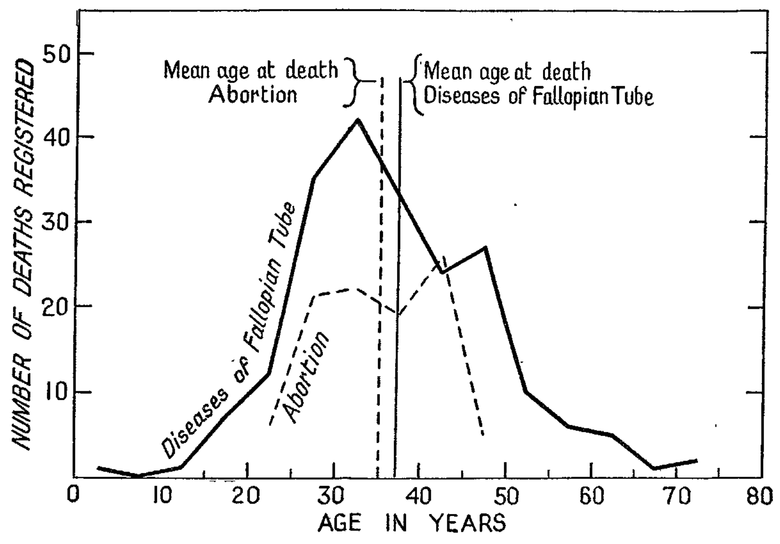

FIG. 5.—Frequency distribution of deaths of women in England and Wales, 1934, from (*a*) diseases of the Fallopian tube, and (*b*) abortion not returned as septic.

to give only the mean and the range; as previously pointed out, the frequency distribution itself should be given—even if no further calculations are made from it.

STANDARD DEVIATION

The further calculation that the statistician invariably makes is of the Standard Deviation, which is a measure of the scatter of the observations around their mean. Put very briefly the development of this particular measure is as follows. Suppose we have the following twenty observations of systolic blood pressure made on twenty different persons:—

Table III

Twenty observations of systolic blood pressure.	Deviation of each observation from the mean (mean = 128).	Square of each deviation from the mean.
(*1*)	(*2*)	(*3*)
98	−30	900
160	+32	1024
136	+ 8	64
128	0	0
130	+ 2	4
114	−14	196
123	− 5	25
134	+ 6	36
128	0	0
107	−21	441
123	− 5	25
125	− 3	9
129	+ 1	1
132	+ 4	16
154	+26	676
115	−13	169
126	− 2	4
132	+ 4	16
136	+ 8	64
130	+ 2	4
Sum 2560	0	3674

The mean, or average, blood pressure is the sum of the observations divided by 20 and equals 128. It is obvious from cursory inspection that the variability of the individual values round this mean is considerable. They range from 98 to 160; on the other hand, a large proportion of the values lie in the narrower range 125–135 (50 per cent. of them). The mean and range are not sufficient to describe the distribution. As a further step we may calculate the amount by which each observation differs, or deviates, from the average as is shown in col. 2. If these differences are added, *taking their sign into account*, the sum must equal nought, for a characteristic of the arithmetic mean, or average, is that the sums of the positive and of the negative deviations of the observations from itself are equal. In this example the sum of the deviations above the mean is $+93$ and below the mean -93. Two ways of avoiding this difficulty are possible: we may add all the deviations, regardless of their sign, or we may square each deviation so that each becomes positive. If the deviations be added in the example, ignoring their sign, the sum is 186 and the *mean deviation* is, therefore, 186 divided by 20 (the number of observations) or 9·3. This is a valid measure of the variability of the observations around the mean, but it is one which, for reasons involved in the problems of sampling, discussed later, is of less value in general statistical work than the *standard deviation*. To reach the latter the squared deviations are used. Their sum is 3674 so that the mean square deviation is this sum divided by 20 which equals 183·7. The standard deviation is the square root of this value (for having squared the original deviations the reverse step of taking the square root must finally be made) and in this example is, therefore, 13·55.

A large standard deviation shows that the frequency distribution is widely spread out from the mean while a small standard deviation shows that it lies closely concentrated about the mean with little variability between one observation and another. For example, the standard deviation of the widely spread age distribution of deaths attributed to diseases of the Fallopian tube (given in Table II) is 11·3, while of the more concentrated age distribution of deaths attributed to abortion, not returned as septic, it is only 6·8. The frequency distributions themselves clearly show this considerable difference in variability. The standard deviations have the advantage of summarising this difference by measuring the variability of each distribution in a single figure; they also enable us to test, as will be seen subsequently, whether the observed differences between two such means and between two such degrees of variability are more than would be likely to have arisen by chance.

In making a comparison of one standard deviation with another it must, however, be remembered that this criterion of variability is measured in the same units as the original observations. The mean height of a group of school-children may be 48 in. and the standard deviation 6 in.; if the observations were recorded in feet instead of in inches, then the mean would be 4 ft. and the standard deviation 0·5 ft. It follows that it is not possible by a comparison of the standard deviations to say, for instance, that weight is a more variable characteristic than height; the two characteristics are not measured in the same units and the selection of these units—e.g., inches or feet, pounds or kilogrammes—must affect the comparison. In fact it is no more helpful to compare these standard deviations than it is to compare the mean height with the mean weight. Further, a standard deviation of 10 round a mean of 40 must indicate a relatively greater degree of scatter than a standard deviation of 10 round a mean of 400, even though the units of measurements are the same.

COEFFICIENT OF VARIATION

To overcome these difficulties of the comparison of the variabilities of frequency distributions measured in different units or with widely differing means, the Coefficient of Variation is utilised. This coefficient is the standard deviation of the distribution expressed as a percentage of the mean of the distribution—i.e., Coefficient of Variation = (Standard Deviation $\div$ Mean) $\times$ 100. If the standard deviation is 10 round a mean of 40, then the former value is 25 per cent. of the latter; if the standard deviation is 10 and the mean 400, the former value is 2·5 per cent. of the latter. These percentage values are the coefficients of variation. The original unit of measurement is immaterial for this coefficient, since it enters into both the numerator and the denominator of the fraction above. For instance, with a mean height of 48 in. and a standard deviation of 6 in. the coefficient of variation is $6/48 \times 100 = 12{\cdot}5$ per cent. If the unit of measurement is a foot instead of an inch, the mean height becomes 4 ft., the standard deviation is 0·5 ft. and the coefficient of variation is $0{\cdot}5/4 \times 100 = 12{\cdot}5$ per cent. again. Similarly the coefficient of variation of the blood pressure of Table III is $(13{\cdot}55/128) \times 100 = 10{\cdot}6$.

These measures of variability are just as important characteristics of a series of observations as the measures of position—i.e., the average round which the series is centred. The important step is to "get out of the habit of thinking in terms of the average, and think in terms of the frequency distribution. Unless and until he [the investigator] does this, his conclusions will always be liable to fallacy. If someone states merely that the average of something is so-and-so, it should always be the first mental question of the reader: 'This is all very well, but what is the frequency distribution likely to be? How much are the observations likely to be scattered round that average? And are they likely to be more scattered in the one direction than the other, or symmetrically round the average?' To raise questions of this kind is at least to enforce the limits of the reader's knowledge, and not only to render him more cautious in drawing conclusions, but possibly also to suggest the need for further work" (G. U. Yule: Industrial Health Research Board, Report No. 28, 1924, p. 10).

EXAMPLES OF VARIABILITY

The practical application of the measures of variability may be illustrated by the figures below which are taken from a valuable statistical study of blood pressure in healthy adult males by P. L. McKinlay and A. B. Walker (*Edinb. med. J.* 1935, **42**, 407). In the original the full frequency distributions are set out (see Table below).

The variability of these physiological measurements which is apparently compatible with good health is striking. It leads the authors to conclude that we must hesitate to regard as abnormal any isolated measurements in otherwise apparently fit individuals. Some of the measurements they found are definitely within the limits regarded as pathological and study is necessary to determine whether such large deviations from the "normal" have any unfavourable prognostic significance. It is clear that the mean value alone is a very insecure guide to "normality."

As a further example of the importance of taking note of the variability of observations the incubation period of a disease may be considered. If the day of exposure to infection is known for a number of persons we can construct a frequency distribution of the durations of time elapsing between exposure to infection and onset of disease as observed clinically. If these durations cover a relatively wide range, say 10–18 days with an average of 13 days, it is obvious that observation or isolation of those who have been exposed to infection for the *average* duration would give no high degree of security. For security we need to know the proportion of persons who develop the disease on the fourteenth, fifteenth, &c., day after exposure; if these proportions are high—i.e., the standard deviation of the distribution is relatively large—isolation must be maintained considerably beyond the *average* incubation time. In such a case the importance of variability is indeed obvious; but there is a tendency for workers to overlook the fact that in *any* series of observations the variability, large or small, is a highly important characteristic.

Means, Standard Deviations, Coefficients of Variation, and the Range of Measurement (based on 566 observations).

—	Mean.	Standard deviation.	Coefficient of variation.[1]	Range.
Age (years)	23·2	4·02	17·31	18– 40
Heart-rate (beats per minute)	77·3	12·83	16·60	46–129
Systolic B.P. (mm.) ..	128·8	13·05	10·13	97–168
Diastolic B.P. (mm.) ..	79·7	9·39	11·78	46–108
Pulse pressure (mm.)	49·1	11·14	22·25	24– 82

In actual practice the calculation of the standard deviation is not usually carried out by the method shown above—i.e., by computing the deviation of each observation from the mean and squaring it. Shorter methods are available both for ungrouped observations (like the twenty measurements of blood pressure in Table III) and for grouped observations (like those in Table II). These methods are described fully in many text-books—e.g., Woods and Russell.

[1] These are the figures given in the original. According to the values of the mean and standard deviation the final coefficient of variation is 22·69. Probably the original calculation was based upon the values to a further number of decimal places.

SYMMETRY

Another characteristic of the frequency distribution is its symmetry or lack of symmetry. With a completely symmetrical distribution the frequency with which observations are recorded at each point on the graph, or within certain values below the mean, is identical with the frequency of observations at the same point, or within the same values, above the mean. With asymmetry the observations are not evenly scattered on either side of the mean but show an excess on one side or within particular values—e.g., with a mean of 50, observations below the mean may not fall below 20 units from that point, the lowest observation recorded being 30, while on the positive side of the mean observations 40 units above the mean may be observed, values of 90 being recorded. The tabulated distribution and, still more, a graph of it will afford an *indication* of this characteristic.

Summary

As descriptions of the frequency distribution of a series of observations certain values are necessary, the most important of which are, usually, the mean and standard deviation. The mean alone is rarely, if ever, sufficient. In statistical work it is necessary to think in terms of the frequency distribution as a whole, taking into account the central position round which it is spread (the mean or average), the variability it displays round that central position (the standard deviation and coefficient of variation), and the symmetry or lack of symmetry with which it is spread round the central position.

A. B. H.

SPECIAL ARTICLES

SOME ASPECTS OF INFANT NUTRITION *

BY CÉCILE ASHER, M.D., B.Sc. Lond., D.C.H.

MEDICAL REGISTRAR, DUCHESS OF YORK HOSPITAL, MANCHESTER; LATE SENIOR RESIDENT MEDICAL OFFICER TO THE ALDER HEY CHILDREN'S HOSPITAL, LIVERPOOL

THE subject of infant nutrition is a wide one, and this article concerns certain aspects only, and in particular the type of material with which we deal in Alder Hey, our methods of investigation and treatment, and the conclusions at which we have arrived as a result of 3½ years' minute observation of 4656 infants under one year. In order to investigate our material we divided the children of this age into weight groups. Of the deaths in children in these groups, 69 per cent. occurred in those who were more than 20 per cent. below average weight. This serves to show what an important part the infant service plays in the work of the hospital, and how necessary is a right understanding of the disorders and nutritional needs of these small patients.

They require unremitting care and watchfulness on the part of the nursing staff. They are apt to change quickly for the worse; on the other hand one often observes an almost miraculous change as the result of 24 hours' intensive treatment, so that it is our practice never to give up hope until an infant is actually dead.

* A paper read to the Maternity and Child Welfare section of the Society of Medical Officers of Health on July 4th, 1936.

Methods of Investigating Faults

If a baby is brought to us with the complaint of any nutritional disturbance—i.e., diarrhœa, vomiting, or failure to gain weight—we invariably approach the problem by assuming that there is (1) something wrong with the feed, (2) something wrong with the baby, or (3) something wrong with both.

THE FEED

We first examine the history carefully, noting the type of feed, the precise amounts given, the addition of sugar and vitamins, and the intervals at which they are given. The child's reaction to each feed is ascertained—that is to say, whether he has gained or not, the type and frequency of the stools, whether vomiting has occurred or not, and if he seems satisfied. A careful analysis is then made to see if the feed