THE USE OF THE MEDIAN IN TESTS OF SIGNIFICANCE.

BY S. R. SAVUR.
(*Meteorological Office, Poona.*)

Received March 3, 1937.

THE arithmetic mean has been in use for such a long time and has consequently and for some other reasons become such a general favourite—even its name has been shortened to "the mean"—that the mere idea of using some other statistical parameter in its place is very likely to be not only looked down upon but also scoffed at by many Statisticians. The general belief is that no other statistic* calculated from a sample has any claim to the seat of honour that has been given to the mean of that sample. We will examine in this paper some of the advantages which the mean is said to possess over similar statistics, and show that those advantages can also be secured by one other at least.

Let us first consider the likely rivals to the mean. They are the mode and the median, if for the present we leave out of consideration the geometric and the harmonic means. We can dismiss the mode at once because it does not possess the essential quality, namely, that a useful statistic should be capable of being determined accurately for a sample. The smaller the sample, the greater is the uncertainty in the determination of its mode. Besides, there may be more than one mode. Thus we are left with the median.

The median can be determined uniquely just as the mean, but far more quickly. This advantage, however, cannot be given much weight. The reasons why the mean has been preferred to the median are the following :—

1. The median is not affected by the *magnitude* of the values so long as a change in the magnitude of any value does not alter its position with regard to the median, whereas the mean depends upon the individual values.

2. The mean is *consistent* and *sufficient* and is the most *efficient* of all statistics.

* This beautiful word coined by R. A. Fisher did not find favour in the eyes of Karl Pearson, see the first foot-note on page 49 of (6). Nevertheless it is to be hoped that the word will still continue to be used.

The first reason generally put forward is as the most powerful argument in favour of the mean as against the median, because it is believed that as the mean takes into account the individual values in the sample, it is more representative of the sample and gives more information from the sample than the median. Precisely what this extra information is nobody seems to have taken the trouble to state or even to find out.

Fisher has defined the "amount of information" in a sample and the "amount of information" supplied by any statistic calculated from that sample [(see ([12]) and ([13])], but these definitions have only theoretical interest. Besides, the assumption that is made there is that the distribution of the efficient statistic tends to normality with the increase in the number of individuals in the sample, an assumption which is true in the case of samples drawn from a normally distributed population but may not hold in the case of samples from a non-normally distributed population.

Suppose we are told that the mean of a sample is 30. If on the other hand we are told that the median of the sample is 35, our knowledge about the sample or of the population from which the sample was obtained is neither more nor less than if we had been given only the mean of the sample. By itself, each is as equally useful or useless as the other.

Let us consider the practical use to which the mean of a sample is put. If there is only one sample, we test the significance of its mean. If there be more than one sample we try to find out whether these samples could not have been obtained by mere chance from the same population. These tests are based on the assumption of "normal" (or "not far from normal") distribution in the populations from which our samples were obtained. We give below better tests than these, better in the sense that they are *independent of the distributions of the individuals in the populations from which our samples were drawn.* Hence the so-called information possessed by the mean of a sample is not of interest, so long as the original distribution is not known.

Let us now turn to the second reason. It is perfectly true that the mean of a sample possesses these characteristics as defined by Fisher, *provided that the sample was drawn from a normally distributed population.* This proviso is the most important and unfortunately it is the greatest stumbling block also. In other cases the mean may lose one or all of these three properties. Fisher, himself, has given on page 321 of ([14]) an example where he has shown that the mean is an entirely useless statistic for the purpose of determining the mean of that population from a sample drawn from it. Now, it is only in a very few cases such as the tossing of coins, the throwing of dice, the counting of the number of corpuscles with a hæmacytometer, etc., that we know the form of the frequency-distribution in the population from which our sample has been obtained. *In all other cases we have absolutely no idea of this distribution.* Then what is it that is done in such cases? The most usual procedure has been to assume that our sample was obtained from a population in which the frequency-distribution is either normal or not far from normal. Is this assumption justified? We have to confess that so far no satisfactory justification has been forthcoming. The assumption is made with a feeling (or is it with a hope?) that we may not be far wrong.

Let us consider the full significance of this assumption. We calculate some statistics from our sample, apply certain tests to these statistics and draw appropriate conclusions. The tests are such that these conclusions are quite correct if the frequency-distribution in the population from which our sample was drawn is normal. In other cases they are subject to errors the magnitudes of which depend upon how far the frequency-distribution in the population differs from the normal. Our assumption that this distribution is not far from normal means that we consider that these errors are not material. Hence *the same tests are being applied to samples in all those cases in which the frequency-distributions in the populations from which these samples were drawn are unknown to us.* As we cannot prove or disprove the legitimacy of our assumption, it follows that we can neither show that our deductions are correct, nor find out whether they are wrong. This being the case we do not appear to be better off after applying the tests than if we had not applied them.

We see thus that the tests in use at present may not give, in a number of cases, the information that is sought from them. Further, we are unable to find out which of our deductions, after applying the existing tests, are trustworthy because we are not able to determine definitely which of our samples had been drawn from a normally distributed population. The question then arises, "Are there any tests which are applicable to all cases, that is to say are there tests which are independent of the frequency-distributions in the populations from which our samples are drawn?" The answer is *most certainly yes, and an infinite number of such tests can be devised.* We will consider for the present only those tests in which the median is used.

Credit should undoubtedly go to P. R. Crowe for having boldly espoused recently the cause of the median in his paper.[1] His enthusiasm for the median must have caught H. A. Mathews who followed with his paper (2). Still more recently Crowe has given another paper (3) on the same subject.

In his paper (1) Crowe has given many advantages which the median possesses over the mean in statistical tests. In addition to these the median has another advantage which is the most important from the statistical point of view, *namely, that using the median we can devise tests which are applicable to samples irrespective of the nature of the frequency-distributions in the populations from which the samples were obtained.* A preliminary note on these tests was given in (4). Before describing these tests we will start with two definitions :—

The definition of the median of a population is well known, namely, it is the value such that there are as many individuals below it as there are above it in that population.

We will use the following definition for the median of a sample.†

If the values (say n in number) in the sample be arranged in the ascending order of magnitude thus

$$y_1 \quad y_2 \quad y_3 \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad y_{n-1} \quad y_n,$$

the median of the sample is

(1) y_{m+1} when $n = 2m + 1$, and

(2) $\frac{1}{2}(y_m + y_{m+1})$ when $n = 2m$

The phrase "the median of a population" will occur so often that it appears necessary to have a special symbol for it. All the likely symbols M, m, and μ have already been used to represent some parameter or other. For this reason I suggested in (4) that the Sanskrit letter म (pronounced like "me" in "calmer" but with the vowel sound slightly less prolonged) may be used for this purpose. The letter appears to be quite appropriate not only on account of its sound but also because it is the first letter of the Sanskrit word मध्यम which means middle or central. As for the sample, its median will be represented by म′.

Thus for "the median of the population from which sample A was obtained" we will write "म of sample A", or merely "म of A", whereas by "म′ of sample A" or "म′ of A" we will mean "the median of the sample A".

Limits for म.—We will now derive two limits between which the म of our sample should lie in order that the sample could have been obtained from that population by random chance.

---

† For other definitions see Whittaker and Robinson's *Calculus of Observations*, p. 197 (1924 edition). The method for determining the median of a sample suggested by Dunham Jackson is interesting in that he must have been imagining that from that method one will obtain the best approximation to the median of the population from which that sample was drawn.

From the definition of the median of a population we see that the chance is $\frac{1}{2}$ of obtaining a value less than म at random from that population. The chance is also $\frac{1}{2}$ of drawing at random from that population a value greater than म.

Suppose our sample contains n values.

The chance that in a random sample of n values there are not more than l values below म is given by

$$P = \frac{1}{2^n} (1 + {}_nC_1 + {}_nC_2 + \ldots\ldots + {}_nC_l),$$

$$= I_{\frac{1}{2}} (n - l, l + 1) \qquad \ldots \qquad \ldots \qquad \ldots \qquad \ldots \qquad \text{(I)}$$

where $I_{\frac{1}{2}} (n - l, l + 1)$ is an incomplete β-function ratio written after the manner of K. Pearson.[7]

For a given value of P and n we can solve equation (I) for l. A method of solution is given below.

Let us choose some limit for random chance, say 5%. We put $P = 0{\cdot}05$ in (I) and solve for l. In general, we will get two consecutive integers such that the corresponding values of P from (I) are on either side of $0{\cdot}05$. We use l_1 the smaller of the two integers.

We now arrange our values in the sample in the ascending order of magnitude thus,

$$y_1 \;\; y_2 \;\ldots\; y_{l_1} \ldots\; y_{n+1-l_1} \ldots\; y_n.$$

It is clear that when म is equal to or less than y_{l_1} the chance is less than $0{\cdot}05$ of getting in a random sample of n values not more than l_1 values which are below म; that is to say, on our 5% limit for random chance our sample could not have been obtained from a population in which म is equal to or less than y_{l_1}. Hence म $> y_{l_1}$.

Similarly we can see that on the same 5% limit for random chance म $< y_{n+1-l_1}$. Thus

$$y_{l_1} < \text{म} < y_{n+1-l_1} \qquad \text{(i)}$$

Suppose $y_{l_1} < y_{l_1+1}$ and $y_{n-l_1} < y_{n+1-l_1}$. Now the y's are measured correct to a certain number of significant digits. We increase y_{l_1} and decrease y_{n+1-l_1} by unity in the last significant digit. Let the new values so obtained be denoted by y'_{l_1} and y'_{n+1-l_1}. Then our inequality (i) can be written as

$$y'_{l_1} < \text{म} \leqslant y'_{n+1-l_1} \qquad \text{(ii)}$$

Since the limits y'_{l_1} and y'_{n+1-l_1} were determined on the 5% limit for random chance, we will term them the 5% limits for म and the interval y'_{l_1} to y'_{n+1-l_1} (including the limits) the 5% interval for म.

In deriving (ii) we assumed that $y_{l_1} < y_{l_1+1}$ and $y_{n-l_1} < y_{n+1-l_1}$. If, however, $y_{l_1} = y_{l_1+1}$ then $y'_{l_1} = y_{l_1}$, and when $y_{n-l_1} = y_{n+1-l_1}$ we take $y'_{n+1-l_1} = y_{n+1-l_1}$.

In a similar manner we can find the limits and the interval for म on any assumed limit for random chance.

It may be noted in passing that the problem considered here is to some extent the converse of that discussed in [5] and [11], where equations similar to (I) have occurred. In the problems considered in those papers l (or α in the notation used there) was assumed to be known but the population chance of success was unknown and so the limits to the population chance of success were obtained. In the problem solved here, the population chance of success is known—it is equal to $0 \cdot 5$—while l is not known. Hence the value of l is found from equations like (I).

The म *test of significance.*—We will now give a test of significance which may be stated in the two following forms, the first being applicable to a single sample, while the second when there are two or more samples.

1. The median of a sample is significant (*i.e.*, it is significantly different from zero) on any given limit for random chance, if the interval for म calculated from that sample on that limit for random chance does not contain zero. If zero be an end of that interval, the median can be considered to be just significant.

2. The medians of two samples are significantly different from each other if the intervals (on our limit for random chance) for the corresponding म's do not have a common part. If the higher end of one of the intervals be the same as the lower end of the other, the corresponding म's can be considered to be just significantly different from each other.

Before applying these tests to a few numerical examples we will show how to solve equation (I) for l. There are two methods :—

(*a*) In this, the first method, the tables[7] are used. There is, however, a certain upper limit to n, depending upon the value for P (our limit for random chance), beyond which the tables cannot be used. For such cases the method (*b*) is available. For example, when $P = 0 \cdot 05$, n should not be greater than 84, and when $P = 0 \cdot 01$, n must not exceed 79 if tables[7] are to be used. The following is the method.

We start with an integral value, say m, slightly less than $\frac{1}{2}n$ and from tables[7] we find $I_{\frac{1}{2}}(n+1-m, m)$. If this be greater than P, our limit for random chance, we find $I_{\frac{1}{2}}(n+2-m, m-1)$, that is, we increase p and decrease q (both in Pearson's notation) by unity. In this manner we will find in general two values $I_{\frac{1}{2}}(n+1-m_1, m_1)$ and $I_{\frac{1}{2}}(n-m_1,$

$m_1 + 1)$ such that $I_{\frac{1}{2}}\ (n + 1 - m_1,\ m_1) < P < I_{\frac{1}{2}}\ (n - m_1, m_1 + 1)$. We take $l_1 = m_1 - 1$, that is to say we take the samller value, m_1, of q in these two I-functions and diminish it by unity to get our value for l_1.

(*b*) This method, which is approximate, is applicable for all values of n. The error introduced by this method diminishes very quickly with the increase in n. Although the error may be small, our method of selecting l_1 may occasionally magnify the error into $\pm$ 1 in the calculated value of l_1. When n is large, l_1 is also large, and as will be seen later on an error of ± 1 in l_1 in this case will not affect our results materially.

This method depends upon the fact that the binomial distribution $(\frac{1}{2} + \frac{1}{2})^n$ approaches the corresponding normal distribution very rapidly with the increase in n. The following is the method:

From Table II of ([8]) we find the value of x for which the value of $\frac{1}{2}(1 + \alpha)$ is $1 - P$, P being our assumed limit for random chance. Let the value of x be x_1. Then

$$l_1 \text{ is the integral part of } \tfrac{1}{2}(n - 1 - x_1\sqrt{n}) \qquad \text{(iii)}$$

The limits for random chance usually adopted are 5% and 1%.

When 5% is the limit for random chance,

$$l_1 \text{ is the integral part of } \tfrac{1}{2}\,(n - 1 - 1\cdot 6449\sqrt{n}) \qquad \text{(iv)}$$

and when 1% is the limit for random chance,

$$l_1 \text{ is the integral part of } \tfrac{1}{2}(n - 1 - 2\cdot 3263\sqrt{n}) \qquad \text{(v)}$$

When we use 5% as our limit for random chance, it is a very happy accident that for $n \not> 84$, in which case tables[7] can be used, there is no error in l_1 as calculated by the relation (iv). *Hence tables*[7] *are not necessary.* When, however, we use 1% as our limit for random chance, errors in l_1 as calculated from (v) are found only in the following cases for $n \not> 79$, for all of which tables[7] can be used.

TABLE A.

n	l_1 from relation (v)	l_1 correct
19	3	4
27	6	7
41	12	13
53	17	18

With this Table A before us we can dispense with the tables[7] *when we adopt* 1% *as our limit for random chance.* Thus, we can completely do away with tables[7] when we use either 5% or 1% as our limit for random chance.

For higher values of n, errors will be introduced occasionally in the values of l_1 calculated by (iv) or (v), but, as mentioned above, the errors will only be $\pm$ 1 and an allowance can always be made for this as shown in example 3 below.

Example 1. To show that the method (b) does not introduce a large error even for small values of n we shall apply both (a) and (b) to the case where $n = 15$. Suppose we take $P = 0 \cdot 05$.

We will use the method (a) first. Proceeding in the manner described above we find that $I_{\frac{1}{2}}(11, 5) = \cdot 0592$ and $I_{\frac{1}{2}}(12, 4) = \cdot 0176$. The smaller value of q is 4. Hence $l_1 = 3$.

Let us now use the method (b). From (iv) we see that l_1 is the integral part of $\frac{1}{2}(14 - 1 \cdot 645 \times 3 \cdot 872)$ which is $3 \cdot 82$.

Thus $l_1 = 3$, the same result as that obtained by method (a).

We will now apply the म test of significance defined above to some numerical examples.

Example 2. Let us take the example given on page 113 of Fisher's book.[9] This example relates to the test conducted to find out which of two narcotic drugs (*Dextro-*) and (*Lævo-*)‡ is the more potent. The results are reproduced below :—

TABLE I.

Additional hours of sleep gained by the use of hyoscyamine hydrobromide.

Patient	1 (*Dextro-*)	2 (*Lævo-*)	Difference (2 − 1)
1	+ 0·7	+ 1·9	+ 1·2
2	− 1·6	+ 0·8	+ 2·4
3	− 0·2	+ 1·1	+ 1·3
4	− 1·2	+ 0·1	+ 1·3
5	− 0·1	− 0·1	0·0
6	+ 3·4	+ 4·4	+ 1·0
7	+ 3·7	+ 5·5	+ 1·8
8	+ 0·8	+ 1·6	+ 0·8
9	0·0	+ 4·6	+ 4·6
10	+ 2·0	+ 3·4	+ 1·4

‡ These are abbreviations for *Dextro*-hyoscyamine hydrobromide and *Lævo*-hyoscyamine hydrobromide.

(*a*) Suppose each of the patients had been given the two drugs in succession but in a random order. The last column gives the excess in the hours of sleep produced by (*Lævo-*) over those produced by (*Dextro-*). If the two drugs were equally powerful it is easy to see that the population of these differences in the two durations of sleep is symmetrically distributed about zero. Hence, the median as well as the mean of this population is zero. Our test of significance is thus to see whether median of the sample, in the last column of the table, is significantly different from zero.

To apply the test we arrange the values in the ascending order of magnitude thus

0·0, + 0·8, + 1·0, + 1·2, + 1·3, + 1·3, + 1·4, + 1·8, + 2·4, + 4·6.

The median, म′ is + 1·3.

Since $n = 10$, we find that $l = 1$ on the 5% limit for random chance.

Hence y_{l_1} and y_{n+1-l_1} are 0·0 and +4·6, and y'_{l_1} and y'_{n+1-l_1} are + 0·1 and + 4·5, as $y_2 > y_1$ and $y_{10} > y_9$. Thus the 5% interval is from + 0·1 to + 4·5 (limits included).

This interval does not contain zero, and so we conclude that the *median* of the sample is significantly different from zero, that is to sáy (*Lævo-*) is significantly more potent than (*Dextro-*).

Fisher using " Student's " *t* test finds that the *mean* of the sample is significantly different from zero.

(*b*) Suppose now that the drugs had been given to twenty different people. In this case we must use the columns under 1 (*Dextro-*) and 2 (*Lævo-*) separately. Arranging the values in the ascending order of magnitude, the values in the two samples are:

(*a′*) − 1·6, − 1·2, − 0·2, − 0·1, 0·0, + 0·7, + 0·8, + 2·0, + 3·4, + 3·7 and

(*b′*) − 0·1, + 0·1, + 0·8, + 1·1, + 1·6, + 1·9, + 3·4, + 4·4, + 4·6 , + 5·6.

The 5% interval for (*a′*) is − 1·5 to + 3·6 and that for (*b′*) is 0·0 to + 5·5. These two intervals have a common part and so the *medians* of the two samples are not significantly different from each other on the 5% limit for random chance. On the *t* test Fisher finds that the *means* of the two samples are not significantly different from each other on the same limit for random chance.

We will now show how to apply the म test to those cases in which n is quite large. We will take the example 18 on page 110 of ([9]).

Example 3. A test was conducted to see whether the mean difference in height between the two sexes is significant. The following table, which is a reproduction of Table 26 on page 110 of([9]) with an extra row, gives the distribution of the excess in stature of a brother over his sister in 1401 pairs.

TABLE II.

a	−5	−4	−3	−2	−1	0	1	2	3	4	5	6
b	$\frac{1}{4}$	$1\frac{1}{2}$	$1\frac{1}{4}$	$4\frac{1}{2}$	$11\frac{1}{4}$	$27\frac{1}{2}$	$71\frac{3}{4}$	$122\frac{3}{4}$	$171\frac{3}{4}$	$209\frac{3}{4}$	$220\frac{1}{2}$	$205\frac{1}{2}$
c	$\frac{1}{4}$	$1\frac{3}{4}$	3	$7\frac{1}{2}$	$18\frac{3}{4}$	$46\frac{1}{4}$	118	$240\frac{3}{4}$	$412\frac{1}{2}$	$622\frac{1}{4}$	$842\frac{3}{4}$	$1048\frac{1}{4}$

a	7	8	9	10	11	12	13	14	15	16
b	$148\frac{3}{4}$	$95\frac{3}{4}$	57	26	$11\frac{1}{4}$	$8\frac{1}{2}$	$2\frac{3}{4}$	1	1	$\frac{3}{4}$
c	1197	$1292\frac{3}{4}$	$1349\frac{3}{4}$	$1375\frac{3}{4}$	1387	$1395\frac{1}{2}$	$1398\frac{1}{4}$	$1399\frac{1}{4}$	$1400\frac{1}{4}$	1401

a = Stature difference (in inches) ; b = Frequency ; c = Total frequency from the negative end.

(In the following detailed working of the problem brevity has been purposely sacrificed for clarity in many places.)

There are 1401 values in our sample. Hence the median is the 701th value when the values are arranged in the order of magnitude, that is to say it is some value between 4 and 5 inches (see row c). The variable, the difference in stature, obviously varies continuously and so we can very reasonably assume that the portion of the frequency curve between 4 and 5 is straight as the interval is small. Now, half the frequency listed under 4 is below 4 inches and half that listed under 5 is above 5 inches. Thus there are $517\frac{3}{8}$ $(= 412\frac{1}{2} + \frac{1}{2} \times 209\frac{3}{4})$ values below 4 inches, and $215\frac{1}{8}$ $[= \frac{1}{2}(209\frac{3}{4} + 220\frac{1}{2})]$ values in the interval 4 to 5 inches. Assuming that the frequency curve in the interval 4 to 5 inches is straight, we get

$$\text{the 701th value} = 4 + \frac{701 - 517\frac{3}{8}}{215\frac{1}{8}} = 4 \cdot 8537 \text{ inches.}$$

It is very difficult to measure the height of a person correct to 0·1 inch. Hence it will be more than sufficient if we calculate the median value correct to 0·01 inch at the utmost.

Thus the median of our sample is 4·85 inches.

To find the 5% interval we use (iv). We see that l_1 is the integral part of $\frac{1}{2}(1400 - 1 \cdot 645 \sqrt{1401}) = 669 \cdot 2$.

Thus $l_1 = 669$.

The lower limit is the 669th value from the negative end and is 4·71 inches (more accurate value is 4·7095 inches). The upper limit is the 669th value from the positive end and is 5·00 inches (more accurate value is 4·9979 inches), *i.e.*, $y_{669} = 4 \cdot 71''$ and $y_{733} = 5 \cdot 00''$.

To find our y's we must calculate the value of y_{670} and y_{732}. Now the change in height per individual is $1/215\frac{1}{3} = \cdot 0046$. Hence $y_{670} = 4\cdot 71$ and $y_{732} = 4\cdot 99$. We see that $y_{669} = y_{670}$ and $y_{732} < y_{733}$.

Thus the 5% interval is 4·71 to 4·99 (limits included).

This interval does not contain zero. Hence the median of this sample is significant, its म being given by

$$4''\cdot 71 \leqslant \text{म} \leqslant 4''\cdot 99.$$

Fisher's values, on the assumption of the normal law, are :—

Mean of the sample = 4″·895, and the mean of the population is some value between 4″·75 and 5″·00.

While explaining the method (*b*) for calculating l_1 it was said that the error introduced in the calculated value of l_1 is occasionally ± 1, and that when l_1 is large this error will not appreciably affect our results. In this example l_1 is the integral part of 669·2 and so we took $l_1 = 669$. If at all this value is in error by 1, its correct value is 668. Assuming this value we get

$$y_{668} = 4\cdot 70 \text{ and } y_{734} = 5\cdot 00.$$

Since $y_{669} > y_{668}$ and $y_{733} = y_{734}$, we get

$$y'_{668} = 4\cdot 71 \text{ and } y'_{734} = 5\cdot 00.$$

The 5% interval now comes out as 4·71 to 5·00. Thus *only an insignificant change has been produced in the interval by correcting our value of l_1 for a supposed error in it.*

We will take one more example in which Fisher has applied the t test as well as another which he has called a "General" test. This example is given in Fisher's later book.[10]

Example 4. A test was made to see whether cross-fertilised plants grow taller than self-fertilised ones. The data are given in Table 3 on page 41 of ([10]) and are reproduced below in the order of ascending magnitude.

TABLE III.

Differences in eigths of an inch between cross- and self-fertilised plants of the same pair,

− 67, − 48, 6, 8, 14, 16, 23, 24, 28, 29, 41, 49, 56, 60, 75.

Let us apply the म test to this sample, namely to see whether the median of the sample is significant.

म′ of the sample is 24.

Since $n = 15$, $l_1 = 3$ [see example (1)] on the 5% limit for random chance.

y_{l_1} and y_{n+1-l_1} are thus 6 and 56.

Hence y'_{l_1} and y'_{n+1-l_1} are 7 and 55 respectively.

म′ is therefore definitely *significant* as the 5% interval 7 to 55 does not contain zero.

Fisher, using the *t* test and the same limit for random chance, finds that the *mean* is *just significant.* He now applies another and more general test, see pages 50–54 of ([10]). On this test he finds that the *mean just fails to be significant.*

To understand this general test consider an infinite population consisting of the differences in length between cross- and self-fertilised plants of the same pair. Let us draw an unlimited number of random samples of 15 values each from this population. We will thus get another infinite population consisting of these samples. In his general test Fisher has confined himself only to a portion of this second population, namely, to those samples in which the individual values are numerically equal to those in our sample. In reality he has restricted himself still further to a part of this portion of the population, for in using only 2^{15} such samples he has virtually assumed that in the original population of these differences the frequency of each of the values (numerically) found in our sample is the same. There is really no justification for this assumption and for this restriction. Hence Fisher's test is not really as general as the म test described in this paper.

In the case of a normal distribution, the mean and the median are the same. Hence when samples are drawn from this population there is bound to be agreement in a large proportion of the cases between the results obtained from the *t* test and the म test. It should, however, be noted that the converse will not necessarily be true, that is to say, if the results of the two kinds of tests agree, we cannot definitely say that the sample has been obtained from a normally distributed population. On the other hand if there is no agreement between the results, we can more definitely conclude that the sample could not have been drawn from a normally distributed population.

Applying these arguments we can conclude that the samples in examples 2 and 3 may have been obtained from a normally distributed population, while that in example 4 could not have been obtained from such a population.

It will be clear from the foregoing that the test suggested in this paper is *independent of the frequency-distribution in the population,* whereas the tests, which are in use at present and in which the mean is used, *depend upon some specific assumption (usually the normal law) regarding the frequency-distribution in the population.* Hence the test described in this paper is quite general. This is its greatest merit. In addition another great merit of this test is that its application is far simpler and quicker than that of the *t* test, especially, when the number of data is large, as in example 3. Hence the suggestion is put forward that the म test, in which the median is used, may replace the existing ones with advantage.

There is, however, likely to be some opposition in the beginning to this move. This opposition will be mainly due to the *feeling* that as the mean depends upon the individual values in the samples, it must be somehow more representative of the sample than the median. It is intended to discuss this aspect of the matter in more detail in the near future. We will, however, remark in advance that for all practical purposes the so-called advantages of the mean will be found to have but little value and that the median is a far superior statistic for use in significance tests such as those discussed in this paper.

Summary.

In this paper a new test of significance, called the म test, in which the median is used, is given. It is shown that this test is quite independent of the, to us unknown, frequency-distributions in the populations from which our samples are drawn. The application of this test is quite simple. Hence the test appears to be far better than those in use at present and which are based on some assumed frequency-distribution, generally normal distribution. The म test is applied to some illustrative examples and the results compared with those obtained after the application of the existing tests.

REFERENCES.

1. P. R. Crowe .. *Scott. Geog. Mag.*, 1933, **49**, 73–91.
2. H. A. Mathews. *Ibid.*, 1936, **52**, 84–97.
3. P. R. Crowe .. *Geog. Review*, 1936, **26**, 463–84.
4. S. R. Savur .. *Curr. Sci.*, 1936–37, **5**, 419–21.
5. ——— .. *Ibid.*, 1936–37, **5**, 385–86.
6. K. Pearson .. *Biom.*, 1936, **28**, Parts I and II.
7. ——— .. *Tables of the Incomplete β-Function*, "Biometrika" Office, Univ. College, London.
8. ——— .. *Tables for Statisticians and Biometricians*, Part I, Camb. Univ. Press.
9. R. A. Fisher .. *Statistical Methods for Research Workers*, 4th edition, 1932.
10. ——— .. *The Design of Experiments*, 1935, Oliver and Boyd, London.
11. S. R. Savur .. *Proc. Ind. Acad. Sci.*, **A**, 1937, **5**, 222–34.
12. R. A. Fisher .. *Proc. Camb. Phil. Soc.*, 1925, **22**, 700–25.
13. ——— .. *Proc. Roy. Soc.*, **A**, 1934, **144**, 285–307.
14. ——— .. *Phil. Trans.*, **A**, 1922, **222**, 309–68.