MATHEMATICS

THE MEDIAN AND INTERQUARTILE RANGE TEST APPLIED TO FREQUENCY DISTRIBUTIONS PLOTTED ON A CIRCULAR AXIS

BY

J. WESTENBERG

(Communicated by Prof. M. W. WOERDEMAN at the meeting of May 20, 1950)

In statistical science, univariate populations and samples are specified by the distribution of the frequencies for the different values of the variate under consideration. These frequency distributions may be represented graphically by plotting the frequencies against the corresponding values of the variate, measured along a straight axis.

For the purpose of a concise description or a comparison of populations or samples, some general characteristics are introduced viz.

1. a *measure of location* (some mean value, median or mode) i.e. some central value giving the whereabouts of the distribution along the axis.
2. a *measure of dispersion*, giving the degree of scatter about this value.
3. a *measure of skewness*, giving the departure of the distribution from symmetry.

In some less common cases, as for instance the number of events during a given period of time coinciding with different directions of wind or water current, or occurring in different seasons, or at different times of the day &c., a suitable graphical representation should rather be given along a circular axis instead of a straight one (fig. 1). This

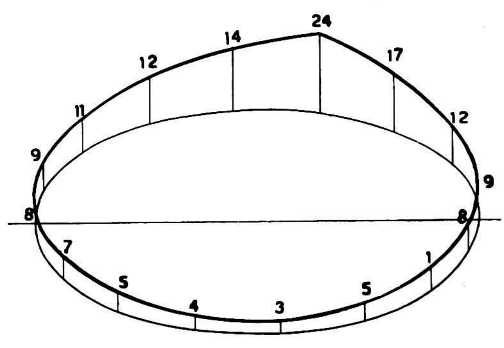

Fig. 1

is especially the case when, in grouped or in discontinuous distributions, all intervals on the axis are occupied.

It is worth noting that in graphical representation on a circular base line with all intervals occupied, the most common measures of location, viz. the means and the median, do not find application. On the contrary, the mode, the least appreciated among the measures of location, is still valid as usual for plottings on a circular axis. This does not permit, however, the application of a significance test in the comparison of two samples.

Although the concept of the median does not apply to this kind of frequency distributions, the principle of the median and the interquartile range test, needs only a slight extension to yield a significance test for the kind of distributions under discussion.

The median and the quartiles together divide a population into four equal portions. Now a division into four equal portions could as well be started at any other point of the straight axis; the two remaining outer parts together will then constitute one portion.

The same procedure could be started at every point of a circular axis. To achieve as close a resemblance as possible to the use of the median in the case of a straight axis, we could start the division in the region of minimum frequencies, between two frequency classes. This starting point will make the zero point of an ogive curve that can be imagined along the cirular axis (fig. 2). The point at which this curve

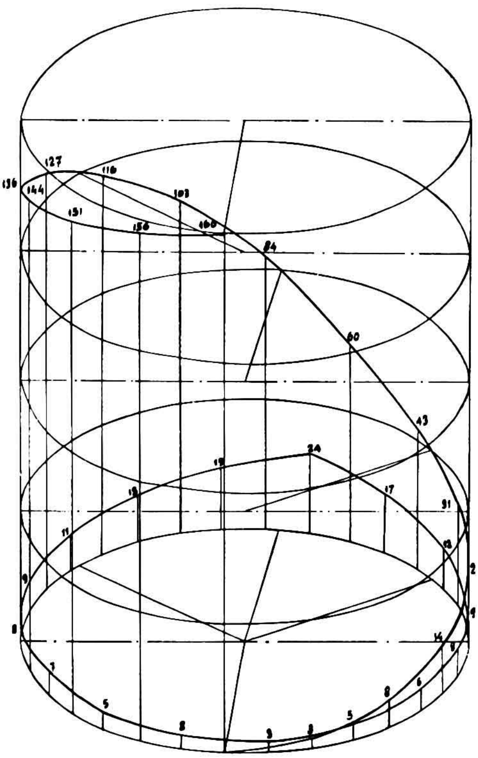

Fig. 2

cuts the horizontal plane, corresponding to a frequency equal to one half of the sample, gives the median, or rather a value measured on the circular axis that can be used as median. Similarly the points of the ogive

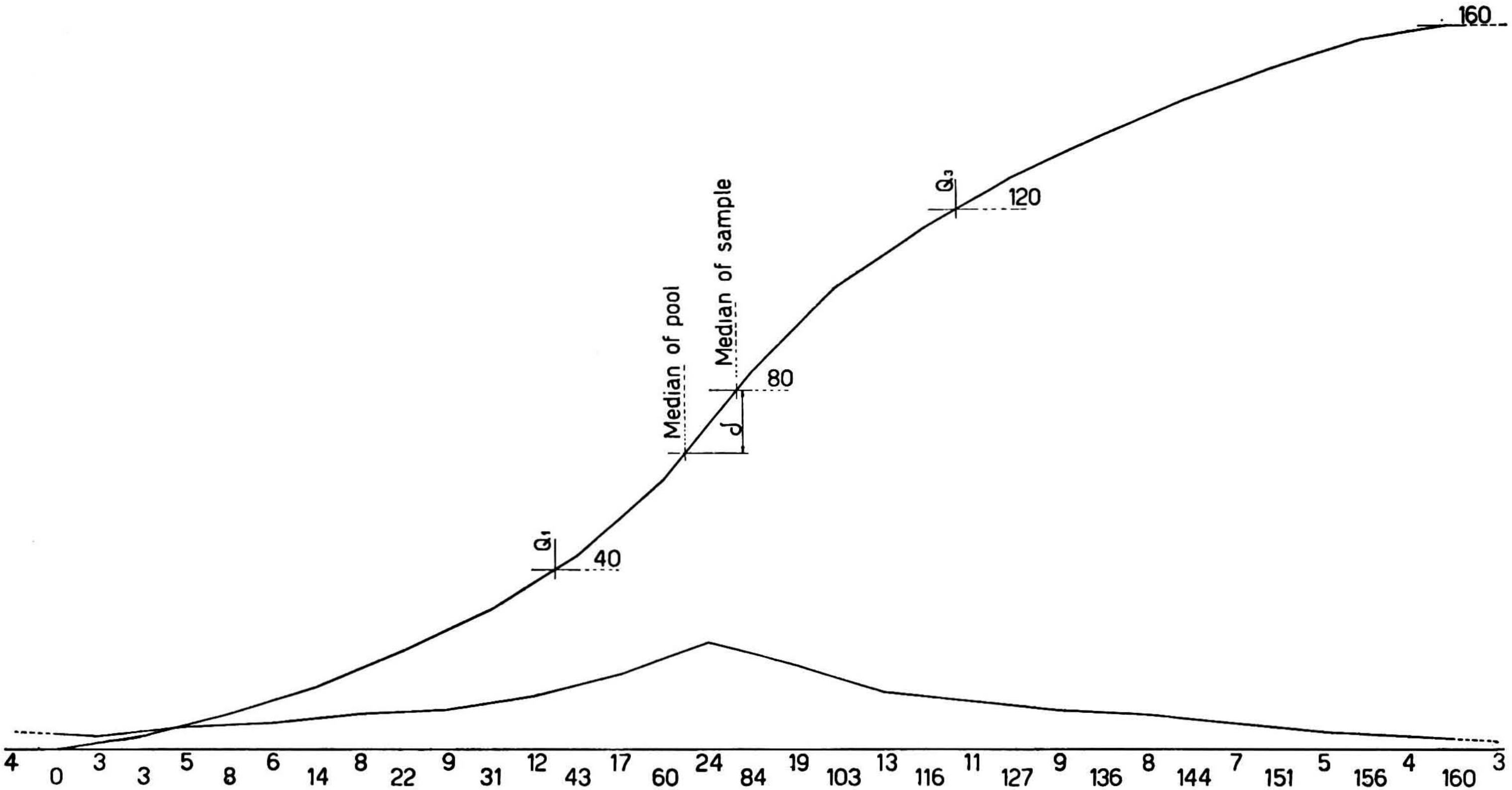
160
120
Q3
80
Median of sample
Median of pool
d
40
Q1
4 3 5 6 8 9 12 17 24 19 13 11 9 8 7 5 4 3
0 3 8 14 22 31 43 60 84 103 116 127 136 144 151 156 160

Fig. 3

corresponding to frequencies of $\frac{1}{4}$ and $\frac{3}{4}$ of the sample respectively, will yield values that can be regarded as the quartiles. This construction can suitably be carried out on ordinary graph paper (fig. 3). The resulting division of the sample in four equal parts can again be demonstrated on the circular axis (fig. 4).

In the comparison between two samples, the same procedure can

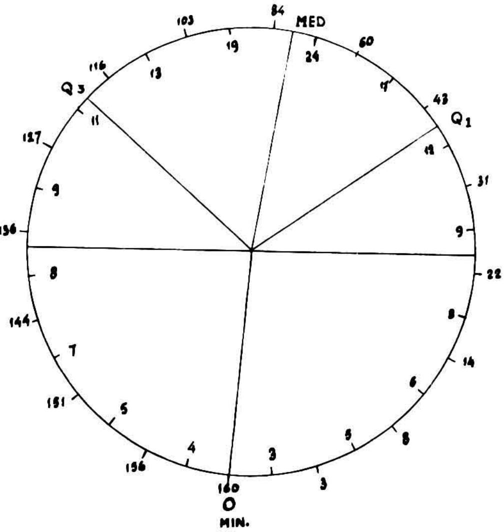

Fig. 4

be carried out with the pool of the samples. The median of the pool is then inserted in the graph representing one of the samples. The number of observations in this sample, between the median of the pool and the median of the sample proper will correspond to $|\delta|$ in the ordinary median test, which can hence be applied without any difficulty.

Similarly, the number of specimens between two quartiles (e.g. on the side of greater density) can be used for testing the significance of differences between two samples. To this end, the quartiles of the pool are again inserted in the graph representing one of the samples. The difference in number of observations between the quartiles of the pool, and the quartiles of the sample proper corresponds to $|\delta|$, so that the significance of differences can be tested with the same $|\delta|$ table.

In the event of no satisfactory significance being revealed, other starting points for the ogive curve may be tried.

LITERATURE

WESTENBERG, J., Significance test for median and interquartile range in samples from continuous populations of any form. Proc. Kon. Ned. Akad. v. Wetensch. Amsterdam, **51**, no. 2, 252—261 (1948).

————, A tabulation of the median test for unequal samples. Proc. Kon. Ned. Akad. v. Wetensch. Amsterdam, **53**, no. 1, 77—82 (1950).