ON THE THEORY OF RUNS WITH SOME APPLICATIONS TO QUALITY CONTROL[1]

By J. Wolfowitz

Columbia University

1. Recent developments in the theory of runs. The increasing number and importance of recent advances in the theory and statistical applications of runs may make a brief paper on the subject of some interest. The large volume of material and its wide dispersal, together with the limitations of space, will of necessity make these remarks far from exhaustive and complete.

I shall not define a run because new advances and applications of new criteria to new problems would probably soon render most definitions obsolete. Runs as used in statistics are best characterized by a philosophy and a technique rather than by the employment of any one specific device. What is always involved is the ordering of observations according to some characteristic and the resultant effect of this ordering on the ordering according to some other characteristic. For example, if the seats at a meeting of statisticians and engineers are numbered and occupied by m engineers and n statisticians, then if we list the numbers of the occupied seats in ascending order and replace each number by E or S according as the seat is occupied by an engineer or statistician, we shall have a sequence of $m + n$ elements, m E's and n S's. Thus, if $m = 7$ and $n = 6$, such a sequence might be

$$E\,E\,E\,S\,E\,E\,S\,S\,S\,E\,S\,S\,E.$$

If we were interested in knowing how well engineers and statisticians are acquainted with one another, we should find it of interest to study the runs of E's and S's in this sequence. Any subsequence of consecutive E's or S's which cannot be enlarged is called a run. Thus in the example above there is a run of E's of length 3, followed in order by a run of S's of length 1, a run of E's of length 2, a run of S's of length 3, a run of E's of length 1, a run of S's of length 2, and a run of E's of length 1. Runs of this kind are usually called runs of two kinds of elements. Naturally the characteristic according to which we order (in the example above, seat number) and the characteristic whose runs are observed (E or S) may be various. They ought in general to have a meaningful connection.

The order of observations has no value if it is known that the observations are independent and random from the same universe and one seeks to estimate a parameter of the universe. Many of the statistical problems treated in the literature are of this character. In quality control of manufactured articles one

[1] Revised from an expository address delivered at a joint meeting of the Institute of Mathematical Statistics and the American Society of Mechanical Engineers at New York, May 29, 1943, at the invitation of the program committee.