THE JOURNAL OF EXPERIMENTAL EDUCATION
(Volume 32, Number 1, Fall 1963)

AN APPLICATION OF SEQUENTIAL ANALYSIS TO PSYCHOLOGICAL RESEARCH AND NONPARAMETRICS

ROBERT SCHWITZGEBEL[1]
Science Committee on Programmed Experimentation
Cambridge, Massachusetts

ONE ESTABLISHED method of evaluating data as they are being collected is known as sequential analysis (1, 6, 7). An outstanding advantage of this technique is that it permits the experimenter to save considerable effort by discontinuing research as soon as a predetermined significance level is reached.

A more recent advance in methodology has been nonparametric techniques (3, 5). These techniques are quite useful in situations where one does not wish to make the usual assumptions about the normality of the population distribution. The sign test (5) is one example of a nonparametric which is extremely easy to use, has wide applicability, and is relatively powerful on sample sizes of 30 or less.

The purpose of this paper is to introduce the combination sequential analysis and nonparametric techniques ("sequential nonparametrics") to psychological researchers. This will be done in a very practical way, it is hoped, by presenting in detail one particular application. The application to be demonstrated is a special case of sequential nonparametrics appropriately used when we have two related samples and assume a binomial distribution. The nonparametric sign test which we mentioned earlier can be used in this instance, and when combined with sequential analysis we have what we shall call "sequential sign analysis."

The application of sequential analysis to data which are assumed to have a binomial distribution has already been explicated by the Columbia Statistical Research Group (6). However, this work is not readily available to most researchers and deals primarily with industrial inspection procedures rather than psychological investigations. The following presentation is intended to be useful even to the psychologist who is relatively unsophisticated in statistical methodology.

FUNCTION

Sequential sign analysis is applicable to the case of two related samples. It is uniquely useful in situations where the experimenter wishes to determine, during the course of his study, the degree of difference between two conditions in which one member of a pair of observations can be ranked only as "larger" or "smaller" than the other member of the pair. Sequential sign analysis can also be used where it is possible to specify quantitatively the difference within the pairs of observations, but as the name implies, it requires only that one note whether the sign of the difference is plus or minus.

Within every pair of observations, relevant variables should be matched if different groups are used, or else each subject should be used "as his own control." There is no assumption that the observed differences between the two conditions are normally distributed. However, it is assumed that the various pairs of observations are independent.

PROCEDURE

To use sequential sign analysis, these steps should be followed:

1. Choose a desired significance level, α. (The power, β, will automatically be fixed at the same level.[2])
2. Let X_1 be the scores under one condition

(or the "Before" treatment if the same subjects are tested twice). Let Y_1 represent the score under the other condition (or "After" treatment). Note the direction of the difference between the scores of every matched pair. Assign a plus to all differences in one direction and a minus to all differences in the other direction. If the scores of a matched pair are equal (the difference being zero), they are not included in the analysis and the total effective cases, N, is correspondingly reduced. The Null Hypothesis is that there will be an equal number of plus and minus signs.

3. Construct a graph making the vertical axis the number of possible plus signs, n. Let the horizontal axis represent the total number of matched pairs, N. See Figure 1.

4. Draw decision lines on the graph to represent the significance level you have chosen. The lines are drawn according to the formula presented in Table 1. "Intercept" refers to the point at which the decision line crosses the vertical axis, and "Slope" to the angle of inclination. Figure 1 shows the position of decision lines for the four most commonly chosen significance levels.

To illustrate how the formula in Table 1 was used to determine the .05 line, let us assume that we have a sample of 10 pairs. To be significant at .05 or below, the number of plus signs should equal 10 (.6309) ± 2.6807. In other words, nine positive cases are required if we are to reject a Null Hypothesis of no difference, and no more than three cases are allowed if we are to accept the Null with the same degree of confidence. These points (more exactly 8.990 and 3.628) can be marked on the graph by following the 10th vertical line of N upward until it is intersected by the third and by the ninth horizontal lines of n.

The other points we require are already given to us in the formula as the intercept. In the present case of the .05 line, these points are +2.6807 and -2.6807.

It is not absolutely necessary to use the formula in Table 1 if the researcher is able to reproduce Figure 1 or to construct on graph paper a figure of identical proportions.

5. Record data on the graph as they are collected. When a decision line is crossed, research may be discontinued and the Null Hypothesis accepted or rejected at the given significance level. On rare occasions, neither line will have been crossed after a large number of observations. The simplest procedure in this case is to accept the Null noting that it is done with a margin of error greater than represented by the decision lines. (It is possible, however, to draw decision lines at any significance and power level. The formula for this is given in the technical notes at the end of this paper.)

ILLUSTRATION OF PROCEDURE

In Table 2 are some data given by Mosteller (2). The tasks are to determine 1) whether the scores on examination B are in fact different than the scores on examination A, 2) whether scores B are greater than scores A, and 3) whether there is a real difference (i.e., $\alpha \leq .05$) between the test scores given a margin of ten points in either direction. Figure 1 shows the application of sequential sign analysis to this data.

1. Whenever a difference between the scores occurred, in either direction, it was noted "0" in Figure 1. Since none of the scores were equal, there was a corresponding and constant increase in the line recording these observations. One might have guessed from a casual examination of the scores that there would be some difference attributable to the test, but the sequential procedure shows how many such differences are required in order to reject the Null at a given significance level. After the 12th case, for example, we could have rejected the Null at the .01 level since the .01 line had been crossed. At the 17th case we could have stopped sampling and rejected the Null at a significance level below .001. Remember, however, that one elegant feature of sequential testing is that a significance level is determined in advance and held to until either the data meet the demand or sampling is discontinued.

2. The hypothesis tested was that the B scores are significantly larger than the A scores. The number of positive signs (i.e., when the B scores exceed the A as indicated by "X") is noted by comparing each pair of observations. For example, at the 10th observation there were seven differences recorded in support of the hypothesis. By the 16th case we could have rejected the Null Hypothesis of no difference and accepted our alternative hypothesis below the .05 significance level.

3. The task was to determine whether there was a difference of 10 units in either direction between the two groups of scores. This test, similar to the first, is two-tailed. If the number of positive occurrences, as indicated by "Δ" crosses the bottom decision line, we accept the Null. If they cross the upper line, then we reject the Null. In this instance the Null Hypothesis of equality of test scores within a 10-unit range could have been accepted at a significance level below .05.

CONCLUSION

A simple method has been outlined which combines sequential analysis and the nonparametric sign test. The possibility of computational errors in research cannot be completely avoided; however, methods such as sequential sign analysis can seem-

FIGURE 1

DISTRIBUTION AND SIGNIFICANCE LEVELS OF EXAMINATION SCORE DIFFERENCES

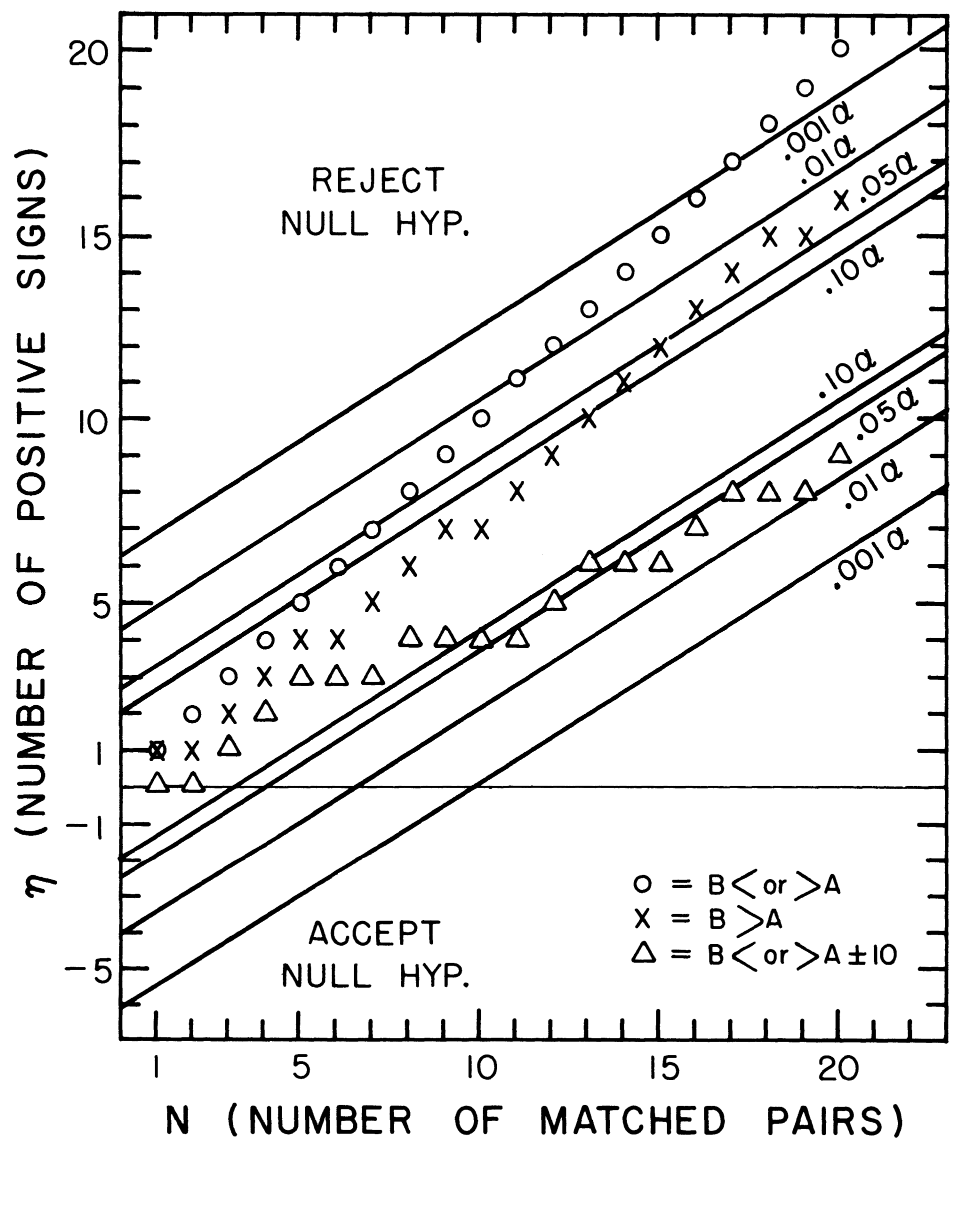

TABLE 1

FORMULAS FOR DECISION LINES OF SEQUENTIAL SIGN ANALYSIS

Significance and Power Levels	Number of + Signs (n)	Intercept	Slope· Cases
$\alpha = \beta = .001$	———	$\pm$ 6.287	+ .6309N
$\alpha = \beta = .01$	———	$\pm$ 4.204	+ .6309N
$\alpha = \beta = .05$	———	$\pm$ 2.681	+ .6309N
$\alpha = \beta = .10$	———	$\pm$ 2.000	+ .6309N

TABLE 2

EXAMINATION SCORES (FICTITIOUS)

Student	Test A	Test B	Student	Test A	Test B
1	87	96	11	91	97
2	92	91	12	80	92
3	66	80	13	52	75
4	72	85	14	84	92
5	75	91	15	72	95
6	96	95	16	70	88
7	85	89	17	66	90
8	82	94	18	82	90
9	69	78	19	81	85
10	70	68	20	75	88

ingly minimize such errors — especially where the researcher is pressed for time or must rely on relatively unskilled assistance. Once the graph is drawn or printed it requires little statistical sophistication beyond simple subtraction to determine significance levels.

Other advantages of sequential sign analysis include the ability to test for significance while the data are being collected, the absence of an assumption of population normality, and the opportunity to explicitly determine both the significance and power levels.

One limitation of this particular technique is its applicability only to matched pairs or twice-tested individuals. Another limitation is the possible loss of power if the researcher is fortunate enough to have interval or ordinal scale data, although it should be clear that he could still use the more conservative sign analysis if he desired.

TECHNICAL NOTES

The following notes are appended for those who may be interested in the derivation of the test or wish to draw decision lines at various significance levels. The Null Hypothesis of the sign test is: $p(X_a < X_b) = p(X_a > X_b) = \frac{1}{2}$ where X_a and X_b are measurements before and after treatment or under two conditions (5). If p is the probability of obtaining a positive difference on a single pair of observations, then the probability of obtaining n positive differences out of N trials when $p = \frac{1}{2}$ is: $P_n = \binom{N}{n}\left(\frac{1}{2}\right)^N$. It is this binomial expansion which permits us to connect the sign test to the established sequential technique for dealing with binomial data.

The linear parameters for sequential analysis have been determined by the Columbia Statistical Research Group (6) as:

$$d_1 = h_2 + sn \text{ (lower line)}$$

$$d_2 = h_2 + sn \text{ (upper line)}$$

where d represents in our case the number of plus signs; n, the number of pairs of scores sampled; h_1 and h_2, the intercepts; and s, the slope.

The intercepts and slope are given by the following formula (6; for alternate forms see reference 4):

$$h_1 = \frac{\log\left(\frac{1-\alpha}{\beta}\right)}{\log\frac{P_2}{P_1}\left(\frac{1-P_1}{1-P_2}\right)} \qquad h_2 = \frac{\log\left(\frac{1-\beta}{\alpha}\right)}{\log\frac{P_2}{P_1}\left(\frac{1-P_1}{1-P_2}\right)}$$

$$s = \frac{\log\left(\frac{1-P_1}{1-P_2}\right)}{\log\frac{P_2}{P_1}\left(\frac{1-P_1}{1-P_2}\right)}$$

where P_1 in the present statistic represents a true Null Hypothesis; P_2, the research hypothesis; α, the probability of rejecting P_1 or committing a Type I error; β, the probability of accepting P_2 or committing a Type II error. The intercepts, h_1 and h_2, are equal (except for sign) when, and only when $\alpha = \beta$. This is the special condition represented earlier in Table 1 of this paper. It would appear that the use of these formulas is one relatively simple way for even the amateur psychological researcher to control the significance level, the power level, and the absolute size of the differences he wishes to detect.

FOOTNOTES

1. The author is indebted to Dean W. Allen Wallis, Dr. Donald Fiske, and Dr. David Ricks for helpful criticisms of this paper in various stages of preparation.
2. The significance and power levels may be chosen independently by using formulas given in the technical notes at the end of this paper. As a matter of convenience, however, they are set equal in the test as described here. In most statistical techniques the power level is left unspecified, and the researcher, in practice, is unable to manipulate this factor. Thus, he may discover too late that he could not possibly obtain enough subjects to reject a given hypothesis, or on the other hand, that he has spent time and money on subjects he would not have needed.

REFERENCES

1. Fiske, D. W., and Jones, L. V. "Sequential Analysis in Psychological Research," Psychological Bulletin, LI (1954), pp. 264-75.
2. Mosteller, F. Syllabus for Social Problems 199, Harvard University, Spring 1959, problems section p. 12 (mimeographed).
3. Mosteller, F., and Bush, R. "Selected Quantitative Techniques," in G. Lindzey (Ed.), Handbook of Social Psychology, I (Cambridge, Massachusetts: Addison-Wesley, 1954), pp. 312-14.
4. Noether, G. "Two Sequential Tests Against Trend," Journal of the American Statistical Association, LI (1956), pp. 440-43.
5. Siegel, S. Nonparametric Statistics for the Behavioral Sciences (New York: McGraw-Hill, 1956).
6. Statistical Research Group, Sequential Analysis of Statistical Data: Applications (New York: Columbia University Press, 1945).
7. Wald, A. Sequential Analysis (New York: Wiley and Sons, 1947).