Editorial

Sequential Analysis in Nutrition Studies

In conducting nutrition surveys and various kinds of experimental work, there is a need for a relatively rapid method of evaluating the data as they are collected. Such a method would not only prevent wasted effort in collecting data after sufficient information is available to make a decision but would also prevent the termination of the work before enough was known to reach a definite conclusion. It would be extremely convenient to have the various hypotheses to be tested stated in advance and to be able to evaluate the results as they are collected. As soon as a decision was reached with regard to any particular hypothesis, within probability limits previously determined, the collection of such data could be stopped and additional effort expended in other more profitable directions.

In a survey, for example, blood samples might be collected for several determinations including hemoglobin and cholesterol. The hypothesis to be tested might be that the average hemoglobin value lies between 14 and 15 gm. per 100 ml. and that the mean serum cholesterol value is between 200 and 240 mg. per 100 ml. There is no reason to suspect that the same number of samples will be required to permit a judgment on each of the questions asked. In areas in which the population is relatively anemic, only a few samples would be required to determine whether or not the hypothesis with regard to hemoglobin was correct and the determination of hemoglobin could be dispensed with or an alternate hypothesis proposed. In other areas a decision with regard to the serum cholesterol level might be made with few samples, but many more might be required to reach a decision regarding the hemoglobin level.

Sequential analysis might well offer possibilities in this area. The technics involved were developed during World War II and were used to increase the efficiency of the inspection of products. The purpose of this paper is to describe briefly the situations in which sequential procedures might be applied and, through the use of an example, to provide a general understanding of the method and its potential value in the nutrition field.

Several recent texts on statistics have chapters devoted to the subject. The simplest description we have found, including an example similar to many of those of interest to nutritionists, appears in the text of Rosander.[1] The example is on page 659 and reference to this will make the development of our example clearer. The reader who is interested in the theoretical basis, development and limitations of the technic is cautiously referred to Wald's "Sequential Analysis,"[2] the standard book in this field. The caution is due to the rather overwhelming mathematics involved. Very likely, the average person will be content with a general understanding of the method and its usefulness and will consult better qualified people to be assured he does not go astray in application.

Sequential procedures have been developed to test most of the hypotheses generally found in nutrition studies. In surveys, for example, they make it possible to decide whether the proportion of "successes" (however defined) in a sample differs significantly from an expected proportion or whether the average value of a measurement is significantly higher or lower than a so-called "normal" value. In experimental studies in which paired observations are used, differences between proportions or between mean values can be readily tested for significance.

Although various types of questions can be answered through the use of sequential

analysis, the actual test procedure is quite similar from situation to situation. In each case appropriate observations are made one by one and their sum analyzed consecutively as each new observation is made. It is this approach that makes the sequential procedures differ from classic tests, such as the t test, and that, quite often, makes it possible to reach a decision with far fewer observations than would otherwise be used. There are other common elements in all sequential tests. What they are is probably best explained through an example.

The example we would like to discuss has to do with deciding whether the mean value in a sample differs significantly from an expected or "normal" value. Evidence[3] indicates that adult men in the United States have an average serum cholesterol value of approximately 240 mg. per 100 ml. with a standard deviation of approximately 48. The standard deviation of other populations with quite different mean cholesterol values appears to be of the same order of magnitude.[4] This is fortunate since this type of sequential test assumes that the standard deviation is known in advance. In this example, cholesterol determinations for four groups will be used, namely, a sample of Ethiopian adults, of preparatory school boys between fourteen and eighteen years of age, a random group of adult men, and a group from a cardiac clinic who might be suspected of hypercholesterolemia. For each group, the null hypothesis is that the mean does not differ significantly from the expected mean of 240 mg. per 100 ml. Before testing begins, the experimenter must decide on practical grounds how much smaller or larger the sample mean must be to indicate a real difference. For example, a mean of 239 is certainly less than 240, but the difference is so small that it would probably not be considered an important one to detect. Let us decide for our example that the means must differ by at least 25 mg. per 100 ml. from 240 in order to be of practical importance. With this range of acceptable values specified, we then test whether the mean in question is at least as small as 215 or as large as 265.

In sequential analysis, as in all significance testing, there is a possibility of drawing the wrong conclusion from the sample results. To control this potential error, the experimenter must specify in advance the risk of being wrong that he is willing to tolerate. There are actually two ways of wrongly rejecting the null hypothesis in our example. First, we may conclude the mean is as large as 265 when, in fact, it is as small as 215. This risk of doing this is called alpha. Secondly, we may conclude the mean is as small as 215 when it is really as large as 265. This risk is called beta. For our example, we will specify that both alpha and beta should be 2 per cent. That is, we are willing to be wrong two times out of 100 in rejecting the null hypothesis either above or below our specified range.

With this much information available, it is possible to construct boundary lines which will specify for each sample size how large (or small) the sum of the cholesterol values must be to conclude that the mean of the group involved is as large as 265 (or as small as 215).

As given in the example of Rosander,[1] the formulas for these lines, in an example like this, are:

$$y_1 = \Sigma x = \frac{-bs^2}{\bar{x}_2 - \bar{x}_1} + \left(\frac{\bar{x}_1 + \bar{x}_2}{2}\right) n$$

$$y_2 = \Sigma x = \frac{as^2}{\bar{x}_2 - \bar{x}_1} + \left(\frac{\bar{x}_1 + \bar{x}_2}{2}\right) n$$

where $\bar{x}_1$ and $\bar{x}_2$ are the lower and upper means of interest, namely 215 and 265

s = Expected standard deviation, namely 48

n = Number of observations

a = $\log_e \frac{1-\beta}{\alpha}$ α: Risk of wrongly accepting $\bar{x}_2$, namely, 0.02

b = $\log_e \frac{1-\alpha}{\beta}$ β: Risk of wrongly rejecting $\bar{x}_2$, namely, 0.02

For our example, the lines, which are plotted in Figure 1, become:

$$y_1 = \Sigma x = -176.2 + 240n$$

$$y_2 = \Sigma x = +176.2 + 240n$$

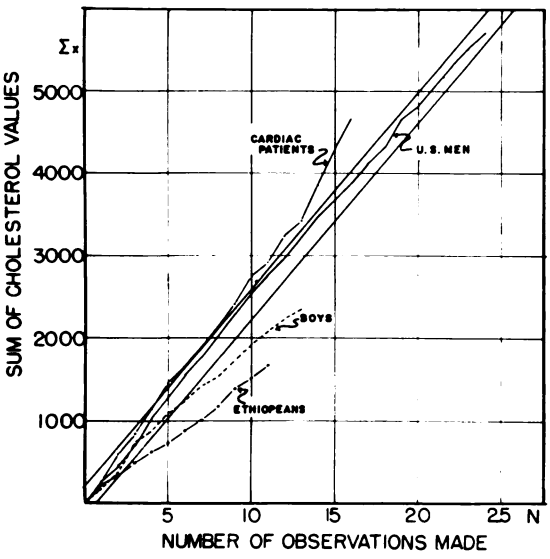

FIG. 1. The test chart for the evaluation of average serum cholesterol values.

These have been plotted in Figure 1, the points being obtained by substituting various values for n in the equations.

It is worth a comment here that these lines do not depend in any way on observations from the four samples under investigation. This means that they can be used by anyone interested in comparing the cholesterol level of a particular group with the average value for adult American males, provided, of course, they are willing to accept 215 and 265 as reasonable dividing values and are willing to be wrong 2 per cent of the time in either direction.

The test procedure itself is a simple one. As each observation is made, i.e., as each cholesterol level is determined, it is added to the preceding ones and their sum is plotted on Figure 1. As long as the sum remains between the boundary lines, observations continue to be made. As soon as the sum crosses a line, either above or below, testing stops and the appropriate conclusion is drawn.

For each of our four groups, successive values as they were determined and the cumulative sums are given in Table I. The sample mean values are included for information but not used in the test. The sums are plotted in Figure 1.

For the Ethiopians, the sum of the values rapidly falls below the limits established by the hypothesis, indicating that their average cholesterol level is significantly lower than that of American males and is at least as small as 215. In fact, only three determinations needed to be made in this group since their sum at that point fell below the boundary line. For illustrative reasons only, a few more observations were made and their sum recorded to show that the difference becomes more striking as the sample size increases.

The prep-school boys also have an average cholesterol level that is significantly low, but it was necessary to observe six or seven boys before being able to draw this conclusion.

The cardiac patients can be defined as hypercholesterolemic, above 265 mg. per 100 ml. by the tenth observation although only after the fourteenth is this unequivocal.

Finally, the random sample of adult males shows no tendency to depart from the boundary lines. This is to be expected since they are a sample from the population with which they are being compared. In a situation like this, namely the sum remaining within the lines as the sample size increases indefinitely, there are methods available to determine the maximum number of observations that need to be made.

Other sequential tests differ from the one illustrated in the actual statistics on which the test is made and in the formulas for the boundary lines, but all are based on the use of such boundary lines. They have in common the need for specifying a range of acceptable values, such as the range from 215 to 265 mg. per 100 ml., and for specifying the risk that will be tolerated for being wrong in either direction, such as α and β both 0.02. Most of the methods of sequential analysis require a reasonable estimate of the underlying standard deviation and, in addition, assume that the distribution involved is essentially normal.

Before extensive use of the method can be made, it will be necessary to know more about the nature of the distributions encountered with various tests, and under varying conditions, and the degree to which errors in the estimated standard deviations influence the

TABLE I
Serum Cholesterol Data from Four Samples Tested by Sequential Analysis

Sample No.	Ethiopians		Boys		Adult Men		Cardiac Patients	
	Serum Cholesterol (mg. %)	Σx	Serum Cholesterol (mg. %)	Σx	Serum Cholesterol (mg. %)	Σx	Serum Cholesterol (mg. %)	Σx
1	227	227	187	187	255	255	258	258
2	103	330	196	383	189	444	338	596
3	163	493	345	728	272	716	249	845
4	134	627	145	873	308	1,024	298	1,143
5	93	720	199	1,072	248	1,272	307	1,450
6	153	873	157	1,229	279	1,551	193	1,643
7	129	1,002	179	1,408	230	1,781	236	1,879
8	163	1,165	126	1,534	238	2,019	248	2,127
9	227	1,392	203	1,737	267	2,286	275	2,402
10	122	1,514	185	1,922	269	2,555	355	2,757
11	155	1,669	171	2,093	222	2,777	214	2,971
12	...	...	137	2,230	206	2,983	279	3,250
13	...	...	122	2,352	254	3,237	179	3,429
14	...	...	...	...	247	3,484	428	3,857
15	...	...	...	...	183	3,667	424	4,281
16	...	...	...	...	228	3,895	389	4,670
17	...	...	...	...	230	4,125	...	...
18	...	...	...	...	187	4,312	...	...
19	...	...	...	...	340	4,652	...	...
20	...	...	...	...	160	4,812	...	...
21	...	...	...	...	246	5,058	...	...
22	...	...	...	...	245	5,303	...	...
23	...	...	...	...	195	5,498	...	...
24	...	...	...	...	193	5,691	...	...
Mean	151.7		180.9		237.1		291.9	

applicability of the test. Too little attention has been paid in the past to this statistical parameter. It is clear that the data are often markedly skewed[5] and some transformation may be needed before the data can be properly treated. However, in many laboratories data, which have been collected in the past under varying conditions, are now available. These might be used for retrospective tests of the method and current surveys could provide tests in the field. Preference for any method, statistical or otherwise, must depend upon its performance under actual usage.

Finally, an advantage of sequential analysis might be that it requires the formulation of the hypotheses to be tested in rather exact form prior to the collection of the data. Most nutrition surveys to date have been performed with only general questions in mind and a "let's see what happens" attitude. This is often useful and necessary. In many other situations, however, particularly in which standard methods are applied, we should have reached the degree of maturity where we know what we want to know, the accuracy with which we wish to make the decision, and desire an answer with a minimum expenditure of time and energy. It is often true that the quality of the answer obtained depends upon the question asked.

D. M. HEGSTED, PH.D.
and M. E. DROLETTE, M.P.H.
Harvard School of Public Health
Boston, Massachusetts

REFERENCES

1. ROSANDER, A. C. Elementary Principles of Statistics. New York. 1951. D. Van Nostrand Company.

2. WALD, A. Sequential Analysis. New York. 1947. John Wiley and Sons, Inc.
3. LAWRY, E. Y., MANN, G. V., PETERSON, A., WYSOCKI, A. P., O'CONNELL, R. and STARE, F. J. Cholesterol and beta-lipoproteins in the serums of Americans. *Am. J. Med.*, 21: 605, 1957.
4. SCHRIMSHAW, N. S., TRULSON, M., TEJADA, C., HEGSTED, D. M. and STARE, F. J. Serum lipoprotein and cholesterol concentrations; comparison of rural Costa Rican, Guatemalan and United States populations. *Circulation*, 15: 805, 1957.
5. HEGSTED, D. M., GERSHOFF, S. N., TRULSON, M. F. and JOLLY, D. H. Variation in riboflavin excretion. *J. Nutrition*, 60: 581, 1956.