Therefore

$$S_{2p+1}(\alpha, \alpha) = 0. \tag{18}$$

When n is an integer, either $n + 1$ or $n + 2$ is odd. Therefore when (15) holds, one of either (7) or (8) will be satisfied identically if we take $\beta = \alpha$. The other may then be solved for α.

As an example, suppose one had the moments $\mu_0 = 1$, $\mu_1 = \frac{1}{2}$, $\mu_2 = \frac{7}{24}$, $\mu_3 = \frac{3}{16}$, $\mu_4 = \frac{31}{240}$, and wished to obtain an $f(x)$ such that $f(0) = 0, f(1) = 0$. In this case $n = 2$, and (15) is satisfied. It follows that (7) is satisfied identically when $\beta = \alpha$, and (8) gives

$$\frac{\Gamma(2\alpha + 5)}{\Gamma(\alpha + 1)} + 4\,\frac{\Gamma(2\alpha + 6)}{\Gamma(\alpha + 2)}\left(-\frac{1}{2}\right) + 6\,\frac{\Gamma(2\alpha + 7)}{\Gamma(\alpha + 3)}\left(\frac{7}{24}\right) + 4\,\frac{\Gamma(2\alpha + 8)}{\Gamma(\alpha + 4)}\left(-\frac{3}{16}\right) + \frac{\Gamma(2\alpha + 9)}{\Gamma(\alpha + 5)}\left(\frac{31}{240}\right) = 0.$$

This easily reduces to

$$1 - 4\,\frac{\alpha + 5/2}{\alpha + 1} + 7\,\frac{(\alpha + 5/2)(\alpha + 3)}{(\alpha + 1)(\alpha + 2)} - 6\,\frac{(\alpha + 5/2)(\alpha + 7/2)}{(\alpha + 1)(\alpha + 2)} + \frac{31}{240}\,\frac{(\alpha + 5/2)(\alpha + 7/2)}{(\alpha + 1)(\alpha + 2)} = 0,$$

which reduces to the quadratic

$$4\alpha^2 - 6\alpha + 5 = 0,$$

from which

$$\alpha = \beta = 3/4 \pm (1/4)\sqrt{11}\,i. \tag{19}$$

These may be substituted into (4)–(6) to complete the solution.

REFERENCE

[1] G. Szegö, *Orthogonal Polynomials*, Amer. Math. Soc. Colloquium Pub., No. 23, 1939.

CONSISTENCY OF SEQUENTIAL BINOMIAL ESTIMATES

By J. Wolfowitz

Columbia University

The notion of consistency of an estimate, introduced by R. A. Fisher, applies to a sequence of estimates which converge stochastically, with boundlessly increasing sample size, to the parameter (or parameters) being estimated. Each estimate is a function of a sample of observations, the number in each sample being determined independently of the observations themselves. In sequential estimation, on the other hand, the number of observations is itself a chance variable, determined by the sequence of observations and the application to them of a rule which may be part of a sequential test. In what follows we shall consider that the operation of sequential estimation *is* associated with a sequential test.[1]

The advantage of using consistent estimates is such as to suggest extension of the idea of consistency to sequential estimation. In the present paper we shall be concerned only with the estimation of a binomial probability (p, say). The obvious extension is that a sequence of estimates, each with its associated test, is consistent if the estimates converge stochastically to p.

Since the number of observations required by a sequential test is a chance variable, a parallel to the classical sequence of samples of increasing size would be a sequence of sequential tests whose average (in some sense) sample sizes increase without limit. It seems reasonable to associate only such a sequence of estimates with this sequence of tests as will converge stochastically to p, i.e., be consistent.

Let z be a chance variable which takes the distinct values c_1 and c_2 with probabilities p, $0 < p < 1$, and $q = 1 - p$, respectively. Let $z_1, \cdots, z_n$ be a sequence of independent observations on z which terminates with the nth according to the specific sequential test under consideration. Denote by x and y, respectively, the number of observations c_2 and c_1 in this sequence. Then x, y and $n = x + y$ are all chance variables. The couple $g = (x, y)$ is called a boundary point of index n (see [1]). The sequence of observations which terminates at g is called a path. Let $k(g)$ denote the number of paths which terminate at g, and let $k^*(g)$ denote the number of these paths whose first observation is c_1. The "points" on the various paths together with all the points g constitute the "region" under discussion.

Let $P\{n = j\}$ denote the probability of the relation in braces. If

$$\sum_{j=1}^{\infty} P\{n = j\} = 1,$$

the region is called closed. Only closed regions will be considered below, so that this assumption will henceforth be made without explicit formulation. It has been shown by Girshick, Mosteller, and Savage [1], that $p(g) = k^*(g)/k(g)$ is an unbiased estimate of p for any closed region R, i.e.,

$$\Sigma\, p(g)k(g)p^y q^x \equiv p,$$

where the summation takes place over all the boundary points g of R. For many important regions this estimate is the unique unbiased estimate.

Let there be given an infinite sequence of sequential tests with each of which we associate the estimate $p(g)$. Consider the ith one of these, and let n_{0i} be the smallest number of observations required for a decision, i.e., n_{0i} is the smallest

[1] Really all that is required is a rule for terminating the observations such that its region R is closed (see below). However, we defer to conventional statistical usage in referring to "tests."

value of j for which $P\{n = j\} \neq 0$. The theorem proved below asserts that if n_{0i} approaches infinity with i the estimate $p(g)$ converges stochastically to p. To put it in other words: if $T_1, T_2, \cdots$ is the sequence of tests, and ϵ_1 and ϵ_2 are arbitrarily small positive numbers, there exists a positive number $J(\epsilon_1, \epsilon_2)$ such that, for all T_i such that $i > J$,

$$P\{| p(g) - p | > \epsilon_1\} < \epsilon_2,$$

when $n_{0i} \to \infty$. An important example of such a sequence is that of the Wald sequential binomial tests [2] obtained as follows: Let $\alpha_1, \alpha_2, \cdots, \alpha_i \cdots$ and $\beta_1, \beta_2, \cdots, \beta_i \cdots$, be two sequences of positive numbers all of which are less than $\frac{1}{2}$ and which approach zero as $i \to \infty$. Let p_0 and p_1, $0 < p_0 < p_1 < 1$, be two fixed numbers,

$$c_1 = \log \frac{p_1}{p_0}, \qquad c_2 = \log \frac{(1 - p_1)}{(1 - p_0)}, \qquad Z_j = \sum_{k=1}^{j} z_k.$$

Finally let the rule for terminating the process of drawing observations be as follows for the ith test T_i: The process of drawing observations terminates at the smallest integer n for which either

$$Z_n \geq \log \frac{1 - \beta_i}{\alpha_i} \quad \text{or} \quad Z_n \leq \log \frac{\beta_i}{1 - \alpha_i}.$$

Since $(1 - \beta_i)/\alpha_i \to \infty$ and $\beta_i/(1 - \alpha_i) \to 0$ while c_1 and c_2 are constant, it is evident that the hypothesis of the theorem is satisfied.

The property of being unbiased is not generally considered an indispensable characteristic of an optimum estimate, while consistency is generally so regarded. Our theorem shows that $p(g)$ enjoys the latter property with respect to important sequences of sequential tests.

THEOREM: *Let $T_1, \cdots, T_i, \cdots$ be a sequence of sequential binomial tests. For the ith test T_i let n_{0i} be the smallest integer such that $P\{n = n_{0i}\} \neq 0$. Finally let $n_{0i} \to \infty$ as $i \to \infty$. Then $p(g)$ converges stochastically to p as $i \to \infty$.*

PROOF: For typographic simplicity we shall use n_0 as the designation of the generic element of the sequence $n_{01}, n_{02}, \cdots$. No confusion will be caused thereby.

Let $n' = n_0 - 1$, and $\delta_1 > 0$ and $\delta_2 > 0$ be arbitrarily small fixed numbers. Let $k'(g)$ be the number of paths which end at the point g and are such that $| y'/n' - p | < \delta_1$, where y' is the number of observations c_1 among the first n' observations. We then have

LEMMA 1. *For n_0 sufficiently large*

(1) $$\sum_{g \epsilon B} k'(g) p^y q^x > 1 - \delta_2$$

where B is the set of boundary points of R.

PROOF: Consider the totality $\{h\}$ of all points $h = (x', y')$, with $x' + y' = n'$. Here x' and y' denote, respectively, the number of observations c_2 and c_1 in the sequence of the first n' observations on z. Let $k_0(h)$ denote the number of paths to h. Let C denote the set of points h such that $| y'/n' - p | < \delta_1$. If n_0 is large enough we have, by the law of large numbers,

$$\sum_{h \epsilon C} k_0(h) p^{y'} q^{x'} > 1 - \delta_2 .$$

Let $k(h, g)$ be the number of paths from h to g. From Theorem 2′ of [3] it follows that

$$\text{(A)} \qquad \sum_{g \epsilon B} k(h, g) p^y q^x = p^{y'} q^{x'}.$$

Also from the definitions of the various symbols involved it readily follows that

$$k'(g) = \sum_{h \epsilon C} k_0(h) k(h, g).$$

Hence

$$\sum_{g \epsilon B} k'(g) p^y q^x = \sum_{g \epsilon B} \Big(\sum_{h \epsilon C} k_0(h) k(h, g)\Big) p^y q^x = \sum_{g \epsilon B} \Big(\sum_{h \epsilon C} k_0(h) k(h, g) p^y q^x\Big)$$

$$= \sum_{h \epsilon C} k_0(h) \Big(\sum_{g \epsilon B} k(h, g) p^y q^x\Big) = \sum_{h \epsilon C} k_0(h) p^{y'} q^{x'} > 1 - \delta_2 .$$

This proves Lemma 1.

Let $\xi(g) = [k(g) - k'(g)]k(g)$. Thus $\xi(g)$ is a chance variable, being a function of the chance point g.

LEMMA 2. *Let δ_3 and δ_4 be arbitrarily small positive numbers. For n_0 sufficiently large*

$$\text{(2)} \qquad P\{\xi(g) \leqq \delta_3\} > 1 - \delta_4 .$$

PROOF: If (2) were not true, we would have

$$\text{(3)} \qquad E\,\frac{k'(g)}{k(g)} = \Sigma\, k'(g) p^y q^x \leq (1 - \delta_4) + (1 - \delta_3)\delta_4 = 1 - \delta_3 \delta_4 .$$

Choose the δ_2 of Lemma 1 so that $\delta_2 < \delta_3 \delta_4$. For some large value of n_0 we would then have a contradiction between (1) and (3). This proves the lemma.

Let g be any boundary point. Consider any path whose y' is such that $| y'/n' - p | < \delta_1$; let us call such a path one of type T. Consider the terminal sequence S of this path,

$$S : z_{n_0}, z_{n_0+1}, \cdots, z_n$$

This sequence, together with $g = (x, y)$, uniquely determines y'. Any permutation of y' elements c_1 and $n' - y' = x'$ elements c_2 may serve as the initial sequence of n' observations of a path which terminates at g and has the terminal sequence S. For no boundary point is of index smaller than n_0, so that under permutation of the first n' observations a path remains a path, i.e., the process of taking observations will not terminate prematurely as a result of the permuting of the elements. Of these permutations a proportion y'/n' begin with the element c_1. We deal in this manner with all the different terminal sequences of the paths of type T which end at g. Let $k^{*\prime}(g)$ be the number of these which begin with c_1. We obtain

LEMMA 3. *For all g such that* $k'(g) \neq 0$

$$\left| \frac{k^{*\prime}(g)}{k'(g)} - p \right| < \delta_1.$$

Putting Lemmas 2 and 3 together we have

LEMMA 4. *As* $n_0 \to \infty$, $k^{*\prime}(g)/k(g)$ *converges stochastically to* p.

Now it follows in a manner similar to that of Lemma 2 that, as $n_0 \to \infty$, $k^{*\prime}(g)/k^*(g)$ converges stochastically to one. This, together with Lemma 4, proves the theorem.

REFERENCES

[1] M. A. GIRSHICK, FREDERICK MOSTELLER, AND L. J. SAVAGE, "Unbiased estimates for certain binomial sampling problems, with applications," *Annals of Math. Stat.*, Vol. 17 (1946), pp. 13–23.
[2] A. WALD, "Sequential tests of statistical hypotheses," *Annals of Math. Stat.*, Vol. 16 (1945), pp. 117–186.
[3] J. WOLFOWITZ, "On sequential binomial estimation," *Annals of Math. Stat.*, Vol. 17 (1946), pp. 489–492.