values p_1 and p_2 for a given n and c can be calculated, using a table of the 5 per cent points of the F (variance ratio) distribution. We may take

$$n_1 = 2(n - c)$$

$$n_2 = 2(c + 1)$$

$$F_1 = F(n_1, n_2)$$

$$F_2 = F(n_2, n_1).$$

Then

$$p_1 = \frac{n_2}{n_2 + n_1 F_1}$$

and

$$p_2 = \frac{n_2 F_2}{n_1 + n_2 F_2},$$

utilizing a property of the F distribution pointed out in [3], page 2.

REFERENCES

[1]. A. Wald, "Sequential tests of statistical hypotheses", *Annals of Math. Stat.*, Vol. 16 (1945), pp. 117-186.

[2]. H. F. Dodge and H. G. Romig, "A method of sampling inspection", *Bell System Technical Journal*, Vol. 8 (1929), pp. 613–631.

[3]. R. A. Fisher and F. Yates, *Statistical Tables*, 2nd edition, Oliver and Boyd, 1943.

ON AN EQUATION OF WALD

By David Blackwell

Howard University

Let $X_1, X_2, \cdots$ be a sequence of independent chance variables with a common expected value a, and let $S_1, S_2, \cdots$ be a sequence of mutually exclusive events, S_k depending only on $X_1, \cdots, X_k$, such that $\sum_{k=1}^{\infty} P(S_k) = 1$. Define the chance variables $n = n(X_1, X_2, \cdots) = k$ when S_k occurs and $W = X_1 + \cdots + X_n$. We shall consider conditions under which the equation

$$E(W) = aE(n), \tag{1}$$

due to Wald [3, p. 142], holds.

This equation has various interpretations:

A. n may be considered as defining a sequential test on the X_i. If a and $E(W)$ are known, (1) may be used to determine $E(n)$, the expected number of observations required by the sequential test, [3, p. 142 et seq].

B. n may be considered as representing a gambling system, i.e. it represents the point at which a player decides to stop. W then represents his winnings, and (1), in the special case $a = 0$, says that, if each play is a fair game, then the system leads to a fair game.

C. n may be considered as the duration of a random walk. The meaning of W and (1) is obvious.

More exactly, we shall investigate conditions on X_i under which (1) holds for every test n of finite expected value. Our results, Theorems 1 and 2, are that (1) holds if the X_i have identical distributions, or if they are uniformly bounded. Theorem 1 is a generalization of a result of Wald [3, p. 142].

The test n may be considered as a test on the variables $Y_i = X_i - a$. Then $W' = Y_1 + \cdots + Y_n = W - na$, so that $E(W') = 0$ is equivalent to (1) for tests of finite expected value. Thus it is no loss of generality to assume $a = 0$ and to seek conditions under which $E(W) = 0$. We remark that if $E(n)$ does not exist, then $E(W)$ need not be zero. For example define $X_i = \pm 1$ with probability $\frac{1}{2}$, and n as the smallest integer k for which $X_1 + \cdots + X_k = 1$. Then $E(W) = 1$. (It follows from Theorem 1 or 2 that $E(n)$ cannot exist, which can also be shown directly.)

Theorem 1. *If $X_1, X_2, \cdots$ have identical distributions, $E(X_i) = 0$, $E(n) < \infty$, then $E(W) = 0$.*

Proof: Define chance variables n_k inductively as follows: $n_1 = n$. Supposing $n_1, \cdots, n_k$ to be defined, define $n_{k+1} = n(X_{n_1+\cdots+n_k+1}, X_{n_1+\cdots+n_k+2}, \cdots)$ i.e. $n_1, n_2, \cdots$ are the successive values of n obtained by iterating the test. Then

$$P(n_1, \cdots, n_k; n_{k+1} = j) = P(S_j). \tag{2}$$

For the event $\{n_1 = a_1, \cdots, n_k = a_k\} = R$ depends only on $X_1, \cdots, X_{a_1+\cdots+a_k}$, while under the hypothesis R the event $\{n_{k+1} = j\}$ coincides with the event $S = \{n(X_{a_1+\cdots+a_k+1}, \cdots) = j\}$. Thus $P_R(S) = P(S)$. Finally $P(S) = P(S_j)$ since S is defined by imposing the same conditions on $X_{a_1+\cdots+a_k+1}, \cdots$ that S_j imposes on $X_1, \cdots, X_j$. (2) shows inductively that $n_1, n_2, \cdots$ are defined everywhere and are mutually independent with identical distributions. Now define $W_k = X_{n_1+\cdots+n_{k-1}+1} + \cdots + X_{n_1+\cdots+n_k}$. A similar argument shows that $W_1 (= W), W_2, \cdots$ are also independent variables with identical distributions. The strong law of large numbers [2, p. 488] asserts that, with probability one,

$$\frac{X_1 + \cdots + X_N}{N} \to 0 \qquad \text{as } N \to \infty. \tag{3}$$

It follows that, with probability one,

$$\frac{W_1 + \cdots + W_k}{n_1 + \cdots + n_k} \to 0 \qquad \text{as } k > 0.$$

$$\text{For if } \left| \frac{W_1 + \cdots + W_k}{n_1 + \cdots + n_k} \right| > \epsilon \qquad \text{for an infinite number of } k,$$

$$\text{then } \left| \frac{X_1 + \cdots + X_N}{N} \right| > \epsilon \qquad \text{for an infinite number of } N,$$

which by (3) is an event of probability zero. Also from the strong law of large numbers $\frac{n_1 + \cdots + n_k}{k} \to E(n)$ with probability one. Then

$$\frac{W_1 + \cdots + W_k}{k} = \left(\frac{W_1 + \cdots + W_k}{n_1 + \cdots + n_k}\right)\left(\frac{n_1 + \cdots + n_k}{k}\right) \to 0$$

with probability one. It follows from the converse of the strong law of large numbers [2, p. 488] that $E(W_i) = E(W) = 0$.

Write $S_1 + \cdots + S_k = U_k$, $C(U_k) = V_k$ so that $V_k = \{n > k\}$. Then (a) V_k depends only on $X_1, \cdots, X_k$, (b) $V_1 \supset V_2 \supset \cdots$, $P(V_k) \to 0$. Conversely any sequence of sets V_k satisfying (a) and (b) defines a sequential test on X_i; define $n = k$ on $V_{k-1}C(V_k)$. Moreover $E(n) < \infty$ if and only if (c) $\sum_{k=1}^{\infty} P(V_k)$ converges [1, p. 297]. Now

$$E(W) = \lim_{N\to\infty} \sum_{k=1}^{N} \int_{S_k} (X_1 + \cdots + X_k)\, dP = \lim_{N\to\infty} \sum_{k=1}^{N} \int_{S_k} (X_1 + \cdots + X_N)\, dP$$

$$= \lim_{N\to\infty} \int_{U_N} (X_1 + \cdots + X_N)\, dP = -\lim_{N\to\infty} \int_{V_N} (X_1 + \cdots X_N)\, dP.$$

This establishes the following

LEMMA: *If* $E(X_i) = 0$, *then* $E(W) = 0$ *for every test of finite expected value if and only if for every sequence of sets* V_N *satisfying* (a), (b), (c),

$$\int_{V_N} (X_1 + \cdots + X_N)\, dP \to 0.$$

From this condition we obtain easily

THEOREM 2. *If* $E(X_i) = 0$, $|X_i| < M$, $E(n) < \infty$, *then* $E(W) = 0$.

PROOF: If V_N is a sequence of sets satisfying (a), (b), (c), then

$$\left| \int_{V_N} (X_1 + \cdots + X_N)\, dP \right| < MNP(V_N).$$

Now the series $\Sigma\, P(V_N)$ is a convergent series with decreasing positive terms. It is well known that under these conditions $NP(V_N) \to 0$. It follows from the lemma that $E(W) = 0$.

The question of finding sufficient conditions for $E(W) = 0$ more general than those given in Theorems 1 and 2 is of interest. The bare condition $E(X_i) = 0$ is not sufficient, as the following example (which is simply the system of doubling the stake) shows: $X_i \pm 2^i$ with probability $\frac{1}{2}$, n is the smallest integer k for which $X_k > 0$. A simple computation shows $E(n) = E(W) = 2$. It is well known that the expected amount of capital required for the above system is infinite. That this is generally true for such systems is shown by the following theorem, in which no hypothesis is made concerning the existence of $E(n)$.

THEOREM 3. *If* $E(X_i) = 0$, $E(W) > 0$, *then* $E(Z) = -\infty$, *where*

$$Z = \min_{k \leq n} (X_1 + \cdots + X_k).$$

PROOF: It follows from the proof of the lemma that

$$\int_{V_N} (X_1 + \cdots + X_N)\, dP \rightarrow -E(W).$$

Now on V_N, $Z \leq (X_1 + \cdots + X_N)$. Hence

$$\lim_{N\to\infty} \int_{V_N} Z\, dP \leq -E(W).$$

Thus $E(Z)$ cannot exist if $E(W) > 0$, since $P(V_N) \rightarrow 0$. Since $Z \leq X_1$, $\int_{Z \geq 0} Z\, dP$ exists; consequently $E(Z) = -\infty$.

REFERENCES

[1] J. L. DOOB, "The law of large numbers for continuous stochastic processes," *Duke Math. Jour.*, Vol. 6 (1940), pp. 290–306.
[2] A. KOLMOGOROFF, "Bemerkungen zu meiner Arbeit 'Uber die Summen Zuffalliger Grossen,'" *Math. Ann.*, Vol. 102 (1929–30), pp. 494–488.
[3] A. WALD, "Sequential tests of statistical hypotheses," *Annals of Math. Stat.* Vol., 16 (1945), pp. 117–186.

CORRECTION TO THE PAPER "ON A PROBLEM OF ESTIMATION OCCURING IN PUBLIC OPINION POLLS"

BY H. B. MANN
Ohio State University

In the paper "On a problem of estimation occurring in public opinion polls" (*Annals of Math. Stat.*, Vol. 16 (1945), pp. 85–90) the author made the assertion that, in the notation of the paper, $E[(\epsilon_i - r_i)^2]$ is always smaller than $E[(\epsilon_i - e_i)^2]$. This statement is incorrect and its supposed proof contains a numerical error in the fourth line from above on p. 90.

We have

$$E(r_i^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \int_{1/2}^{\infty} \int_{1/2}^{\infty} \frac{1}{2\pi\sigma_i^3} \exp\left[-\frac{1}{2\sigma_i^2} Q(x, y, p_i)\right] dx\, dy\, dp_i$$

$$= \frac{1}{2\pi} \frac{2}{\sqrt{3}} \int_{c/\sqrt{2}}^{\infty} \int_{c/\sqrt{2}}^{\infty} \exp\left[-\frac{1}{2}\frac{4}{3}(x^2 + y^2 - xy)\right] dx\, dy$$

$$c = \frac{\frac{1}{2} - \pi_i}{\sigma_i}.$$