A LOWER BOUND FOR THE VARIANCE OF SOME UNBIASED SEQUENTIAL ESTIMATES

By D. Blackwell and M. A. Girshick

Howard University and *Bureau of the Census*

Consider a sequence of independent chance variables $x_1, x_2, \cdots$ with identical distributions determined by an unknown parameter θ. We assume that $E\, x_i = \theta$ and that $W_k = x_1 + \cdots + x_k$ is a sufficient statistic for estimating θ from $x_1, \cdots, x_k$. A sequential sampling procedure is defined by a sequence of mutually exclusive events S_k such that S_k depends only on $x_1, \cdots, x_k$ and $\Sigma\, P(S_k) = 1$. Define $W = W_k$ and $n = k$ when S_k occurs. In a previous paper by one of the authors [1] it was shown that if $S_k = W_k\, C(S_1 + \cdots + S_{k-1})$, (where $C(A)$ denotes the event that A does not occur), the function $V(W, n) = E(x_1 \mid W, n)$ is an unbiased estimate of θ, and $\sigma^2(V) \leq \sigma^2(x_1)$. It is the purpose of this note to obtain a lower bound for $\sigma^2(V)$. Our result is:

Theorem I. $\sigma^2(V) \geq \dfrac{\sigma^2(x_1)}{E(n)}$.

We remark that the lower bound is actually attained in the classical case of samples of constant size N. For in this case, (see [1]), $V = E(x_1 \mid W_N) = W_N/N$. In fact we shall show that in a sense this is the only case in which the lower bound is attained.

The proof of Theorem I depends on certain properties of sums of independent chance variables. These, formulated more generally than is required for the proof of Theorem I, are given in

Theorem II. *Let $x_1, x_2, \cdots$ be independent chance variables with identical distributions, having mean θ and variance $\sigma^2(x_1)$. Let furthermore $\{S_k\}$ be any sequential test for which $E(n)$ is finite. Let $W = x_1 + \cdots + x_k$ when $n = k$. Then*

(a) $\sigma^2(W - \theta n) \leq \sigma^2(x_1)\, E(n)$.

(b) *If $\sigma^2(n)$ is finite, the equality sign holds in* (a).

(c) $E[x_1(W - \theta n)] = \sigma^2(x_1)$.

Proof of (a). Write $y_i = x_i - \theta$, and define $Y = y_1 + \cdots + y_k$ when $n = k$. By definition,

$$\sigma^2(W - \theta n) = \sum_{k=1}^{\infty} \int_{S_k} (y_1 + \cdots + y_k)^2\, dP. \tag{1}$$

To prove (a), we must verify that the series on the right of expression (1) converges and has sum $\leq \sigma^2(x_1)E(n)$. Now

$$\begin{aligned} &\sum_{k=1}^{N} \int_{S_k} (y_1 + \cdots + y_k)^2\, dP \\ &\leq \sum_{k=1}^{N-1} \int_{S_k} (y_1 + \cdots + y_k)^2\, dP + \int_{n \geq N} (y_1 + \cdots + y_N)^2\, dP \\ &= \sum_{k=1}^{N} \int_{n \geq k} y_k^2\, dP + 2 \sum_{k=2}^{N} \int_{n \geq k} y_k(y_1 + \cdots + y_{k-1})\, dP. \end{aligned} \tag{2}$$

Since the event $\{n \geq k\}$ is independent of y_k, each term in the second sum vanishes and the first sum becomes

$$\begin{aligned}\sum_{k=1}^{N} \int_{\{n \geq k\}} y_k^2 \, dP &= \sigma^2(x_1) \sum_{k=1}^{N} P\{n \geq k\} \\ &= \sigma^2(x_1)[P\{n = 1\} + 2P\{n = 2\} + \cdots NP\{n = N\} \\ &\qquad + NP\{n > N\}] \leq \sigma^2(x_1)E(n).\end{aligned} \tag{3}$$

This establishes Theorem II(a).

Proof of Theorem II(b). Write $z_i = | y_i |$ and let $Z = z_1 + \cdots + z_k$ when $n = k$. From (a) it follows that $\sigma^2[(Z - nE(z_i)]$ is finite. If in addition, $\sigma^2(n) < \infty$ then $E(Z^2) < \infty$. Thus the series

$$\sum_{k=1}^{\infty} \int_{S_k} (z_1 + \cdots + z_k)^2 \, dP = \sum_{1 \leq i,j \leq k < \infty} \int_{S_k} z_i z_j \, dP \tag{4}$$

converges, so that the series

$$\sum_{1 \leq i,j, \leq k < \infty} \int_{S_k} y_i y_j \, dP \tag{5}$$

converges absolutely. The terms of the latter series may be arranged to yield

$$\text{(A)}: \sum_{k=1}^{\infty} \int_{S_k} (y_1 + \cdots + y_k)^2 \, dP = \sigma^2 (W - \theta n)$$

or to yield

$$\text{B}: \sum_{k=1}^{\infty} \int_{\{n \geq k\}} y_k^2 \, dP + 2 \sum_{k=2}^{\infty} \int_{\{n \geq k\}} y_k(y_1 + \cdots + y_{k-1}) \, dP = \sigma^2(x_1)E(n).$$

This proves Theorem II(b).

Proof of Theorem II(c). It follows from Theorem II(a) that $Ex_1(W - \theta n)$ is finite. If we show that

(6) $E(W - \theta n \mid x_1) = x_1 - \theta$, i.e. $E(Y \mid y_1) = y_1$, it will follow [1] that

(7) $E[x_1(W - \theta n)] = E[x_1(x_1 - \theta)] = \sigma^2(x_1)$.

To verify (6), it is sufficient to show that if $f(x_1)$ is the characteristic function of an event depending only on x_1 (i.e. $f(x_1) = 1$ when the event occurs, $f(x_1) = 0$ otherwise)

$$E(fy_1) = E(fY). \tag{8}$$

Write $\phi_1 = 0, \phi_i = f \cdot (y_2 + \cdots + y_i), i \geq 2$.
Then it easily verified that

$$E(\phi_j \mid x_1, \cdots, x_i) = \phi_i \text{ for } j \geq i \tag{9}$$

$$E\phi_i \leq \sum_{k=1}^{i} | y_k | \tag{10}$$

$$E(\phi_i) = 0. \tag{11}$$

Hence it follows [2] that $E\phi = 0$ where $\phi = \phi_i$ when $n = i$. In our case $\phi = fY - fy_1$, and $E\phi = 0$ yields (6). This completes the proof of Theorem II.

PROOF OF THEOREM I. In [1] it is proved that $E(x_1(W - \theta n)) = E[V(W - \theta n)]$. Hence employing Theorem II we get

$$\sigma^2(x_1) = E[V(W - \theta n)] = \sigma(V)\sigma(W - \theta n)\rho \tag{12}$$

where ρ, $(0 \leq \rho \leq 1)$, is the coefficient of correlation between V and $W - \theta n$. Substituting for $\sigma(W - \theta n)$ we get

$$\begin{aligned} \sigma^2(x_1) &\leq \sigma(V)\sigma(x_1)\sqrt{E(n)}\,\rho \\ &\leq \sigma(V)\sigma(x_1)\sqrt{E(n)}. \end{aligned} \tag{13}$$

Solving for $\sigma(V)$ we finally obtain

$$\sigma^2(V) \geq \frac{\sigma^2(x_1)}{E(n)} \tag{14}$$

which proves Theorem I.[1]

If $\sigma^2(n)$ is finite, the equality sign in (14) will hold if and only if $\rho = 1$. We shall now prove the following.

THEOREM III. *Let N be the minimum value of n for which $P(n = N) \neq 0$. Then, a necessary and sufficient condition that $\rho = 1$ is that $P(n = N) = 1$.*

PROOF. The sufficiency of this condition follows from the fact that if $P(n = N) = 1$, $V = W/N$. To prove the necessity of this condition, we observe that if $\rho = 1$, V is a linear function of $W - n\theta$. That is,

$$V = \alpha(W - n\theta) + \beta. \tag{15}$$

Now, since $EV = \theta$ and $E(W - n\theta) = 0$, it follows that $\beta = \theta$. Also, since by hypothesis $\sigma^2(V) = \sigma^2(x_1)/E(n)$ and $\sigma^2(W - n\theta) = \sigma^2(x_1)E(n)$, it follows that $\alpha = 1/\mathrm{E}(n)$. Hence the estimate V is given by

$$V = \frac{W - n\theta}{E(n)} + \theta. \tag{16}$$

---

[1] Under certain regularity conditions Cramér has obtained the inequality

$$\sigma^2(x) \geq 1/E\left(\frac{\partial \log f}{\partial \theta}\right)^2$$

where $f = f(x, \theta)$ is the density function of x ([3], p. 475). Thus with the same regularity conditions, our inequality yields

$$\sigma^2(V) \geq 1/E(n)E\left(\frac{\partial \log f}{\partial \theta}\right)^2,$$

which is a special case of the results presented by J. Wolfowitz in this issue of the *Annals*.

Let N be defined as above. We note that $N < \infty$ since by hypothesis $E(n) < \infty$. Let V_N be the estimate of θ when the sequential test terminates with $n = N$. Then $V_N = W/N$. Substituting this value in (16) we get

$$\frac{W}{N} - \theta = \frac{N}{E(n)}\left[\frac{W}{N} - \theta\right]. \tag{17}$$

We exclude the trivial case where $W \equiv N\theta$. Then (16) yields $E(n) = N$. That is $P(n = N) = 1$. This proves the theorem.

We remark that N may be a function of θ but for a fixed θ, $n = N$ is fixed when $\rho = 1$.

REFERENCES

[1] D. Blackwell, "Conditional expectation and unbiased sequential estimation." Submitted to *Annals of Math. Stat.*

[2] D. Blackwell and M. A. Girshick, "On sums of sequences of independent chance vectors, with applications to the random walk in k dimensions," *Annals of Math. Stat.*, Vol. 17 (1946).

[3] Harald Cramér, *Mathematical Methods of Statistics*, Princeton Univ. Press, 1946.

AN EXTENSION TO TWO POPULATIONS OF AN ANALOGUE OF STUDENT'S t-TEST USING THE SAMPLE RANGE

By John E. Walsh

Princeton University

1. Summary. The modified t-test considered by Daly[1] (see [1]) is used to develop one-sided significance tests to decide whether the mean of a new normal population exceeds the mean of an old normal population having the same variance. Significance tests are also developed to decide whether the mean of the new population is less than the mean of the old population. These tests require very little computation for their application and are approximately as powerful as the most powerful tests of these hypotheses.

2. Introduction. Let $r_1, \cdots, r_n$, $(n \leq 10)$, be independently distributed according to a normal distribution with zero mean and unit variance. Let $r_{(u)}$ denote the uth largest of the r's. Then Daly has shown how to determine numbers g_α such that

$$\begin{aligned} &Pr[\bar{r}/(r_{(n)} - r_{(1)}) > g_\alpha] = \alpha \\ &Pr[\bar{r}/(r_{(n)} - r_{(1)}) < -g_\alpha] = \alpha. \end{aligned} \tag{1}$$

This note will use these relations to develop easily applied significance tests to decide whether the mean ν of a new normal population exceeds the mean μ of

[1] This problem is also considered by Lord in [2]. This note was in proof when [2] appeared.