The Annals of Statistics
1980, Vol. 8, No. 1, 221–225

SEQUENTIAL POINT ESTIMATION OF THE DIFFERENCE OF TWO NORMAL MEANS[1]

By Malay Ghosh and Nitis Mukhopadhyay

Iowa State University and University of Minnesota and University of Missouri, Columbia

A sequential procedure for estimating the difference of two normal means when the variances are unknown and not necessarily equal is proposed, and an asymptotic expression for "regret" is given. This generalizes the corresponding one sample result of Woodroofe.

1. Introduction. Consider two mutually independent sequences of random variables $X_1, X_2, \cdots$ and $Y_1, Y_2 \cdots$ where the X_i's are i.i.d. $N(\mu_1, \sigma_1^2)$ and the Y_j's are i.i.d. $N(\mu_2, \sigma_2^2)$; $-\infty < \mu_1, \mu_2 < \infty$, $0 < \sigma_1, \sigma_2 < \infty$, μ_1 and μ_2 both unknown. The problem is to find a point estimator of $\mu_1 - \mu_2$. Taking samples of sizes r and s from the X's and the Y's respectively, suppose the loss incurred in estimating $\mu = \mu_1 - \mu_2$ by $W = \bar{X}_r - \bar{Y}_s$ ($\bar{X}_r = r^{-1}\Sigma_1^r X_i$, $\bar{Y}_s = s^{-1}\Sigma_{j=1}^s Y_j$) is

$$L_{r,s} = A(W - \mu)^2 + c(r + s), \tag{1.1}$$

where $A(> 0)$ is the known weight and $c(> 0)$ is the known cost per unit observation. Then the risk is

$$v_{r,s}(c) = A\left(r^{-1}\sigma_1^2 + s^{-1}\sigma_2^2\right) + c(r + s). \tag{1.2}$$

For known σ_1 and σ_2, the pair (r^*, s^*) for which (1.2) is a minimum, is given by

$$r^* = b\sigma_1, s^* = b\sigma_2, \tag{1.3}$$

where $b = (A/c)^{\frac{1}{2}}$. For this pair

$$r^*/s^* = \sigma_1/\sigma_2, n^* = r^* + s^* = b(\sigma_1 + \sigma_2), \tag{1.4}$$

and the corresponding minimum risk is

$$v(c) = v_{r^*, s^*}(c) = 2cn^*. \tag{1.5}$$

When σ_1 and σ_2 are unknown, no fixed sample size minimizes (1.2) simultaneously for all $0 < \sigma_1, \sigma_2 < \infty$. Sequential procedures determining r and s as random variables were proposed by Mukhopadhyay (1975, 1977) as follows:

Define for $i \geqslant 2, j \geqslant 2$,

$$u_i^2 = (i - 1)^{-1}\Sigma_{k=1}^i\left(X_k - \bar{X}_i\right)^2, v_j^2 = (j - 1)^{-1}\Sigma_{k=1}^j\left(Y_k - \bar{Y}_j\right)^2. \tag{1.6}$$

Received June 1978; revised November 1978.

[1]Research supported by the Army Research Office Durham Grant Numbers DAAG29-76-G-0057 and DAAG29-76-G-0038.

AMS 1970 *subject classifications.* 62L12.

Key words and phrases. Difference of two normal means, sequential estimation, risk efficiency, regret.

Take $m(\geqslant 3)$ observations on X and Y to start with. Then, if at any stage i observations are taken on X and j observations are taken on Y and the process does not stop, the next observation is taken on X or Y according as

(1.7) $$i/j \leqslant u_i/v_j \quad \text{or} \quad i/j > u_i/v_j.$$

The stopping time $N \equiv N_c$ is defined by $N =$ first integer $n(\geqslant 2m)$ such that if $R = r$ observations on X and $S = s$ observations on Y are taken, with $r + s = n$,

(1.8) $$r \geqslant bu_r, \qquad s \geqslant bv_s.$$

The risk involved in this sequential procedure is

(1.9)

$$R(c) = AE\left(\bar{X}_R - \bar{Y}_S - \mu\right)^2 + cE(R + S) = AE\left(\sigma_1^2 R^{-1} + \sigma_2^2 S^{-1}\right) + cE(R + S).$$

Following Starr (1966) and Starr and Woodroofe (1969), the "risk efficiency" and the "regret" are defined respectively by

(1.10) $$R_e(c) = R(c)/\nu(c) = (\sigma_1 + \sigma_2)^{-1}\left[\sigma_1\{r^* E(R^{-1}) + E(R/r^*)\} + \sigma_2\{s^* E(S^{-1}) + E(S/s^*)\}\right]$$

and

(1.11) $$R_g(c) = R(c) - \nu(c) = c\left\{E(R - r^*)^2/R + E(S - s^*)^2/S\right\}.$$

Mukhopadhyay (1975) showed that $R_g(c) = 0(c)$ as $c \to 0$. Since $\nu(c) = 0_e(c^{1/2})$, where 0_e denotes the exact order, this implies that $R_e(c) \to 1$ as $c \to 0$.

In this note we prove the stronger result

(1.12) $$R_g(c) = c + o(c) \text{ as } c \to 0, \quad \text{when } m \geqslant 3.$$

To prove (1.12), we proceed as follows. Mukhopadhyay (1975, 1977) showed that $(R - r^*)/(r^*)^{1/2} \to_L N(0, \frac{1}{2})$, $(S - s^*)/(s^*)^{1/2} \to_L N(0, \frac{1}{2})$ as $c \to 0$. Also, it was shown there that $R/r^* \to 1$ a.s., $S/s^* \to 1$ a.s. as $c \to 0$. Thus, $(R - r^*)^2/R \to_L \frac{1}{2}\chi_1^2$, $(S - s^*)^2/S \to_L \frac{1}{2}\chi_1^2$ as $c \to 0$. Hence, for proving (1.12) it suffices to prove the following result.

THEOREM. *If* $m \geqslant 3$, $(R - r^*)^2/R$ *and* $(S - s^*)^2/S$ *are uniformly integrable in* $c \leqslant c_0$ *for some* $c_0 > 0$.

We shall only outline the proof of this theorem in the next section omitting most of the details. Note that in the one sample normal case Woodroofe (1977) proved a similar uniform integrability result by appealing to a more general theorem. Woodroofe's method can be used to cover the present situation as well. However, although on similar lines, our method of proof is not quite the same as Woodroofe's. We can especially avoid the complicacies involved in his estimation of entities of the type $P(|R - r^*| > x(r^*)^{1/2}, (1 - \varepsilon)r^* \leqslant R \leqslant 2r^*)$ etc. where x is sufficiently large, and $0 < \varepsilon < 1$. It should be emphasized that the simplifications in our proof are not merely the results of normality assumptions, because, although, stated in terms of chi-squared random variables, our method of proof uses only the moment bounds for the tail probabilities of centered means of i.i.d. rv's having finite moments of order $2 + \delta(\delta > 0)$, an assumption needed as well by Woodroofe (1977). Also, unlike Woodroofe (1977), we need the starting sample size $m \geqslant 3$ rather than $m \geqslant 4$.

The motivation behind the use of the sampling scheme (1.7) or the stopping rule (1.8) originates in the work of Robbins, Simons and Starr (1967) who considered the fixed length interval estimation of the difference of two normal means in the presence of unknown and possibly unequal variances.

2. Proof of the Theorem. We prove only that $(R - r^*)^2/R$ is uniformly integrable in $c \leqslant c_0$ when $m \geqslant 3$. A similar proof works for $(S - s^*)^2/S$. First show that for $m \geqslant 3$, $(R - r^*)^2/r^*$ is uniformly integrable in $c \leqslant c_0$. In what follows K is a generic constant, positive but not depending on c, and I is the usual indicator function. Write for any $a > 0$,

$$(2.1)\quad E\left[\{(R - r^*)^2/r^*\}I_{[(R-r^*)^2 > a^2 r^*]}\right] = a^2 P\left(|R - r^*| > a(r^*)^{\frac{1}{2}}\right) + 2\int_a^\infty x P\left(|R - r^*| > x(r^*)^{\frac{1}{2}}\right)dx.$$

Since, $(r^*)^{-\frac{1}{2}} = (A/c)^{-\frac{1}{4}}\sigma_1^{-\frac{1}{2}} = Kc^{\frac{1}{4}}$, choose c_1 such that $a > 2(r^*)^{-\frac{1}{2}}$ for $c \leqslant c_1$. Write $k = [r^* + x(r^*)^{\frac{1}{2}}]$, where $[y]$ denotes the integer part of y. Then for $x \geqslant a$ and $c \leqslant c_1$, one has the inequalities (i) $k - 1 \geqslant r^* + x(r^*)^{\frac{1}{2}} - 2 > r^*$ and (ii) $k \geqslant r^* + x(r^*)^{\frac{1}{2}} - 1 \geqslant r^* + \frac{1}{2}x(r^*)^{\frac{1}{2}}$. Using (i) and (ii) and Markov's inequality, for $x \geqslant a$,

$$(2.2)\quad \begin{aligned} P\left(R > r^* + x(r^*)^{\frac{1}{2}}\right) &= P(R > k) \leqslant P(k^2 < b^2 u_k^2) = P\left(\chi_{k-1}^2 > (k-1)(k/r^*)^2\right) \\ &\leqslant P\left(\chi_{k-1}^2 - (k-1) > x(r^*)^{\frac{1}{2}}\right) \leqslant K(k-1)^p (r^*)^{-p} x^{-2p} \\ &\leqslant Kx^{-2p}. \end{aligned}$$

Again, for any $a > 0$, choose c_2 such that $(r^*)^{\frac{1}{2}} > 2a$ for all $c \leqslant c_2$. Then, for $x \geqslant a$,

$$(2.3)\quad \begin{aligned} \int_a^\infty x P\left(R < r^* - x(r^*)^{\frac{1}{2}}\right)dx &= \int_a^{(r^*)^{\frac{1}{2}}} x P\left(R < r^* - x(r^*)^{\frac{1}{2}}\right)dx \\ &\leqslant \int_a^{\frac{1}{2}(r^*)} x\left[P\left(R \leqslant \tfrac{1}{2}r^*\right) + P\left(\tfrac{1}{2}r^* < R < r^* - x(r^*)^{\frac{1}{2}}\right)\right]dx \\ &\quad + \int_{\frac{1}{2}(r^*)^{\frac{1}{2}}}^{(r^*)^{\frac{1}{2}}} x P\left(R \leqslant \tfrac{1}{2}r^*\right)dx \\ &= \int_a^{(r^*)^{\frac{1}{2}}} x P\left(R \leqslant \tfrac{1}{2}r^*\right)dx + \int_a^{\frac{1}{2}(r^*)^{\frac{1}{2}}} x P\left(\tfrac{1}{2}r^* < R < r^* - x(r^*)^{\frac{1}{2}}\right)dx. \end{aligned}$$

Now, following the lines of proof of lemma 5 of Ghosh et al. (1976), one gets for $c \leqslant c_3$,

$$(2.4)\quad P\left(R \leqslant \tfrac{1}{2}r^*\right) \leqslant Kc^{\frac{1}{2}(m-1)}.$$

Hence, for $c \leqslant c_4 = \min(c_2, c_3)$ and $m \geqslant 3$,

$$\int_{\frac{1}{2}(r^*)^{\frac{1}{2}}}^{(r^*)^{\frac{1}{2}}} xP\left(R \leqslant \tfrac{1}{2}r^*\right)dx \leqslant K(r^*)P\left(R \leqslant \tfrac{1}{2}r^*\right) \leqslant Kc^{\frac{1}{2}(m-2)} \leqslant Kc_4^{\frac{1}{2}(m-2)} \leqslant K(a^4)^{-\frac{1}{2}(m-2)}. \tag{2.5}$$

Also, for $a \leqslant x \leqslant \frac{1}{2}(r^*)^{\frac{1}{2}}$, $c \leqslant c_4$, writing $k_1 = [\frac{1}{2}r^*]$, $k_2 = [r^* - x(r^*)^{\frac{1}{2}}]$, one gets by using the Kolmogorov inequality for sum of i.i.d. rv's,

$$\begin{aligned} P\left(\tfrac{1}{2}r^* < R < r^* - x(r^*)^{\frac{1}{2}}\right) &= P\left(\cup_{r=k_1+1}^{k_2}(r \geqslant bu_r)\right) \\ &\leqslant P\left(\cup_{r=k_1+1}^{k_2}\left\{\chi_{r-1}^2 - (r-1) \leqslant -2xk_1(r^*)^{-\frac{1}{2}}\right\}\right) \\ &\leqslant E\left(\chi_{k_2-1}^2 - (k_2-1)\right)^{2p} / \left(Kx(r^*)^{\frac{1}{2}}\right)^{2p} \leqslant K(k_2-1)^{2p}x^{-2p}(r^*)^{-p} \leqslant Kx^{-2p}. \end{aligned} \tag{2.6}$$

Combining (2.2) – (2.6) it follows from (2.1) that for $c \leqslant c_0 = \min(c_1, c_4)$, $p > 1$ and $m \geqslant 3$,

$$\begin{aligned} &E\left[\left\{(R - r^*)^2/r^*\right\} I_{[(R-r^*)^2 > a^2r^*]}\right] \\ &\quad \leqslant K\left[a^{2-2p} + \int_a^\infty x^{1-2p}dx + (a^4)^{-\frac{1}{2}(m-2)} + \int_a^\infty x^{1-2p}dx\right] \\ &\quad \to 0 \text{ as } a \to \infty. \end{aligned} \tag{2.7}$$

This shows that $(R - r^*)^2/r^*$ is uniformly integrable in $c \leqslant c_0$. Next observe that

$$E\left[\left\{(R - r^*)^2/R\right\} I_{\left[R > \frac{1}{2}r^*\right]} I_{[(R-r^*)^2 > a^2R]}\right] \leqslant 2E\left\{(R - r^*)^2/r^*\right\} I_{\left[(R-r^*)^2 > \frac{1}{2}a^2r^*\right]}. \tag{2.8}$$

Also, choosing $a > c_0^{-1}$ for $c \leqslant c_0$,

$$\begin{aligned} &E\left[\left\{(R - r^*)^2/R\right\} I_{\left[R < \frac{1}{2}r^*\right]} I_{[(R-r^*)^2 > a^2R]}\right] \\ &\quad \leqslant E\left[\left\{(R - r^*)^2/R\right\} I_{\left[R < \frac{1}{2}r^*\right]}\right] \\ &\quad \leqslant Kr^*P\left(R \leqslant \tfrac{1}{2}r^*\right) \leqslant Kc_0^{(m-2)/2} \leqslant Ka^{-(m-2)/2}. \end{aligned} \tag{2.9}$$

The uniform integrability of $(R - r^*)^2/R$ in $c \leqslant c_0$ now follows from (2.8) and (2.9).

REFERENCES

[1] Ghosh, M., Sinha, B. K., and Mukhopadhyay, N. (1976). Multivariate sequential point estimation. *J. Multivariate Anal.* **6** 281–294.

[2] Mukhopadhyay, N. (1975). Sequential methods in estimation and prediction. Ph.D. dissertation, Indian Statistical Institute, Calcutta.

[3] Mukhopadhyay, N. (1977). Remarks on sequential estimation of a linear function of two means: the normal case. *Metrika* **24** 197–201.

[4] ROBBINS, H., SIMONS, G., AND STARR, N. (1967). A sequential analogue of the Behrens-Fisher problem. *Ann Math. Statist.* **38** 1384–1391.

[5] STARR, N. (1966). On the asymptotic efficiency of a sequential procedure for estimating the mean. *Ann. Math. Statist.* **37** 1173–1185.

[6] STARR, N., AND WOODROOFE, M. (1969). Remarks on sequential point estimation. *Proc. Nat. Acad. Sci. USA* **63** 285–288.

[7] STARR, N., AND WOODROOFE, M. (1972). Further remarks on sequential point estimation: the exponential case. *Ann. Math. Statist.* **43** 1147–1154.

[8] WOODROOFE, M. (1977). Second order approximations for sequential point and interval estimation. *Ann. Statist.* **5** 984–995.

DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
SNEDECOR HALL
AMES, IOWA 50011

DEPARTMENT OF STATISTICS
OKLAHOMA STATE UNIVERSITY
STILLWATER, OKLAHOMA 74074