TWO-STAGE PROCEDURES FOR ESTIMATING THE DIFFERENCE BETWEEN MEANS*

By

S. G. Ghurye and Herbert Robbins

Institute of Statistics

University of North Carolina

Institute of Statistics
Mimeograph Series No. 80

August, 1953

* Work supported by the U. S. Air Force under Contract AF 18(600)-83.

TWO-STAGE PROCEDURES FOR ESTIMATING THE DIFFERENCE BETWEEN MEANS*

By S. G. Ghurye and Herbert Robbins

Institute of Statistics
University of North Carolina

Introduction and summary. Given two populations P_i $(i = 1, 2)$ with unknown means θ_i and variances σ_i^2, we wish to estimate the difference $\theta_1 - \theta_2$. Let $t_i(n)$ be the mean $(X_{i,1} + X_{i,2} + \ldots + X_{i,n})/n$ of a sample from P_i. Then $t_1(n_1) - t_2(n_2)$ is an unbiased estimate of $\theta_1 - \theta_2$, with variance $\sum_i (\sigma_i^2/n_i)$. Assuming the cost of sampling to be a known linear function of the number of observations, the cost of taking n_1 observations from P_1 and n_2 from P_2 is $a_1n_1 + a_2n_2 + a_3$. If there is a prescribed upper bound A_o to the cost of sampling, n_1 and n_2 are subject to the restriction

$$(0.1) \qquad a_1n_1 + a_2n_2 \leq A = A_o - a_3.$$

The quantity $\sum_i (\sigma_i^2/n_i)$, equal to the variance of $t_1(n_1) - t_2(n_2)$ for integer values of the n_i, is minimized for continuous $n_i > 0$ subject to (0.1) by taking $n_i = n_i^o$, where

$$(0.2) \qquad n_i^o = (A/a_i)\, a_i^{1/2}\, \sigma_i / \sum_j a_j^{1/2}\, \sigma_j \; ;$$

the minimum value being equal to

$$(0.3) \qquad V^o(A) = \sum_i (\sigma_i^2/n_i^o) = A^{-1}(\sum_i a_i^{1/2}\, \sigma_i)^2.$$

* Work supported by the U. S. Air Force under Contract AF 18(600)-83.

When the ratio σ_2/σ_1 on which the optimum values (0.2) depend is not known, we can use a two-stage procedure for estimating $\Theta_1 - \Theta_2$, first taking a sample of $m_1 + m_2$ observations, m_i from P_i, and then using estimates of σ_i obtained from this preliminary sample to distribute the remaining observations between the P_i. We shall investigate the performance of this estimation procedure.

For previous work done on problems of this kind, reference may be made to Putter [1] and the literature cited in that paper. Putter considers the problem of estimating the mean of a population composed of a known number of normally distributed strata whose relative proportions are known. See also Robbins [2; p. 528].

In section 1, we assume the P_i to be normal and evaluate the variance of the two-stage estimate. In section 2, we show that as m_1, m_2 and $A \longrightarrow \infty$ in a certain way, the ratio of this variance to the minimum variance $V^o(A)$ tends to unity, and we also prove the asymptotic result for more general populations.

1. Normal populations. When the P_i are known to be normal, we choose positive integers m_i such that $a_1 m_1 + a_2 m_2 < A$ and take m_i observations from P_i. Let

$$(1.1) \quad s_i^2(m_i) = \left\{ \sum_{j=1}^{m_i} X_{i,j}^2 - m_i t^2(m_i) \right\} / (m_i - 1) = \text{estimated variance of } P_i \text{ *},$$

---

* Putter [1] uses $s_i^2(m_i)(m_i - 1)/(m_i - 2)$ instead of $s_i^2(m_i)$ because it minimizes an expression σ_D^2 which is the variance of the estimate obtained by ignoring the fact that the sample-sizes prescribed by the two-stage procedure are truncated short of the extreme limits possible under (0.1).

$$(1.2)\qquad u(m_1, m_2) = a_1^{1/2} s_1(m_1)/\sum_i a_i^{1/2} s_i(m_1);$$

$$(1.3)\qquad n_1^* = \begin{cases} m_1 & \text{if } Au(m_1, m_2)/a_1 < m_1, \\ (A - a_2 m_2)/a_1 & \text{if } (A - a_2 m_2)/a_1 < Au(m_1, m_2)/a_1, \\ Au(m_1, m_2)/a_1 & \text{if } m_1 \leq Au(m_1, m_2)/a_1 \leq (A - a_2 m_2)/a_1; \end{cases}$$

$$(1.4)\qquad n_2^* = (A - a_1 n_1^*)/a_2;$$

and

$$(1.5)\qquad \tilde{n}_i = [n_i^*],$$

where $[x]$ is the largest integer in x. Having computed $\tilde{n}_i$, we take $(\tilde{n}_i - m_i)$ more observations ($X_{i,j}$, $j = m_i + 1, \ldots, \tilde{n}_i$) from P_i, and estimate $\theta_1 - \theta_2$ by

$$(1.6)\qquad t_1(\tilde{n}_1) - t_2(\tilde{n}_2) = \sum_{j=1}^{\tilde{n}_1} X_{1,j}/\tilde{n}_1 - \sum_{j=1}^{\tilde{n}_2} X_{2,j}/\tilde{n}_2.$$

Let

$$(1.7)\qquad \tilde{V}(A) = \text{Var}\left\{ t_1(\tilde{n}_1) - t_2(\tilde{n}_2) \right\}.$$

Now,

$$(1.8)\qquad t_i(\tilde{n}_i) = m_i t_i(m_i)/\tilde{n}_i + (\tilde{n}_i - m_i) t_i(\tilde{n}_i - m_i)/\tilde{n}_i,$$

where $t_i(\tilde{n}_i - m_i) = \sum_{j=m_i+1}^{\tilde{n}_i} X_{i,j}/(\tilde{n}_i - m_i)$.

Since the $\tilde{n}_i$ depend only on the $s_i(m_i)$, for fixed s_i the random variables $t_1(m_1)$, $t_2(m_2)$, $t_1(\tilde{n}_1 - m_1)$, $t_2(\tilde{n}_2 - m_2)$ are mutually independent, the conditional distributions of $t_i(m_i)$ and $t_i(\tilde{n}_i - m_i)$ being respectively $\mathcal{N}\{\theta_i, \sigma_i^2/m_i\}$ and $\mathcal{N}\{\theta_i, \sigma_i^2/(\tilde{n}_i - m_i)\}$. Hence for fixed s_i the conditional distribution of $t_1(\tilde{n}_1) - t_2(\tilde{n}_2)$ is $\mathcal{N}\{\theta_1 - \theta_2, \sum_i (\sigma_i^2/\tilde{n}_i)\}$. Consequently,

$$(1.9) \qquad E\left\{ t_1(\tilde{n}_1) - t_2(\tilde{n}_2) \right\} = \theta_1 - \theta_2, \text{ and } \tilde{V}(A) = E\left\{ \sum_i (\sigma_i^2/\tilde{n}_i) \right\}.$$

Let

$$(1.10) \qquad F(u) = \Pr\left\{ u(m_1, m_2) \leq u \right\}.$$

Then from (1.5), (1.9) and (1.10), we have

$$(1.11) \qquad \begin{aligned} \tilde{V}(A) = {} & \left\{ \sigma_1^2 m_1^{-1} + \sigma_2^2 [(A - a_1 m_1)/a_2]^{-1} \right\} F(a_1 m_1/A) \\ & + \left\{ \sigma_1^2 [(A - a_2 m_2)/a_1]^{-1} + \sigma_2^2 m_2^{-1} \right\} \left\{ 1 - F(1 - a_2 m_2/A) \right\} \\ & + \int_{a_1 m_1/A}^{1 - a_2 m_2/A} \left\{ \sigma_1^2 [Au/a_1]^{-1} + \sigma_2^2 [A(1-u)/a_2]^{-1} \right\} dF(u). \end{aligned}$$

In what follows we shall denote by $V^*(A)$ the expression obtained by dropping the square brackets in (1.11).

Let

$$(1.12) \qquad \left\{ \begin{aligned} & \rho = \sigma_2/\sigma_1, \quad c = (a_2/a_1)^{1/2}, \; b_1 = (A - a_1 m_1)/(a_2 m_1), \\ & b_2 = a_1 m_2/(A - a_2 m_2), \quad r_i = (m_i - 1)/2, \text{ and } q = r_2/r_1. \end{aligned} \right.$$

Then from (0.3) we have

$$V^o(A) = A^{-1} a_1 \sigma_1^2 (1 + \rho c)^2,$$

and we can reduce the expression for $V^*(A)$ to

$$(1.13)\qquad A\left\{V^*(A) - V^o(A)\right\} = a_1\sigma_1^2(\rho - cb_1)^2 b_1^{-1}F(a_1m_1/A) + a_1\sigma_1^2(\rho - cb_2)^2 b_2^{-1}\left\{1 - F(1 - a_2m_2/A)\right\} + \int_{a_1m_1/A}^{1-a_2m_2/A} a_1\sigma_1^2\left\{(1-u)u^{-1} + \rho^2c^2(1-u)^{-1}u - 2\rho c\right\} dF(u).$$

Finally, making the substitution

$$(1.14)\quad w = \rho^2c^2u^2/\left\{\rho^2c^2u^2 + q(1-u)^2\right\} = (m_1 - 1)(s_1^2/\sigma_1^2)/\sum_i (m_i - 1)(s_i^2/\sigma_i^2) = \chi^2_{m_1-1}/(\chi^2_{m_1-1} + \chi^2_{m_2-1}),$$

so that w has the density function

$$(1.15)\qquad f(w) = \left\{1/B(r_1,r_2)\right\} w^{r_1-1}(1-w)^{r_2-1}, \quad 0 \le w \le 1$$

we reduce (1.13) to

$$(1.16)\qquad V^*(A)/V^o(A) = 1 + (1+\rho c)^{-2}\left\{\rho c I_1 + (\rho - cb_1)^2\, b_1^{-1}I_2 + (\rho - cb_2)^2\, b_2^{-1}I_3\right\},$$

where

$$(1.17)\qquad \begin{cases} I_1 = \int_{\beta_1}^{\beta_2} \left\{(qw)^{1/2}(1-w)^{-1/2} + (qw)^{-1/2}(1-w)^{1/2} - 2\right\} f(w)\,dw, \\ I_2 = \int_0^{\beta_1} f(w)\,dw, \quad I_3 = \int_{\beta_2}^{1} f(w)\,dw, \quad \beta_i = \rho^2/\left\{\rho^2 + c^2qb_i^2\right\}. \end{cases}$$

Hence, V^*/V° can be computed by means of tables of the incomplete β-function. We have done this for $a_1 = a_2 = a$, $N = A/a = 30, 50$, $m_1 = m_2 = m = (0.2)N$, $(0.3)N$, $(0.4)N$ and various values of ρ. The results are given in the table.

For $a_1 = a_2$, the usual procedure for estimating $\theta_1 - \theta_2$ consists in taking $n_1 = n_2 = N/2$ and using the estimate $t_1(N/2) - t_2(N/2)$. The variance is

$$V' = 2(\sigma_1^2 + \sigma_2^2)/N.$$

For comparing V^* with V', the values of V'/V° are given in the last row of the table.

TABLE

Comparison of V^* with V° and V' for normal populations.

	m/N \ ρ	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
V^*/V°	0.2	1.064	1.062	1.058	1.054	1.049	1.044	1.039	1.034	1.030
for	0.3	1.034	1.032	1.028	1.023	1.018	1.016	1.016	1.018	1.022
N = 30	0.4	1.017	1.014	1.012	1.014	1.025	1.039	1.056	1.075	1.094
V^*/V°	0.2	1.032	1.031	1.031	1.029	1.027	1.025	1.022	1.019	1.017
for	0.3	1.021	1.019	1.017	1.014	1.011	1.010	1.008	1.011	1.016
N = 50	0.4	1.013	1.009	1.007	1.010	1.021	1.036	1.055	1.074	1.094
V'/V°		1.000	1.012	1.040	1.074	1.111	1.148	1.184	1.218	1.250

The two-stage procedure seems to effect considerable improvement over the usual one-stage procedure for values of ρ away from 1; and the performance seems to be best for m/N in the neighbourhood of $\sigma_1/(\sigma_1 + \sigma_2)$ if $\sigma_1 < \sigma_2$.

<u>2. Asymptotic efficiency</u>. The idea of substituting $s_i(m_i)$ for σ_i in (0.2) is based on the belief that as $m_i \longrightarrow \infty$, the ratio $\tilde{n}_1/\tilde{n}_2$ will approach the optimum value n_1^o/n_2^o and $\tilde{V}/V^o$ will approach unity. We prove that this is the case when the populations are normal, and then we shall prove a similar result which is true also for other populations.

THEOREM 1. Let P_i be normal, and consider $V^*(A)/V^o(A)$ as given by (1.16). Let a_1, a_2, and ρ remain fixed while m_1, m_2 and A become infinite in such a way that

$$(2.1) \qquad \begin{array}{l} 0 < h \leq m_1/m_2 \leq h' < \infty \text{, where } h, h' \text{ are fixed,} \\ \text{and } m_i/A \longrightarrow 0, \qquad i = 1, 2. \end{array}$$

Then

$$(2.2) \qquad V^*(A)/V^o(A) \longrightarrow 1.$$

PROOF. In (1.16),

$$I_1 \leq q^{1/2}\beta(r_1+1/2, r_2-1/2)/\beta(r_1,r_2) + q^{-1/2}\beta(r_1-1/2,r_2+1/2)/\beta(r_1,r_2) - 2$$

$$= \left\{q^{1/2}\Gamma(r_1+1/2)\Gamma(r_2-1/2) + q^{-1/2}\Gamma(r_1-1/2)\Gamma(r_2+1/2)\right\} / \left\{\Gamma(r_1)\Gamma(r_2)\right\}$$

$$-2$$

which converges to zero, since

$$(2.3) \qquad x^h\Gamma(x - h)/\Gamma(x) \longrightarrow 1 \text{ as } x \longrightarrow \infty .$$

$$I_2 \leq \Pr\left\{ w < \rho^2/(\rho^2 + c^2qb_1^2)\right\} \leq \Pr\left\{ w^{-1} > c^2qb_1^2/\rho^2 \right\}$$

$$\leq E\left\{ w^{-1} \right\} \rho^2/(c^2qb_1^2),$$

and

$$I_3 \leq Pr\left\{ w > \rho^2/(\rho^2 + c^2qb_2^2)\right\} \leq Pr\left\{ (1 - w)^{-1} > \rho^2/(c^2qb_2^2)\right\}$$

$$\leq E\left\{(1 - w)^{-1}\right\} c^2qb_2^2/\rho^2.$$

Now,

$$E\left\{ w^{-1} \right\} = \Gamma(r_1 - 1)\Gamma(r_1 + r_2)/\left\{\Gamma(r_1)\Gamma(r_1 + r_2 - 1)\right\},$$

and

$$E\left\{ (1 - w)^{-1} \right\} = \Gamma(r_2 - 1)\Gamma(r_1 + r_2)/\left\{\Gamma(r_2)\Gamma(r_1 + r_2 - 1)\right\},$$

both of which remain bounded on account of (2.1) and (2.3). Since $b_1 \longrightarrow \infty$ and $b_2 \longrightarrow 0$, we have

$$(\rho - cb_1)^2 b_1^{-1} I_2 \longrightarrow 0 \quad \text{and} \quad (\rho - cb_2)^2 b_2^{-1} I_3 \longrightarrow 0.$$

Hence (2.2) is proved.

From the expression (1.11) for $\tilde{V}(A)$, it follows that

$$\tilde{V}(A)/V^o(A) \longrightarrow 1.$$

Next, we remove the restriction that the P_i be normal, assuming only that they have finite variances σ_i^2 and that we know functions f_i such that the statistics

(2.4) $\quad s_i(n) = f_i(X_{i,1}, \ldots, X_{i,n}; n), \qquad i = 1, 2,$

satisfy conditions (I)-(III) below.

Let $F_i(s; n) = Pr\left\{ s_i(n) \leq s \right\}$, and $C_i(\epsilon) = [\sigma_i - \epsilon, \sigma_i + \epsilon]$.

We assume:

(I) There exists an $\alpha > 1$ such that for every fixed $\epsilon > 0$,

$$n^{\alpha}\Pr\left\{ s_i(n)/C_i(\epsilon)\right\} \quad \text{is bounded for all } n > 0,\ i = 1, 2;$$

(II) There exists an $\epsilon > 0$ and $< \min(\sigma_1, \sigma_2)$ such that $n\left\{ \int_{C_i(\epsilon)} s^k dF_i(s; n) - \sigma_i^k\right\}$ is bounded for all $n > 0$ and $k = 2, -2$;

(III) Either $t_i(n)$ and $s_i(n)$ from the same sample are a pair of mutually independent random variables, or P_i has a finite fourth moment.

We shall follow the two-stage procedure given by (1.2) - (1.6) with $s_i(m_i)$ as given by (2.4) instead of (1.1). Then we have

THEOREM 2. Let a_1, a_2 and ρ remain fixed while m_1, m_2 and A become infinite in such a way that

(2.5) $\left\{ \text{(2.1) holds, and } A/m_i^{(1+\alpha)/2} \text{ is bounded.} \right.$

Then

(2.6) $E\left\{ t_1(\tilde{n}_1) - t_2(\tilde{n}_2)\right\} \longrightarrow \Theta_1 - \Theta_2$,

(2.7) $\tilde{V}(A)/V^{o}(A) \longrightarrow 1.$

PROOF. We shall first prove the following statements:

(2.8) Assumption (II) also holds for $k = 1$ and -1;

(2.9) $Am_i\left\{ E(\tilde{n}_i^{-k}) - (n_i^{o})^{-k}\right\}$ is bounded for $k = 1, 2$;

(2.10) $A^2 m_i^2 E(\tilde{n}_i^{-4}) \longrightarrow 0.$

To prove (2.8) for $k = -1$, we note that

$$\int_{C_i(\epsilon)} s^{-1}\, dF_i(s;\, n) \leq \left\{ \int_{C_i(\epsilon)} s^{-2}\, dF_i(s;\, n) \right\}^{1/2} = \sigma_i^{-1} \left\{ 1 + O(n^{-1}) \right\};$$

and since $s_i^{-1} = \sigma_i^{-1} \left\{ 1 + (s_i^2\sigma_i^{-2} - 1) \right\}^{-1/2} \geq \sigma_i^{-1} \left\{ 1 - (1/2)(s_i^2\sigma_i^{-2} - 1) \right\}$

for $s_i \epsilon C_i(\epsilon)$, we have

$$\int_{C_i(\epsilon)} s^{-1} dF_i(s;n) \geq \sigma_i^{-1} \left\{ (3/2) \int_{C_i(\epsilon)} dF_i(s;n) - (1/2) \int_{C_i(\epsilon)} s^2\sigma_i^{-2} dF_i(s;n) \right\}$$

$$= \sigma_i^{-1} \left\{ 1 + O(n^{-1}) \right\} \text{ , by (I) and (II).}$$

Hence (2.8) is proved for $k = -1$, and in the same way for $k = 1$. We need only prove (2.9) and (2.10) for $i = 1$, since the proof for $i = 2$ is similar. Moreover, from the definitions of $\tilde{n}_i$ and n_i^*, it is evident that (2.9) and (2.10) hold if and only if they hold with $\tilde{n}_i$ replaced by n_i^*.

Let $m = (m_1, m_2)$ and set

$$v_m = s_2(m_2)/s_1(m_1), \tag{2.11}$$

and

$$v_m^* = \begin{cases} N_m & \text{if } v_m > N_m = (A - a_1m_1)/ca_1m_1), \\ \delta_m & \text{if } v_m < \delta_m = a_2m_2/(cA - ca_2m_2), \\ v_m & \text{if } \delta_m \leq v_m \leq N_m. \end{cases} \tag{2.12}$$

Then

$$\delta_m = O(m_iA^{-1}) \longrightarrow 0 \text{ and } N_m = O(m_i^{-1}A) \longrightarrow \infty \ ; \tag{2.13}$$

and

$$(n^*)^{-1} = a_1(1 + cv_m^*)/A. \tag{2.14}$$

Consequently, to prove (2.9), we need only show that

$$m_1 \left\{ E(v_m^*)^k - \rho^k \right\} \text{ is bounded for } k = 1, 2. \tag{2.15}$$

Let us choose an ϵ to satisfy (II); we can, by (2.13), choose m_1, m_2, A large enough so that

$$\delta_m \leq (\sigma_2 - \epsilon)/(\sigma_1 + \epsilon) \quad \text{and} \quad N_m \geq (\sigma_2 + \epsilon)/(\sigma_1 - \epsilon),$$

and hence such that

$$\left\{ s_i(m_i) \epsilon C_i(\epsilon), \quad i = 1, 2 \right\} \Longrightarrow \delta_m \leq v_m \leq N_m.$$

Under these circumstances, we have from (2.12)

$$0 \leq E(v_m^*)^k - \int_{C_1(\epsilon)} \int_{C_2(\epsilon)} s_1^{-k} s_2^k dF_2(s_2; m_2) dF_1(s_1; m_1) \leq N_m^k p(m_1, m_2), \tag{2.16}$$

where $p(m_1, m_2) = 1 - \prod_{i=1,2} \Pr \left\{ s_i(m_i) \epsilon C_i(\epsilon) \right\} = O(m^{-\alpha})$.

By (II), the second term in the middle in (2.16) is $\rho^k + O(m^{-1})$, and by (2.13) the last term is $O(A^k m^{-k} m^{-\alpha}) = O(m^{-1})$ for $k = 1, 2$. Thus, we have (2.15) and hence (2.9). We can prove (2.10) similarly.

Now, let

$$\left\{ \begin{aligned} T_1 &= m_1 t_1(m_1)/\tilde{n}_1 - m_2 t_2(m_2)/\tilde{n}_2, \\ T_2 &= (\tilde{n}_1 - m_1) t_1(\tilde{n}_1 - m_1)/\tilde{n}_1 - (\tilde{n}_2 - m_2) t_2(\tilde{n}_2 - m_2)/\tilde{n}_2 \end{aligned} \right. \tag{2.17}$$

where $(\tilde{n}_i - m_i)t_i(\tilde{n}_i - m_i) = \sum_{j=m_i+1}^{\tilde{n}_i} X_{i,j}$.

Then

(2.18) $$E\left\{t_1(\tilde{n}_1) - t_2(\tilde{n}_2)\right\} = E\left\{T_1\right\} + E_{\tilde{n}_1} E\left\{T_2 \mid \tilde{n}_1\right\} .$$

Since $\tilde{n}_i$ depends only on $\left\{X_{i,j}, \ j = 1, 2, \ldots, m_i\right\}$ and $t_i(\tilde{n}_i - m_i)$, for fixed $\tilde{n}_1$, only on $\left\{X_{i,j}, \ j = m_i+1, \ldots, \tilde{n}_i\right\}$, we have

(2.19) $$\Pr\left\{t_i(\tilde{n}_i - m_i) \leq x_i, \ i = 1, 2 \mid \tilde{n}_i = n_i\right\} = \prod_{i=1,2} \Pr\left\{t_i(n_i - m_i) \leq x_i\right\} .$$

Therefore,

$$E\left\{T_2 \mid \tilde{n}_1\right\} = (\tilde{n}_1 - m_1)\theta_1/\tilde{n}_1 - (\tilde{n}_2 - m_2)\theta_2/\tilde{n}_2;$$

so that

$$E_{\tilde{n}_1} E\left\{T_2 \mid \tilde{n}_1\right\} = \theta_1 - \theta_2 - m_1\theta_1 E(1/\tilde{n}_1) + m_2\theta_2 E(1/\tilde{n}_2)$$

$$\longrightarrow \theta_1 - \theta_2 \text{ by (2.5) and (2.9).}$$

Moreover,

$$m_i E\left\{t_i(m_i)/\tilde{n}_i\right\} \leq m_i\left\{Et_i^2(m_i)E(1/\tilde{n}_i^2)\right\}^{1/2} = A^{-1}m_i O(1)\left\{A^2 E(1/\tilde{n}_i^2)\right\}^{1/2}$$

$$\longrightarrow 0 \text{ by (2.5) and (2.9).}$$

Therefore, $ET_1 \longrightarrow 0$ and hence, we have (2.6).

Finally,

(2.20) $$A\tilde{V}(A) = A \operatorname{Var}(T_1 + T_2) = A \operatorname{Var} T_1 + A \operatorname{Var} T_2 + 2A \operatorname{Cov}(T_1 \ T_2).$$

But

$$(2.21) \quad A \text{ Var } T_1 = A \text{ Var } \left\{ m_1 t_1(m_1)/\tilde{n}_1 - m_2 t_2(m_2)/\tilde{n}_2 \right\}$$

$$\leq 2Am_1^2 \text{ Var } \left\{ t_1(m_1)/\tilde{n}_1 \right\} + 2Am_2^2 \text{ Var } \left\{ t_2(m_2)/\tilde{n}_2 \right\}$$

$$\leq 2\sum_i Am_i^2 \text{ Var } \left\{ [t_i(m_i) - \theta_i]/\tilde{n}_i + \theta_i/\tilde{n}_i \right\}$$

$$\leq 4 \sum_i Am_i^2 E \left\{ [t_i(m_i) - \theta_i]^2/\tilde{n}_i^2 \right\}$$

$$+ 4 \sum_i Am_i^2 \theta_i^2 \text{ Var } (1/\tilde{n}_i).$$

From (2.5) and (2.9), we know that $Am_i^2 \text{ Var } (1/\tilde{n}_i) \longrightarrow 0$. As for the other term on the right hand side in (2.21), if $t_i(m_i)$ and $s_i(m_i)$ are independent, we have

$$Am_i^2 E \left\{ [t_i(m_i) - \theta_i]^2/\tilde{n}_i^2 \right\} = Am_i^2 E \left\{ [t_i(m_i) - \theta_i]^2 \right\} E(1/\tilde{n}_i^2)$$

$$= Am_i \sigma_i^2 E(1/\tilde{n}_i^2)$$

$$\longrightarrow 0 \text{ by } (2.9).$$

If $t_i(m_i)$ and $s_i(m_i)$ are not independent, we still have

$$Am_i^2 E \left\{ [t_i(m_i) - \theta_i]^2/\tilde{n}_i^2 \right\} \leq Am_i^2 \left\{ E[t_i(m_i) - \theta_i]^4 E(1/\tilde{n}_i^4) \right\}^{1/2}$$

$$= \left\{ m_i^2 E[t_i(m_i) - \theta_i]^4 A^2 m_i^2 E(1/\tilde{n}_i^4) \right\}^{1/2}$$

$$\longrightarrow 0 \text{ by (III) and } (2.10).$$

Hence, $A \text{ Var}(T_1) \longrightarrow 0$.

We shall see below that $A \text{ Var}(T_2)$ is bounded, so that

$$A\ \mathrm{Cov}\ (T_1, T_2) \leq \left\{ A\ \mathrm{Var}\ (T_1)\ A\ \mathrm{Var}\ (T_2) \right\}^{1/2} \longrightarrow 0.$$

Therefore,

$$(2.22)\ \lim A\tilde{V}(A) = \lim A\ \mathrm{Var}\ T_2 = \lim A\ E_{\tilde{n}_1} \mathrm{Var}(T_2 \mid \tilde{n}_1) + \lim A\ \mathrm{Var}\ E(T_2 \mid \tilde{n}_1)$$

$$= \lim AE_{\tilde{n}_1} \mathrm{Var} \left\{ (\tilde{n}_1 - m_1)t_1(\tilde{n}_1 - m_1)/\tilde{n}_1 - (\tilde{n}_2 - m_2)t_2(\tilde{n}_2 - m_2)/\tilde{n}_2 \right\}$$

$$+ \lim A\ \mathrm{Var} \left\{ (\tilde{n}_1 - m_1)\theta_1/\tilde{n}_1 - (\tilde{n}_2 - m_2)\theta_2/\tilde{n}_2 \right\}$$

$$= \lim AE_{\tilde{n}_1} \left\{ \sum_i (\tilde{n}_i - m_i)\sigma_i^2/\tilde{n}_i^2 \right\} + \lim A\ \mathrm{Var} \left\{ m_2\theta_2/\tilde{n}_2 - m_1\theta_1/\tilde{n}_1 \right\}$$

$$= \lim \sum_i \sigma_i^2 AE(1/\tilde{n}_i) - \lim \sum_i Am_i\sigma_i^2 E(1/\tilde{n}_i^2) + \lim A\ \mathrm{Var}\left\{ m_2\theta_2/\tilde{n}_2 - m_1\theta_1/\tilde{n}_1 \right\},$$

The last two terms are zero on account of (2.5) and (2.9); and from (2.9), we see that the first term on the right hand side of (2.22) is the required limit in (2.7).

If we use sample sizes $n'_i = (A/a_i)a_i^{1/2}/\sum_j a_j^{1/2}$, which by (0.2) is what we would be led to do if we thought that $\sigma_1 = \sigma_2$, the variance of the estimate of $\theta_1 - \theta_2$ would be

$$V' = V^o \left\{ 1 + c(1 - \rho)^2/(1 + \rho c)^2 \right\} > V^o \quad \text{for} \quad \rho > 1.$$

Hence, asymptotically, the two-stage procedure is more efficient than this one stage procedure if $\rho > 1$.

EXAMPLES. (1) If the P_i are normal, the conditions of Theorem 2 are satisfied for every $\alpha > 0$, and hence in (2.5) A may increase as any power of m_i. We have seen in Theorem 1 that it is actually not necessary to restrict A to be of the order of a power of m_i.

(2) If the P_i are Poisson, $\sigma_i^2 = \Theta_i$, and $s_i^2(n) = t_i(n)$. From the fact that $nt_i(n)$ has a Poisson distribution, it can be seen that the conditions of Theorem 2 are satisfied for every $\alpha > 0$.

(3) If the P_i are binomial, with $\Pr\left\{X_i = 1\right\} = \Theta_i$ and $\Pr\left\{X_i = 0\right\} = 1 - \Theta_i$, we have $\sigma_i^2 = \Theta_i(1 - \Theta_i)$ and $s_i^2(n) = t_i(n)\left\{1 - t_i(n)\right\}$. Using the fact that $nt_i(n)$ is a binomial variate, we can show that the conditions of the theorem are satisfied for every $\alpha > 0$.

(4) If we do not know the forms of P_i, we would use the estimate $s_i(m_i)$ of σ_i given by (1.1). If we know that P_i has a sufficient number, say 8, of moments finite, we can show that Theorem 2 is true for the procedure given by (1.1) - (1.6).

REFERENCES

[1] J. Putter, "Sur une méthode de double échantillonnage, etc.", Revue de l'Institut International de Statistique, (1951), part 3, pp. 1-8.

[2] H. Robbins, "Some aspects of the sequential design of experiments", Bull. Am. Math. Soc., Vol. 58 (1952), pp. 527-535.