The Annals of Statistics
1977, Vol. 5, No. 5, 1047-1050

ESTIMATING THE COMMON MEAN OF SEVERAL NORMAL POPULATIONS

BY THOMAS E. NORWOOD, JR. AND KLAUS HINKELMANN

Norwich Pharmacal Company and Virginia Polytechnic Institute and State University

For $x_i \sim N(\mu, \sigma_i^2)$ $(i = 1, 2, \cdots, n)$ and the x_i's independent, this paper gives necessary and sufficient conditions under which the weighted average of the x_i's, with weights proportional to inverses of the sample variances, has uniformly smaller variance than any of the x_i's.

1. Introduction and summary. If n independent samples are available from normal populations with a common mean but different variances, a natural estimator for the common mean is the weighted average of the sample means with the weights proportional to the inverses of the sample variances. To be more specific, if $x_1, \cdots, x_n, s_1^2, \cdots, s_n^2$ are mutually independent random variables such that $x_i \sim N(\mu, \sigma_i^2)$ and $m_i s_i^2/\sigma_i^2 \sim \chi^2_{m_i}$ for $i = 1, \cdots, n$, then the estimator

$$\hat{\mu} = \sum_{i=1}^n (x_i/s_i^2)/(\sum_{i=1}^n 1/s_i^2) \tag{1.1}$$

is a natural estimator for μ. When the variances are known the minimum variance unbiased estimator for μ is

$$\hat{\mu}^* = \sum_{i=1}^n (x_i/\sigma_i^2)/(\sum_{i=1}^n 1/\sigma_i^2),$$

so the appeal of $\hat{\mu}$ is apparent. Many authors investigated the properties of this estimator or variations of it; among these are Cochran (1937), Meier (1953), Cochran and Carroll (1953), Graybill and Deal (1959), Zacks (1966), Bement and Williams (1969), and Rao and Subrahmaniam (1971). In this paper necessary and sufficient conditions are given under which $\hat{\mu}$ has uniformly smaller variance than any one of the x_i's. These conditions are different from those given by Graybill and Deal (1959) for $n = 2$ and from a correction to their results quoted by Bement and Williams (1969).

2. Two lemmas. The following lemmas are needed to establish the main result.

LEMMA 1. *Let W be a continuous random variable with support $(0, \omega)$ and X be a continuous random variable with support (ϕ, ∞), $\phi < 0$. If for every $t > 0$ and $\phi < a \leqq 0 \leqq b$, $P(W > t \mid X = a) \geqq P(W > t \mid X = b)$ then $E(X) \leqq 0$ implies $E(WX) \leqq 0$ provided $E(X)$ exists.*

LEMMA 2. *Let Y_i $(i = 1, 2, 3)$ be mutually independent random variables with*

Received April 1976; revised March 1977.

AMS 1970 *subject classification.* 62F10.

Key words and phrases. Normal populations, common mean, different unknown variances, estimation, weighted average of sample means.

distribution function F_i *and density function* f_i. *If* $f_i(x) > 0$ *for* $0 < x < \infty$, $= 0$ *otherwise, is continuous, then*

$$P(Y_1/Y_2 < \xi \mid Y_1/Y_3 = \eta) = \frac{\int_0^\infty x[1 - F_2(x/\xi)]f_3(x/\eta)f_1(x)\,dx}{\int_0^\infty xf_3(x/\eta)f_1(x)\,dx}.$$

The proof of Lemma 1 is based on the relationships

$$E(W \mid X = a) = \int_\phi^\infty P(W > t \mid X = a)\,dt, \quad \phi \leqq a$$

and

$$E(WX) = \int_\phi^\infty xE(W \mid X = x)g(x)\,dx,$$

where g is the density of X.

To prove Lemma 2 one obtains the joint distribution function of Y_1/Y_2 and Y_1/Y_3, then the required conditional probability in the usual manner.

3. Main theorem.

THEOREM. (i) *The estimator* $\hat{\mu}$ *of* (1.1) *is unbiased for* μ. (ii) $\operatorname{Var}(\hat{\mu}) < \sigma_i^2$ *for all values of* σ_i^2 $(i = 1, \cdots, n)$ *if and only if either*

(A) $m_i > 9$ $(i = 1, \cdots, n)$

or

(B) $m_i = 9$ *for some* i *and* $m_j > 17$ $(j = 1, \cdots, n; j \neq i)$.

PROOF. Let $\rho_i = \sigma_1^2/\sigma_i^2$ and $\hat{\rho}_i = s_1^2/s_i^2$ for $i = 1, \cdots, n$ so that

$$\hat{\mu} = \sum_{i=1}^n x_i\hat{\rho}_i / \sum_{i=1}^n \hat{\rho}_i \quad \text{and} \quad \hat{\mu}^* = \sum_{i=1}^n x_i\rho_i / \sum_{i=1}^n \rho_i.$$

(i) Since the $\hat{\rho}_i$'s are distributed independently of the x_i's, the conditional expectation of $\hat{\mu}$ given $\hat{\rho}_1, \cdots, \hat{\rho}_n$ is μ, and $\hat{\mu}$ is unbiased.

(ii) *Sufficient conditions*. It is not difficult to show (Norwood, 1974) that

$$\hat{\mu} - \hat{\mu}^* = \sum_{i<i'} (x_i - x_{i'})(\hat{\rho}_i\rho_{i'} - \rho_i\hat{\rho}_{i'})/(\textstyle\sum_{i=1}^n \rho_i)(\sum_{i=1}^n \hat{\rho}_i), \tag{3.1}$$

and

$$\operatorname{Cov}(\hat{\mu} - \hat{\mu}^*, \hat{\mu}^*) = E\{(\hat{\mu} - \hat{\mu}^*)\hat{\mu}^*\} = 0. \tag{3.2}$$

Hence,

$$\operatorname{Var}(\hat{\mu}) = \operatorname{Var}(\hat{\mu}^*) + E\{(\hat{\mu} - \hat{\mu}^*)^2\}. \tag{3.3}$$

Clearly,

$$\operatorname{Var}(\hat{\mu}^*) = \sigma_1^2(\textstyle\sum_{i=1}^n \rho_i)^{-1}, \tag{3.4}$$

and evaluating $E\{(\hat{\mu} - \hat{\mu}^*)^2\}$, which may be interpreted as the part of the variation of $\hat{\mu}$ which is due to estimating the ρ_i's, we first obtain

$$\begin{aligned} E\{[\textstyle\sum_{i<i'} (x_i - x_{i'})(\hat{\rho}_i\rho_{i'} - \rho_i\hat{\rho}_{i'})]^2 \mid \hat{\rho}_1, \cdots, \hat{\rho}_n\} \\ = \sigma_1^2(\textstyle\sum_{i=1}^n \rho_i) \sum_{i<i'} \rho_i\rho_{i'}(f_i - f_{i'})^2, \end{aligned} \tag{3.5}$$

where $f_i = \hat{\rho}_i/\rho_i$, so that from (3.1) and (3.5)

$$E\{(\hat{\mu} - \hat{\mu}^*)^2\} = \sigma_1^2(\textstyle\sum_{i=1}^n \rho_i)^{-1}E\{\sum_{i<i'} \rho_i\rho_{i'}(f_i - f_{i'})^2/(\sum_{i=1}^n \rho_i f_i)^2\}. \tag{3.6}$$

Now, after some manipulation and using $\rho_1 = \hat{\rho}_1 = 1$,

$$E\{\sum_{i<i'} \rho_i \rho_{i'} (f_i - f_{i'})^2 / (\sum_{i=1}^n \rho_i f_i)^2\} < \sum_{i=2}^n \rho_i [1 + \sum_{j=1; j\neq i}^n \rho_j E\{(f_i^2 - 2f_i)/(\sum_{i=1}^n \rho_i f_i)^2\}] ,$$

so from (3.3), (3.4), and (3.6)

$$\text{(3.7)} \qquad \mathrm{Var}\,(\hat{\mu}) < \sigma_1^2 (\sum_{i=1}^n \rho_i)^{-1} [\sum_{i=1}^n \rho_i + \sum_{i=2}^n \rho_i (\sum_{j=1; j\neq i}^n \rho_j) E\{(f_i^2 - 2f_i)/(\sum_{i=1}^n \rho_i f_i)^2\}] .$$

Obviously, $\mathrm{Var}\,(\hat{\mu}) < \sigma_1^2$ if $E\{(f_i^2 - 2f_i)/(\sum_{i=1}^n \rho_i f_i)^2\} < 0$ for $i = 2, \cdots, n$. Letting $X = f_i^2 - 2f_i$, $W = (\sum_{i=1}^n \rho_i f_i)^{-2}$ and assuming $m_i > 4$, Lemma 1 may be applied provided

$$\text{(3.8)} \qquad P(W > t \mid X = a) > P(W > t \mid X = b)$$

for all $t > 0$ and $-1 < a < 0 < b$. To demonstrate this we show that

$$\text{(3.9)} \qquad P(W > t \mid f_i = a_1) > P(W > t \mid f_i = a_2) , \qquad \text{for } 0 < a_1 \leqq a_2 , \quad \text{using Lemma 2} .$$

It is not difficult to show that (3.9) implies (3.8), so that, from Lemma 1, a sufficient condition for $\hat{\mu}$ to have uniformly smaller variance than x_1 is

$$E\{f_i^2 - 2f_i\} < 0 , \qquad \text{for} \quad i = 2, \cdots, n .$$

Since $f_i \sim F_{m_1, m_i}$ the above condition is equivalent to

$$\frac{m_i(m_1 + 2)}{m_1(m_i - 4)} \leqq 2 , \qquad \text{for} \quad i = 2, \cdots, n .$$

Hence,

$$\mathrm{Var}\,(\hat{\mu}) < \min\,(\sigma_1^2, \cdots, \sigma_n^2)$$

if for all pairs (m_i, m_j) $(i, j = 1, 2, \cdots, n; i \neq j)$

$$\frac{m_i(m_j + 2)}{m_j(m_i - 4)} \leqq 2 .$$

This holds whenever all the m_i's are greater than 9 or one of the m_i's is equal to 9 and the others are larger than 17.

Necessary conditions. To prove that these conditions are also necessary, suppose that for some $j = 2, \cdots, n$

$$E\{f_j^2 - 2f_j\} > 0$$

or this expectation fails to exist. One can show that then $\mathrm{Var}\,(\hat{\mu}) > \sigma_1^2$ for sufficiently small values of $\rho_2, \cdots, \rho_n$.

Acknowledgment. The authors would like to thank Dr. Theodore Bolis for his suggestions.

REFERENCES

[1] BEMENT, T. R. and WILLIAMS, J. S. (1969). Variance of weighted regression estimators when sampling errors are independent and heteroscedastic. *J. Amer. Statist. Assoc.* **64** 1369–1382.

[2] COCHRAN, W. G. (1937). Problems arising in the analysis of a series of experiments. *J. Roy. Statist. Soc.* **4** 102–118.

[3] COCHRAN, W. G. and CARROLL, S. P. (1953). A sampling investigation of the efficiency of weighting inversely as the estimated variance. *Biometrics* **9** 447–459.

[4] GRAYBILL, F. A. and DEAL, R. B. (1959). Combining unbiased estimators. *Biometrics* **15** 543–550.

[5] MEIER, P. (1953). Variance of a weighted mean. *Biometrics* **9** 59–73.

[6] NORWOOD, T. E. (1974). The analysis of a series of incomplete block designs under the assumption of additivity. Ph. D. dissertation, Virginia Polytechnic Inst. and State Univ.

[7] RAO, J. N. K. and SUBRAHMANIAM, K. (1971). Combining independent estimators and estimation in linear regression with unequal variances. *Biometrics* **27** 971–990.

[8] ZACKS, S. (1966). Unbiased estimation of the common mean. *J. Amer. Statist. Assoc.* **61** 467–476.

NORWICH PHARMACAL COMPANY
105 SOUTH BROAD STREET
NORWICH, NEW YORK 13815

VIRGINIA POLYTECHNIC INSTITUTE
AND STATE UNIVERSITY
BLACKSBURG, VIRGINIA 24061