The Annals of Mathematical Statistics
1968, Vol. 39, No. 5, 1486–1492

A NOTE ON ROBUST ESTIMATION IN ANALYSIS OF VARIANCE

By Emil Spjøtvoll

University of Oslo

1. Introduction. Consider the c-sample model, in which the observations are

$$X_{i\alpha} = \xi_i + U_{i\alpha}, \qquad \alpha = 1, 2, \cdots, n_i; \qquad i = 1, 2, \cdots, c, \tag{1.1}$$

where the variables $U_{i\alpha}$ are independently distributed with cumulative distribution function F. Let

$$Y_{ij} = \text{med}\,(X_{i\alpha} - X_{j\beta}) \tag{1.2}$$

be the median of the $n_i n_j$ differences $X_{i\alpha} - X_{j\beta}(\alpha = 1, 2, \cdots, n_i, \beta = 1, 2, \cdots, n_j)$. It has been shown by Hodges and Lehmann [2] that the estimate Y_{ij} of $\xi_i - \xi_j$ has more robust efficiency than the standard estimate $T_{ij} = X_{i\cdot} - X_{j\cdot}$, where $X_{i\cdot} = \sum X_{i\alpha}/n_i$.

The estimates Y_{ij} do not satisfy the linear relations satisfied by the differences they estimate. To remedy this, the *raw* estimates Y_{ij} were by Lehmann [3] replaced by *adjusted* estimates Z_{ij} of the form $\hat{\xi}_i - \hat{\xi}_j$. This was done by minimizing the sum of squares

$$\textstyle\sum_{i \neq j} (Y_{ij} - (\xi_i - \xi_j))^2 \tag{1.3}$$

giving (see [3])

$$Z_{ij} = Y_{i\cdot} - Y_{j\cdot} \tag{1.4}$$

where $Y_{i\cdot} = (1/c) \sum Y_{ij}$ and where Y_{ii} is defined to be zero for all i.

The purpose of this note is to argue that in the sum of squares (1.3) there should be used weights according to the number of observations on which the different Y_{ij} are based.

For purpose of reference we state a theorem of Lehmann. Let the sample sizes n_i tend to infinity in such a way that $n_i \to \rho_i N (N = \sum n_i)$, $0 < \rho_i < 1$. Then we have the following theorem (Theorem 2 of [3]).

Theorem 1. *Let the density f of F satisfy the regularity conditions of Lemma* 3(a) *of* [1].

(i) *The joint distribution of* $(V_1, V_2, \cdots, V_{c-1})$ *where*

$$V_i = N^{\frac{1}{2}}(Y_{ic} - (\xi_i - \xi_c))$$

is asymptotically normal with zero mean and covariance matrix

$$\text{Var}\,(V_i) = (\tfrac{1}{12})(1/\rho_i + 1/\rho_c)/(\textstyle\int f^2(x)\,dx)^2$$

$$\text{Cov}\,(V_i, V_j) = (\tfrac{1}{12}\rho_c)/(\textstyle\int f^2(x)\,dx)^2.$$

Received 26 September 1967.