ESTIMATES OF REGRESSION PARAMETERS BASED ON RANK TESTS[1]

By J. N. Adichie

University of Nigeria, Nsukka

0. Introduction and summary. In the linear regression model $Y_j = \alpha + \beta x_j + Z_j$, it is usual to estimate α and β by the method of least squares. This method has, among other things, the nice property of providing "best" linear unbiased estimates, under very general conditions. Various other methods of estimation of the parameters are well known, see for example [12] and [14]. Most of these methods however make use of the actual values of the observations, and the estimates they yield are generally vulnerable to gross errors. For some alternative approaches to the problem, see [7], [11] and [13].

In a recent paper [9], Hodges and Lehmann proposed a general method of obtaining robust point estimates for the location parameter, from statistics used to test the hypothesis that this parameter has a specified value. In Section 1 of this paper, this method is used to define point estimates $\hat{\alpha}$ and $\hat{\beta}$ of α and β, in terms of certain test statistics. It is shown that the least squares estimates are obtainable as special cases from the general method of estimation discussed. In Section 2, the existence of 'rank score' estimates is proved, and in Section 3, computing techniques are given and illustrated with an example. Both the small sample and asymptotic properties of the estimates are discussed. It is shown, for example, that the joint distribution of the estimates $\hat{\alpha}$ and $\hat{\beta}$ is symmetric with respect to the parameter point (α, β)—and hence that $\hat{\alpha}$ and $\hat{\beta}$ are unbiased—if the underlying distribution of the observations is symmetric. In Section 5, the joint asymptotic normality of $\hat{\alpha}$ and $\hat{\beta}$ is proved, and in Section 6, it is shown that the asymptotic efficiency of $(\hat{\alpha}, \hat{\beta})$ is the same as the Pitman efficiency of the rank tests [1], on which they are based, relative to the classical tests. Finally in Section 7, the $(\hat{\alpha}, \hat{\beta})$-estimates are compared with the Brown and Mood median estimates with respect to their efficiencies.

1. Estimation of α and β. As in [1], let $Y_1, \cdots, Y_n$ be independent random variables with distributions

$$P_{\alpha\beta}(Y_j \leqq y) = F(y - \alpha - \beta x_j) \tag{1.1}$$

where x_j are the known regression constants that are not all equal and which satisfy the limiting conditions given in [1], and $P_{\alpha\beta}$ denotes the probability computed for the parameter values α and β. As before we shall assume that the underlying distribution function F belongs to a class $\mathfrak{F}$ of absolutely continuous

Received 1 September 1966; revised 30 January 1967.

[1] This research was based in part on the author's Ph.D. dissertation submitted to the University of California, Berkeley; and was prepared with the partial support of the Agency for International Development under contract MSU AIDc-1398, and of the National Science Foundation, Grant GP-5059.

symmetric distribution functions, the densities of which are also absolutely continuous and square integrable. With these regularity conditions, we shall define estimates $\hat{\alpha}$ and $\hat{\beta}$ which are similar to the Hodges and Lehmann estimates for shift [9].

Let $T_1(Y_1, \cdots, Y_n)$ and $T_2(Y_1, \cdots, Y_n)(T_1(Y)$ and $T_2(Y)$ for short) be two statistics for testing hypotheses about α and β in (1.1). Assume that T_1 and T_2 satisfy the following two conditions:

(A) for fixed b, $T_1(y + a + bx)$ is non-decreasing in a; and for every a, $T_2(y + a + bx)$ is non-decreasing in b, for each y and x. Here $y + a + bx$ stands for $(y_1 + a + bx_1, \cdots, y_n + a + bx_n)$.

(B) When $\alpha = \beta = 0$, the distributions of $T_1(Y)$ and $T_2(Y)$ are symmetric about fixed points μ and ν, independent of $F \varepsilon \mathfrak{F}$.

Let

$$\beta^* = \sup \{b: T_2(y - a - bx) > \nu, \text{ for all } a\}, \tag{1.2}$$

$$\beta^{**} = \inf \{b: T_2(y - a - bx) < \nu, \text{ for all } a\};$$

$$\hat{\beta} = \tfrac{1}{2}(\beta^* + \beta^{**}); \tag{1.3}$$

$$\alpha^* = \sup \{a: T_1(y - a - \hat{\beta}x) > \mu\}, \tag{1.4}$$

$$\alpha^{**} = \inf \{a: T_1(y - a - \hat{\beta}x) < \mu\};$$

$$\hat{\alpha} = \tfrac{1}{2}(\alpha^* + \alpha^{**}). \tag{1.5}$$

For suitable functions T_1 and T_2, we propose $\hat{\alpha}$ and $\hat{\beta}$ as estimates of α and β.

It may be remarked that many existing estimates of α and β, belong to the class of (1.3)- and (1.5)-estimates. In particular, the least squares estimates $\tilde{\alpha}$ and $\tilde{\beta}$ are obtainable as special cases of $\hat{\alpha}$ and $\hat{\beta}$. To see this, take $T_1(Y) = \sum_j Y_j$ and $T_2(Y) = \sum_j (x_j - \bar{x}_n)(Y_j - \bar{Y}_n)$ where as before $\bar{x}_n = n^{-1}\sum_j x_j$, and all the summations are from 1 to n. With these choices of the functions T_1 and T_2, it is easy to see that conditions (A) and (B) are satisfied with $\mu = \nu = 0$. Furthermore,

$$\begin{aligned} &\sup \{b: T_2(y - a - bx) > 0, \text{ for all } a\} \\ &\qquad = \inf \{b: T_2(y - a - bx) < 0, \text{ for all } a\} \\ &\qquad = \{\textstyle\sum_j (x_j - x_n)(y_j - \bar{y}_n)\}/\{\sum_j (x_j - \bar{x}_n)^2\} = \tilde{\beta}. \end{aligned} \tag{1.6}$$

In the same way,

$$\begin{aligned} \sup \{a: T_1(y - a - \tilde{\beta}x) > 0\} &= \inf \{a: T_1(y - a - \tilde{\beta}x) < 0\} \\ &= (\bar{y}_n - \tilde{\beta}\bar{x}_n) = \tilde{\alpha}. \end{aligned} \tag{1.7}$$

2. Estimates based on rank tests. Since our main interest is in robust estimates, we shall be primarily concerned with $\hat{\alpha}$ and $\hat{\beta}$ based on rank (or mixed rank) statistics. As in [1], we shall need the following functions:

$$\psi(u) = -[g'(G^{-1}(\tfrac{1}{2}u + \tfrac{1}{2}))/g(G^{-1}(\tfrac{1}{2}u + \tfrac{1}{2}))], \qquad 0 < u < 1, \tag{2.1}$$

$$(2.2) \qquad \psi_0(u) = -[g'(G^{-1}(u)/g(G^{-1}(u))], \qquad 0 < u < 1,$$

$$(2.3) \qquad \varphi_0(u) = -[f'(F^{-1}(u))/f(F^{-1}(u))], \qquad 0 < u < 1,$$

where G^{-1} is the inverse of G, and G is any distribution function belonging to the class $\mathfrak{F}$. Consider the following pair of statistics:

$$(2.4) \qquad T_1(Y) = n^{-1}\sum_j \psi_n(\mathfrak{R}_j/n+1) \text{ Sign } Y_j$$

and

$$(2.5) \qquad T_2(Y) = n^{-1}\sum_j(x_j - \bar{x}_n)\psi_{0n}(R_j/n+1)$$

where $\mathfrak{R}_j$ is the rank of $|Y_j|$ in the sequence of absolute values $|Y_1|, \cdots, |Y_n|$ of the n observations, while R_j is the rank of Y_j in the ordered sample $V_1 < \cdots < V_n$, i.e. $Y_j = V_{R_j}$, $j = 1, \cdots, n$, and

$$(2.6) \qquad \psi_n(u) = \psi(j/n+1),$$

$$\psi_{0n}(u) = \psi_0(j/n+1) \quad \text{for} \quad (j-1)/n < u \leqq j/n.$$

Observe that the statistic T_1 of (2.4) is the same as the one studied in [1], while T_2 of (2.5) is studied by Hájek [6]. It can be shown by arguments similar to those in [9], that with (2.4) and (2.5) as choices for T_1 and T_2, conditions (A) and (B) are satisfied, and hence that the estimates based on them, exist and are well defined. If in (2.1) and (2.2) we choose G to be the logistic distribution function, then $\psi(u)$ becomes u and $T_1(Y)$ and $T_2(Y)$ of (2.4) and (2.5) coincide with the Wilcoxon one-sample and two-sample statistics respectively. Denoting by $\hat{\alpha}_w$ and $\hat{\beta}_w$ the resulting estimates, we then have that

$$(2.7) \qquad T_1(y - a - \hat{\beta}_w x) = [n(n+1)]^{-1}\sum_j \hat{\mathfrak{R}}_j \text{ Sign } (y_j - a - \hat{\beta}_w x_j)$$

$$= [n(n+1)]^{-1}[2N^+ - n(n+1)/2]$$

where N^+ is the number of pairs (i, j) with $1 \leqq i \leqq j \leqq n$, such that $y_i + y_j - \hat{\beta}_w(x_i + x_j) - 2a$ is positive, and $\hat{\mathfrak{R}}_j$ is the rank of $|y_j - a - \hat{\beta}_w x_j|$ in the sequence of absolute values $|y_1 - a - \hat{\beta}_w x_1|, \cdots, |y_n - a - \hat{\beta}_w x_n|$. The estimate $\hat{\alpha}_w$ is then given by

$$(2.8) \qquad \hat{\alpha}_w = \text{med}_{i \leqq j} \tfrac{1}{2}\{Y_i + Y_j - \hat{\beta}_w(x_i + x_j)\},$$

where $\hat{\beta}_w$ is obtained from (1.2) and (1.3) with

$$(2.9) \qquad T_2(Y) = n^{-1}\sum_j(x_j - \bar{x}_n)(R_j/n+1).$$

3. Computation of the Wilcoxon estimate $\hat{\beta}_w$. An explicit expression, for $\hat{\beta}_w$ in terms of Y's does not seem to be available, without restrictive assumptions on the regression constants x_j. As an example of such an assumption if we take

$$x_j = c_1, \qquad j = 1, \cdots, k,$$

$$= c_2, \qquad j = k+1, \cdots, n,$$

the estimate $\hat{\beta}_w$ so obtained would coincide with the Hodges and Lehmann estimate for shift in the two-sample problem, i.e. $\hat{\beta}_w = \text{med } (Y - Z)$ where $Y = (Y_1, \cdots, Y_k)$ and $Z = (Y_{k+1}, \cdots, Y_n)$. However for any given constants x_j an iterative method may be used to compute $\hat{\beta}_w$, for moderate sample size n. The procedure described below depends mainly on the monotonicity of $T_2(y + bx)$ as a function of b. There are however some values of b, for which ties occur. For any such b we define

$$(3.1) \quad T_2(y + bx) = \tfrac{1}{2}[\sup_{b'<b} T_2(y + b'x) + \inf_{b'>b} T_2(y + b'x)].$$

Note that the use of mid-ranks is not acceptable here for it badly dislocates the monotonicity of $T_2(y + bx)$.

Let $y_1, \cdots, y_n$ be the values of the observations, taken at levels $x_1, \cdots, x_n$, respectively. Choose any b_0 for which there is no tie, (usually $b_0 = 0$ for a start) and rank the n differences $y_j - b_0 x_j$, $j = 1, \cdots, n$. Then compute $T_2(y - b_0 x)$. If the result is positive (negative) increase (decrease) b_0 to b_1 and compute $T_2(y - b_1 x)$. Continue this iteration increasing (decreasing) b at each step until $T_2(y - bx)$ becomes zero. The b-value that achieves this would be the estimate. If $T_2(y - bx) = 0$ for $b' \leqq b \leqq b''$, then the estimate would be $\frac{1}{2}(b' + b'')$. If on the other hand, $T_2(y - bx)$ does not assume the value zero, there would be, by condition (A), a certain b_0 say, such that for $b > b_0$, $T_2(y - bx) < 0$, while for $b < b_0$, $T_2(y - bx) > 0$.

In such a case b_0 would be the estimate. To find b_0, continue the iteration described above, in close steps of b, until $T_2(y - bx)$ changes sign. Then go back and forth to determine where the first change of sign occurs.

Example. Consider the following set of data, taken from Graybill's *Introduction to Linear Statistical Models* Vol. 1.

x	1	2	3	4	10	12	18
y	9	15	19	20	45	55	78

The results of the iterative method described above, are set out below in the form of Tables 3.1(a) and 3.1(b). The estimate $\hat{\beta}_w$ obtained from the tables has the value 4.00 while the corresponding least squares estimate $\tilde{\beta}$ has the value 4.02. Observe that for $b = 4.00$, $T_2(y - bx)$ is defined by (3.1).

4. Invariance and symmetry properties. As in [9], the estimates $\hat{\alpha}$ and $\hat{\beta}$ have useful invariance and symmetry properties in a sense to be made precise in the following two lemmas.

Lemma 4.1. *For any real a and b, the* (1.3)- *and* (1.5)*-estimates possess the following translation invariant properties:*

$$(4.1) \qquad \hat{\beta}(y + a + bx) = \hat{\beta}(y) + b$$

and

$$(4.2) \qquad \hat{\alpha}(y + a + bx) = \hat{\alpha}(y) + a.$$

Proof. (4.1) is immediate from definition (1.3) while (4.2) follows from (4.1) and (1.5).

TABLE 3.1(a)

b		$x = 1$	2	3	4	10	12	18
$b = 0$	$(y - bx) =$	9	15	19	20	45	55	78
	ranks =	1	2	3	4	5	6	7
$b = 2.0$	$(y - bx) =$	7	11	13	12	25	31	42
	rank =	1	2	4	3	5	6	7
$b = 3.0$	$(y - bx) =$	6	9	10	8	15	19	30
	rank =	1	3	4	2	5	6	7
$b = 3.99$	$(y - bx) =$	5.01	7.02	7.03	4.04	5.1	7.12	6.18
	ranks =	2	5	6	1	3	7	4
$b = 4.01$	$(y - bx) =$	4.99	6.98	6.97	3.96	4.9	6.86	5.82
	ranks =	3	7	6	1	2	5	4

TABLE 3.1(b)

$b =$	0	2	3	3.99	4.01
$T_2(y - bx) =$	78.24	77.24	75.24	19.80	−9.14

$\hat{\beta}_w = 4.00$.

From (4.1) and (4.2) it follows that, for all real a and b,

$$(4.3) \qquad P_{\alpha\beta}\{(\hat{\alpha} - \alpha), (\hat{\beta} - \beta) \leqq (a, b)\} = P_{00}\{(\hat{\alpha}, \hat{\beta}) \leqq (a, b)\}.$$

In computing the distributions of the estimates, we may therefore assume that $\alpha = \beta = 0$.

One would like to have the distributions of $\hat{\alpha}$ and $\hat{\beta}$, centered in some sense, on the true parameter values. The next lemma gives conditions under which $\hat{\alpha}$ and $\hat{\beta}$ are symmetrically distributed about α and β.

LEMMA 4.2. *Let $T_1(Y)$ and $T_2(Y)$ be given by* (2.4) *and* (2.5), *with ψ non-decreasing, and let $\hat{\alpha}$ and $\hat{\beta}$ be the* (1.5)- *and* (1.3)-*estimates. If $F \varepsilon \mathfrak{F}$, then $\hat{\beta}$ is symmetrically distributed about β, and $\hat{\alpha}$ is symmetrically distributed about α, and hence $\hat{\alpha}$ and $\hat{\beta}$ are unbiased.*

PROOF. Similar to the proof of Theorem 2 of [9].

5. Limiting distributions. The study of the asymptotic distributions of $\hat{\alpha}$ and $\hat{\beta}$ is based on a result of Hodges and Lehmann (see Theorem 4 of [9]) which gives the connection between the distribution of the estimate and that of the test statistic, on which it is based. Using this result, it can be seen that, under the regularity conditions of Lemma 4.1 of [1],

$$(5.1) \quad \lim \mathcal{L}(n^{\frac{1}{2}}(\hat{\beta} - \beta) \mid P_\beta) = \lim \mathcal{L}(n^{\frac{1}{2}}\hat{\beta} \mid P_0) = N(0, k^2(\psi_0)/c^2)$$

where

$$(5.2) \qquad k^2(\psi_0) = k^2(\psi) = \left(\int_0^1 \psi^2(u)\, du\right)/\left(\int_0^1 \psi(u)\varphi(u)\, du\right)^2$$

with $c^2 = \lim n^{-1} \sum_j (x_j - \bar{x}_n)^2$, and $\varphi_0(u)$ is defined as in (2.3).

For the joint limiting distribution of $(\hat{\alpha}, \hat{\beta})$, it is convenient to consider first the asymptotic distribution of

$$\hat{\delta}_n = \hat{\alpha} + \hat{\beta}\bar{x}_n . \tag{5.3}$$

Observe that if we write the identity

$$Y_j = \alpha + \beta x_j + Z_j = \delta_n + \beta\xi_j + Z_j$$

where $\sum_j \xi_j = 0$, then δ_n depends on n only through a known quantity $\bar{x}_n$, and the estimate $\hat{\delta}_n$ of δ_n is still based on $T_1(Y - \hat{\beta}x)$. Furthermore, the invariance property (4.1) and (4.2) of the estimates, yield equation (5.3).

The following slight generalization of a result in [9], shall be used in the sequel.

THEOREM 5.1 (Hodges and Lehmann). *Let $\bar{x}_n \to \bar{x}$ as $n \to \infty$, and let*

$$\Delta_n = -n^{-\frac{1}{2}}(a + b) \tag{5.4}$$

where a and b are real constants. Let Φ be the distribution function of a normal random variable with mean zero and unit variance, and suppose

$$\lim P_n\{n^{\frac{1}{2}}T_1(Y) \leqq y\} = \Phi[(y + dB)/A] \tag{5.5}$$

where $d = (a + b\bar{x})$, and P_n denotes that the probability is computed for the sequence of parameter values Δ_n. Then for any sequence

$$\delta_n = \alpha + \beta\bar{x}_n \text{ that tends to } \alpha + \beta\bar{x} = \delta, \tag{5.6}$$

$$\lim P_{\delta_n}\{n^{\frac{1}{2}}(\hat{\delta}_n - \delta_n) \leqq d\} = \Phi(dB/A).$$

To establish the limiting distribution of $\hat{\delta}_n$, we need the asymptotic distribution under Δ_n of the statistic

$$n^{\frac{1}{2}}T_1(\hat{Z}) = n^{-\frac{1}{2}}\sum_j \psi_n(\hat{\mathfrak{R}}_j/n + 1) \text{ Sign } \hat{Z}_j \tag{5.7}$$

where $\hat{\mathfrak{R}}_j$ is the rank of $|\hat{Z}_j|$ in the sequence of absolute values $|\hat{Z}_1|, \cdots, |\hat{Z}_n|$, with $\hat{Z}_j = Y_j - \hat{\beta}x_j$.

This asymptotic distribution may be obtained by the help of the following theorem, the proof of which is given in the appendix.

THEOREM 5.2. *Let $\dot{\beta}$ be any estimate of β in* (1.1) *such that*

$$n^{\frac{1}{2}}(\dot{\beta} - \beta) \text{ is bounded in probability as } n \to \infty. \tag{5.8}$$

Assume that the regularity conditions of Lemma 4.1 *of* [1] *are satisfied. If*

$$|d/dy\psi(G(y))| \leqq K(\text{a constant}), \tag{5.9}$$

then

$$\lim \mathfrak{L}(n^{\frac{1}{2}}T_1(\dot{Z}) \mid P_n) = \lim \mathfrak{L}(n^{\frac{1}{2}}T_1(Y) \mid P_n) \tag{5.10}$$

where $\dot{Z}_j = Y_j - \dot{\beta}x_j$, P_n denotes the distribution under Δ_n of (5.4).

We remark that (5.8) is satisfied by all the (1.3)-estimates based on (2.5); it can also be easily verified that (5.9) is satisfied by the usual symmetric distributions such as the normal, the double exponential and the logistic. It follows from Theorem 4.1 of [1], that under the assumptions of Theorem 5.2, $T_1(\hat{Z})$ and $T_2(Y)$ defined in (2.4) and (2.5) are asymptotically independent with limiting distributions

$$\lim \mathcal{L}(n^{\frac{1}{2}}T_1(\hat{Z}) \mid P_n) = N(d \int \psi\varphi, \int \psi^2),$$
$$\lim \mathcal{L}(n^{\frac{1}{2}}T_2(Y) \mid P_n) = N(b \int \psi_0\varphi_0, c^2 \int \psi_0^2),$$

hence $\hat{\delta}_n$ and $\hat{\beta}$ are asymptotically independent with the limiting distribution of $\hat{\delta}_n$ given by

$$\lim \mathcal{L}(n^{\frac{1}{2}}(\hat{\delta}_n - \delta_n) \mid P_{\delta_n}) = \lim \mathcal{L}(n^{\frac{1}{2}}\hat{\delta}_n \mid P_0) = N(0, k^2(\psi)) \tag{5.11}$$

where $k^2(\psi)$ is defined in (5.2).

To compute the limiting distribution of $(\hat{\alpha}, \hat{\beta})$, we use the following simple lemma.

LEMMA 5.1. *Let* (X_n, Y_n) *be a sequence of random vectors and* $\{u_n\}, \{v_n\}$ *be two sequences of constants, such that* $u_n \to u$, $v_n \to v$ *as* $n \to \infty$. *If* $\lim \mathcal{L}(X_n, Y_n) = \mathcal{L}(X, Y)$, *then* $\lim \mathcal{L}(u_nX_n + v_nY_n, Y_n) = \mathcal{L}(uX + vY, Y)$.

With this lemma, it is straight forward to establish the following main result of this section.

THEOREM 5.3. *Let* $\hat{\alpha}$ *and* $\hat{\beta}$ *be the* (1.5)- *and* (1.3)-*estimates based on* (2.4) *and* (2.5) *respectively. Assume that the regularity conditions of Theorem* 5.2 *are satisfied. Then* $\mathcal{L}(n^{\frac{1}{2}}(\hat{\alpha} - \alpha, \hat{\beta} - \beta) \mid P_{\alpha\beta})$ *tends to the bivariate normal distribution with means* (0, 0) *and covariance matrix* $k^2(\psi)\Sigma$, *where*

$$\Sigma = \begin{pmatrix} (c^2 + \bar{x}^2)/c^2 & -\bar{x}/c^2 \\ -\bar{x}/c^2 & 1/c^2 \end{pmatrix} \tag{5.12}$$

with $c^2 = \lim n^{-1} \sum_j (x_j - \bar{x}_n)^2$, *and* $\bar{x} = \lim \bar{x}_n$.

6. Asymptotic efficiency. In this section, we determine the asymptotic efficiency of our estimates relative to the classical least squares estimates. In doing this, we make use of the fact that if two vectors $\mathbf{U}_1$ and $\mathbf{U}_2$ have limiting normal distributions with covariance matrices Σ_1 and Σ_2 related by $\Sigma_1 = k^2\Sigma_2$, for some constant k, then the asymptotic efficiency of $\mathbf{U}_2$ relative to $\mathbf{U}_1$ is k^2.

Conditions for the asymptotic normality of a general class of least squares estimates have been given by Eicker [5]. It can easily be checked that under the assumptions of Theorem 5.2, the conditions in [5] are satisfied. If $\tilde{\alpha}$ and $\tilde{\beta}$ denote the least squares estimates then under very general conditions, $\mathcal{L}(n^{\frac{1}{2}}(\tilde{\alpha} - \alpha, \tilde{\beta} - \beta) \mid P_{\alpha\beta})$ tends to the bivariate normal distribution with zero mean and covariance matrix Σ given in (5.12). It follows that the asymptotic efficiency of the estimates $(\hat{\alpha}, \hat{\beta}) = \hat{\mathbf{\Delta}}$ relative to the least squares estimates $(\tilde{\alpha}, \tilde{\beta}) = \tilde{\mathbf{\Delta}}$ is $k^{-2}(\psi)$. If the common variance of Y_j is σ^2 instead of unity as assumed, the efficiency becomes

$$e_{\hat{\mathbf{\Delta}},\tilde{\mathbf{\Delta}}}(\psi) = \sigma^2\left(\int_0^1 \psi(u)\varphi(u)\, du\right)^2 / \left(\int_0^1 \psi^2(u)\, du\right). \tag{6.1}$$

As expected, (6.1) is the same as the Pitman efficiency [1] of the M_n-tests relative to the classical F-test. This is unlike the situation in the multivariate location case where the corresponding efficiencies do not coincide, see Bickel [2] and [3]. From the particular cases of (6.1) discussed in [1], it follows that the estimates $\hat{\alpha}$ and $\hat{\beta}$ have all the desirable properties including robustness [10], of the Hodges and Lehmann estimates for shift.

7. Comparison with the Brown and Mood Estimates. In [11], Brown and Mood proposed the 'median' estimates $\dot{\alpha}$ and $\dot{\beta}$ of α and β. These estimates are determined by the two equations:

$$\text{Median } (Y_j - \dot{\alpha} - \dot{\beta} x_j) = 0 \quad \text{for} \quad x_j \leqq \text{med } x, \tag{7.1}$$

$$\text{Median } (Y_i - \dot{\alpha} - \dot{\beta} x_i) = 0 \quad \text{for} \quad x_i > \text{med } x, \tag{7.2}$$

where med x is the median of the constants x_j, $j = 1, \cdots, n$.

In [8] Hill proved both the existence and the asymptotic normality of $(\dot{\alpha}, \dot{\beta}) = \dot{\mathbf{\Delta}}$. He showed that $\mathcal{L}((\frac{1}{2}n)^{\frac{1}{2}}(\dot{\alpha} - \alpha, \dot{\beta} - \beta))$ tends to the bivariate normal distribution with mean (0, 0) and covariance matrix $\|\tau_{\alpha\beta}\|$ defined by

$$\tau_\alpha^2 = [(\textstyle\int_0^{\frac{1}{2}} h(t)\,dt)^2 + (\int_{\frac{1}{2}}^1 h(t)\,dt)^2](2f^2(\theta)\{\int_{\frac{1}{2}}^1 h(t)\,dt - \int_0^{\frac{1}{2}} h(t)\,dt\}^2)^{-1}, \tag{7.3}$$

$$\tau_\beta^2 = \{8[f(0)\{\textstyle\int_{\frac{1}{2}}^1 h(t)\,dt - \int_0^{\frac{1}{2}} h(t)\,dt\}]^2\}^{-1}, \tag{7.4}$$

$$\tau_{\alpha\beta} = -\textstyle\int_0^1 h(t)\,dt\{8[f(0)\{\int_{\frac{1}{2}}^1 h(t)\,dt - \int_0^{\frac{1}{2}} h(t)\,dt\}]^2\}^{-1}, \tag{7.5}$$

where h is a continuous strictly monotone increasing function on [0, 1] (also called spacing function) defined by

$$x_{nj} = h(j/n), \qquad j = 0, \cdots, n. \tag{7.6}$$

Since the asymptotic covariance matrix of $\dot{\mathbf{\Delta}}$ is not proportional to that of $\hat{\mathbf{\Delta}}$, in order to compute their efficiencies one may use another measure of efficiency based on asymptotic generalized variance (Cramér [4], p. 301).

The asymptotic generalized variance of $n^{\frac{1}{2}}\dot{\mathbf{\Delta}}$ is

$$\text{Var } n^{\frac{1}{2}}\dot{\mathbf{\Delta}} = [4f^2(0)\{\textstyle\int_{\frac{1}{2}}^1 h(t)\,dt - \int_0^{\frac{1}{2}} h(t)\,dt\}]^{-2}, \tag{7.7}$$

and that of the (1.3)- and (1.5)- estimates $\hat{\mathbf{\Delta}}$ is

$$\text{Var } n^{\frac{1}{2}}\hat{\mathbf{\Delta}} = [\textstyle\int_0^1 \psi^2(u)\,du]^2/[c^2\{\int_0^1 \psi(u)\varphi(u)\,du\}]. \tag{7.8}$$

It therefore follows that the asymptotic efficiency of the median estimates $\dot{\mathbf{\Delta}}$ relative to the $\hat{\mathbf{\Delta}}$-estimates is given by

$$e_{\dot{\mathbf{\Delta}},\hat{\mathbf{\Delta}}} = [f^2(0)\{\textstyle\int_{\frac{1}{2}}^1 h(t)\,dt - \int_0^{\frac{1}{2}} h(t)\,dt\} \int_0^1 \psi^2(u)\,du] \cdot\{[\int_0^1 \psi(u)\varphi(u)\,du]^2[\int_0^1 h^2(t)\,dt - (\int_0^1 h(t)\,dt)^2]^{\frac{1}{2}}\}^{-1} \tag{7.9}$$

where we have written c in terms of the function h, by the relation $\lim n^{-1}\sum_j x_j = \int_0^1 h(t)\,dt$. If we consider in particular the estimates $(\hat{\alpha}_s, \hat{\beta}_s) = \hat{\mathbf{\Delta}}_s$, based on the sign statistics $T_1(Y) = n^{-1}\sum_j \text{Sign } Y_j$ and $T_2(Y) = n^{-1}\sum_j(x_j - \bar{x}) \text{ Sign } Y_j$, the efficiency expression in (7.9) reduces to

$$e_{\dot{\mathbf{\Delta}},\hat{\mathbf{\Delta}}_s} = [\textstyle\int_{\frac{1}{2}}^1 h(t)\,dt - \int_0^{\frac{1}{2}} h(t)\,dt][\int_0^1 h^2(t)\,dt - (\int_0^1 h(t)\,dt)^2]^{-\frac{1}{2}}. \tag{7.10}$$

The function h is typically linear with positive slope, and with this, (7.10) simplifies to

$$e_{\hat{\Delta}^{\bullet}, \hat{\Delta}^{s}} = 3^{\frac{1}{2}}/2 < 1 \tag{7.11}$$

which implies that there is some loss in the efficiency of the median estimates. This loss is probably due to the fact that some information is lost in the process of ordering the observations in two separate groups.

APPENDIX

8. Proof of Theorem 5.2. Throughout this section, k will denote a generic constant and θ_j (short for $\theta_{nj}(t)$) a generic sequence of functions that tends to zero uniformly in t for $|t| \leqq k$, as $n \to \infty$.

Let $Y_j^* = Y_j - n^{\frac{1}{2}}(tx_j)$ for $|t| \leqq k$ and let E_0 denote the expectation taken with respect to the distribution under $\alpha = \beta = 0$. Write

$$\begin{aligned} T_1(Y) &= n^{-1} \textstyle\sum_j \psi_n(\mathcal{R}_j/n + 1) \operatorname{Sign} Y_j , \\ T_1(Y^*) &= n^{-1} \textstyle\sum_j \psi_n(\mathcal{R}_j^*/n + 1) \operatorname{Sign} Y_j^*, \end{aligned} \tag{8.1}$$

where $\mathcal{R}_j^*$ is the rank of $|Y_j^*|$ in the sequence of absolute values $|Y_1^*|, \cdots, |Y_n^*|$.

Due to the independence of $(\mathcal{R}_1, \cdots, \mathcal{R}_n)$ and $(\operatorname{Sign} Y_1, \cdots, \operatorname{Sign} Y_n)$.

$$\begin{aligned} \operatorname{Var}_0 [n^{\frac{1}{2}}(T_1 - T_1^*)] & \\ &= n\{E_0(T_1 - T_1^*)^2 - [E_0(T_1 - T_1^*)]^2\} \\ &= n^{-1} \textstyle\sum_j E_0[\psi_n(\mathcal{R}_j^*/n + 1) \operatorname{Sign} Y_j^* - \psi_n(\mathcal{R}_j/n + 1) \operatorname{Sign} Y_j]^2 \\ &\qquad - n^{-1}[\textstyle\sum_j E_0\{\psi_n(\mathcal{R}_j^*/n + 1) \operatorname{Sign} Y_j^*\}]^2 \\ &\leqq n^{-1} \textstyle\sum_j E_0[\psi_n(\mathcal{R}_j^*/n + 1) \operatorname{Sign} Y_j^* - \psi_n(\mathcal{R}_j/n + 1) \operatorname{Sign} Y_j]^2 \\ &= B_1 + B_2 + B_3 , \end{aligned}$$

with

$$\begin{aligned} B_1 &= n^{-1} \textstyle\sum_j E_0[\psi_n(\mathcal{R}_j^*/n + 1) - \psi_n(\mathcal{R}_j/n + 1)]^2, \\ B_2 &= n^{-1} \textstyle\sum_j E_0[\psi_n(\mathcal{R}_j^*/n + 1)(\operatorname{Sign} Y_j^* - \operatorname{Sign} Y_j)]^2, \\ B_3 &= 2n^{-1} \textstyle\sum_j E_0[\{\psi_n(\mathcal{R}_j^*/n + 1) - \psi_n(\mathcal{R}_j/n + 1)\}\{\operatorname{Sign} Y_j^*\} \\ &\qquad \cdot\{\psi_n(\mathcal{R}_j^*/n + 1)\}\{\operatorname{Sign} Y_j^* - \operatorname{Sign} Y_j\}]. \end{aligned}$$

Now

$$\begin{aligned} |B_2| &= 4\, |n^{-1} \textstyle\sum_j \{(tx_j)/n^{\frac{1}{2}}\} f(\theta_j) E_0\psi_n^2(\mathcal{R}_j^*/n + 1)| \\ &\leqq \max_{1 \leqq j \leqq n} k\, |n^{-\frac{1}{2}} x_j f(\theta_j)|\; n^{-1} \textstyle\sum_j \psi_n^2(j/n + 1). \end{aligned}$$

Using the absolute continuity of f, and the boundedness of the regression constants x_j (see (1.2), (1.3) and (1.4) of [1]), we see that $\lim_n \sup_t |B_2| = 0$. For B_3, we have

$$|B_3| = |2n^{-1} \textstyle\sum_j E_0[\psi_n^2(\mathfrak{R}_j^*/n+1) - \psi_n(\mathfrak{R}_j/n+1)\psi_n(\mathfrak{R}_j^*/n+1)] \cdot [1 - \text{Sign } Y_j \text{ Sign } Y_j^*]|$$

$$\leqq \max_{1\leqq j\leqq n} k\,|n^{-\frac{1}{2}}x_j f(\theta_j)|\, n^{-1} \textstyle\sum_j |E_0[\psi_n^2(\mathfrak{R}_j^*/n+1) - \psi_n(\mathfrak{R}_j/n+1)\psi_n(\mathfrak{R}_j^*/n+1)]|$$

$$\leqq \max_{1\leqq j\leqq n} k\,|n^{-\frac{1}{2}}x_j f(\theta_j)|\, n^{-1} \textstyle\sum_j \psi^2(j/n+1).$$

Hence $\lim_n \sup_t |B_3| = 0$ for $|t| \leqq k$. Write $B_1 = \sum_{i=1}^6 B_{1i}$, where

$$B_{11} = n^{-1} \textstyle\sum_j E_0[\psi_n(\mathfrak{R}_j/n+1) - \psi(U_j)]^2,$$

$$B_{12} = n^{-1} \textstyle\sum_j E_0[\psi_n(\mathfrak{R}_j^*/n+1) - \psi(U_j^*)]^2,$$

$$B_{13} = n^{-1} \textstyle\sum_j E_0[\psi(U_j) - \psi(U_j^*)]^2,$$

$$B_{14} = 2n^{-1} \textstyle\sum_j E_0[\psi_n(\mathfrak{R}_j/n+1) - \psi(U_j)][\psi(U_j^*) - \psi_n(\mathfrak{R}_j^*/n+1)],$$

$$B_{15} = 2n^{-1} \textstyle\sum_j E_0[\psi_n(\mathfrak{R}_j/n+1) - \psi(U_j)][\psi(U_j) - \psi(U_j^*)],$$

$$B_{16} = 2n^{-1} \textstyle\sum_j E_0[\psi(U_j^*) - \psi_n(\mathfrak{R}_j^*/n+1)][\psi(U_j) - \psi(U_j^*)],$$

where $U_j = 2F(Y_j) - 1$, $U_j^* = 2F(Y_j^*) - 1$.

By (2.4) of [1], it is immediate that both B_{11} and B_{12} tend to zero uniformly in t for $|t| \leqq k$. On applying the mean value theorem to B_{13}, we obtain

$$|B_{13}| \leqq (k/n^2) \textstyle\sum_j \{x_j f(y+\theta_j)\psi'[2F(y) - 1 + \theta_j]\}^2$$

$$\leqq (k/n^2) \textstyle\sum x_j^2 \qquad \text{by (5.9)},$$

$$|B_{14}| \leqq |B_{11}| + |B_{12}| \to 0 \qquad \text{uniformly in } t,$$

$$|B_{15}| \leqq |B_{11}| + |B_{13}| \to 0 \qquad \text{uniformly in } t, \text{ finally}$$

$$|B_{16}| \leqq |B_{12}| + |B_{13}|.$$

We have therefore proved that

$$\sup_{|t|\leqq k} [n^{\frac{1}{2}}\{T_1(Y) - T_1(Y^*)\}] \to 0 \qquad \text{in } P_0\text{-probability},$$

and it follows from the contiguity of P_0 and P_n that

$$\sup_{|t|\leqq k} -n^{\frac{1}{2}}[T_1(Y) - T_1(Y^*)]^2 \to 0 \qquad \text{in } P_n\text{-probability},$$

and the theorem is proved.

9. Acknowledgment. I should like to thank Professor Erich L. Lehmann, not only for suggesting this problem but also for his continued guidance and encouragement. I am also grateful to Professor J. Hájek for many helpful discussions, and to a referee for useful suggestions.

REFERENCES

[1] ADICHIE, J. N. (1967). Asymptotic efficiency of a class of non-parametric tests for regression parameters. *Ann. Math. Statist.* **38** 884–893.

[2] BICKEL, P. J. (1964). On some alternative estimates for shift in the p-variate one same problem. *Ann. Math. Statist.* **35** 1079–1090.

[3] BICKEL, P. J. (1965). On some asymptotically non-parametric competitors of Hotelling's T^2. *Ann. Math. Statist.* **36** 160–173.

[4] CRAMÉR, H. (1945). *Mathematical Methods of Statistics*. Princeton Univ. Press.

[5] EICKER, F. (1963). Asymptotic normality and consistency of the least squares estimators for families of linear regressions. *Ann. Math. Statist.* **34** 447–456.

[6] HÁJEK, J. (1962). Asymptotically most powerful rank order tests. *Ann. Math. Statist.* **33** 1124–1147.

[7] HEMELRIJK, J. (1949). Construction of a confidence region for a line. *Indag. Mth.* **11** 374–384.

[8] HILL, B. M. (1962). A test of linearity versus convexity of a median regression curve. *Ann. Math. Statist.* **33** 1096–1123.

[9] HODGES, J. L., JR. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

[10] HOYLAND, A. (1965). Robustness of Hodges and Lehmann estimates for shift. *Ann. Math. Statist.* **36** 174–197.

[11] MOOD, A. M. (1950). *Introduction to the Theory of Statistics*. McGraw-Hill, New York.

[12] NAIR, K. R. and SHRIVASTAVA, M. D. (1942). On a simple method of curve fitting. *Sankhya* **6** 121–132.

[13] THEIL, H. (1950). A rank invariant method of linear and polynomial regression analysis. *Indag. Math.* **12** *Fasc.* **2** 85–91, 173–177.

[14] WALD, A. (1940). The fitting of straight lines if both variables are subject to error. *Ann. Math. Statist.* **11** 284–300.