UNIVERSITY OF NORTH CAROLINA
Department of Statistics
Chapel Hill, N. C.

A TEST OF WHETHER TWO PARALLEL REGRESSION LINES ARE THE SAME WHEN THE VARIANCES MAY BE UNEQUAL

by

Richard F. Potthoff

August, 1962

Contract No. AF 49(638)-213

This paper considers a situation where one has two regression lines which are known to be parallel, with the two sets of error terms assumed to be normally distributed but with (possibly) different variances. The paper presents a test of the hypothesis that the two regression lines are identical (i.e., the two α-coefficients are equal). The test is analogous to the Wilcoxon test. The discussion in this paper is on a rather technical level; for a less technical discussion of the test, see Mimeo Series No. 323.

This research was supported in part by Educational Testing Service, and was supported in part by the Air Force Office of Scientific Research.

Institute of Statistics
Mimeograph Series No. 334
August, 1962

A TEST OF WHETHER TWO PARALLEL REGRESSION LINES ARE THE SAME WHEN THE VARIANCES MAY BE UNEQUAL[1]

by

Richard F. Potthoff
University of North Carolina

1. Introduction and summary. We suppose that we have M pairs (Y_1, X_1), $(Y_2, X_2), \ldots, (Y_M, X_M)$ such that

$$(1.1) \qquad Y_i = \alpha_Y + \beta X_i + e_i \qquad (i = 1,2,\ldots, M) \quad ,$$

and N pairs $(Z_1, W_1), (Z_2, W_2), \ldots, (Z_N, W_N)$ such that

$$(1.2) \qquad Z_j = \alpha_Z + \beta W_j + f_j \qquad (j = 1,2,\ldots, N) \quad .$$

In other words, it is known that the two regression lines associated respectively with (1.1) and (1.2) are parallel. The regression parameters α_Y, α_Z, and β are unknown. We assume that the X_i's and W_j's are (known) fixed constants; that the e_i's are $N(0, \sigma_e^2)$ and the f_j's are $N(0, \sigma_f^2)$, with $e_1, e_2, \ldots, e_M, f_1, f_2, \ldots, f_N$ being mutually independent; and that the variances σ_e^2 and σ_f^2 are unknown. We may define $\tau = \sigma_e^2/(\sigma_e^2 + \sigma_f^2)$.

It is desired to test the hypothesis

$$(1.3) \qquad H_o: \alpha_Y = \alpha_Z$$

against alternatives $\alpha_Y \neq \alpha_Z$. In other words, H_o (1.3) is the hypothesis that our two (parallel) regression lines are identical. The test of H_o to be presented here bears some resemblance to the Wilcoxon test [2, 6] and is based on

[1] This research was supported in part by Educational Testing Service, and was supported in part by the Air Force Office of Scientific Research.

the same kind of idea that was used in [3] and [4] to develop Wilcoxon-type tests for a generalized Behrens-Fisher problem [3] and for testing whether two regression lines are parallel when the two variances are (possibly) unequal [4]. Thus, even though we are assuming normality (of the e_i's and f_j's) for our test in this paper, our approach will be that of non-parametric statistics. The behavior of our test under non-normality of the e_i's and f_j's is an area which has not been explored.

For our test of H_o (1.3), we will utilize the statistic

$$w = (1/T) \sum_{\Omega} u(V_{iIjJ}) \quad , \tag{1.4}$$

where

$$V_{iIjJ} = \frac{(X_I - W_j)(Z_J - Y_i) - (W_J - X_i)(Y_I - Z_j)}{X_I + W_J - X_i - W_j} \quad , \tag{1.5}$$

the function $u(V)$ is 0 or 1 according to whether $V \leq 0$ or $V > 0$ respectively, the set Ω (over which the summation is taken) embraces all quadruples (i, I, j, J) such that

$$X_i < X_I \quad , \qquad W_j < W_J \tag{1.6a}$$

and

$$X_i \leq W_J \; , \qquad W_j \leq X_I \quad , \tag{1.6b}$$

and T is the total number of quadruples (i, I, j, J) satisfying (1.6) (i.e., belonging to Ω). In Section 2 we will show that, regardless of what τ is,

$$E(w) = \frac{1}{2} \qquad \text{if and only if } H_o \text{ is true} . \tag{1.7}$$

Section 3 establishes that

(1.8) $$\sup_{0<\tau<1} \operatorname{var}(w) = Q \qquad \text{if } H_0 \text{ is true ,}$$

where the number Q is an involved function of the X_i's and W_j's and is defined by (3.5 - 3.7), (3.9). In Section 4 we utilize a theorem of Hoeffding [1] to show that, under certain mild restrictions, $[\operatorname{var}(w)]^{-1/2}(w - \frac{1}{2})$ is asymptotically $N(0, 1)$ if H_0 (1.3) is true. This result together with (1.7) and (1.8) tells us that a test with critical region

(1.9) $$Q^{-1/2} \left| w - \frac{1}{2} \right| > z_{\alpha/2} \qquad ,$$

where $z_{\alpha/2}$ is defined by

$$(2\pi)^{-1/2} \int_{-\infty}^{z_{\alpha/2}} e^{-z^2/2}\, dz = 1 - (\alpha/2) \qquad ,$$

will be a size-α test of H_0 (disregarding inaccuracies due to the normal approximation); the test (1.9) will of course be a conservative test, since var(w) generally will not be as large as Q. Section 5 is concerned with the consistency of the test (1.9).

2. The expectation of w. We now establish (1.7). From (1.1-1.2) and (1.5) we have

(2.1) $$V_{iIjJ} = p_{iIjJ}(f_J - e_i) + (1 - p_{iIjJ})(f_j - e_I) + (\alpha_Z - \alpha_Y) \quad ,$$

where

$$p_{iIjJ} = \frac{X_I - W_j}{X_I + W_J - X_i - W_j} \quad .$$

Since the e's and f's are normal and independent, and all have zero means, it follows that V_{iIjJ} is normally distributed with mean $(\alpha_Z - \alpha_Y)$. Hence V_{iIjJ} has median $(\alpha_Z - \alpha_Y)$, so that

$$(2.2) \qquad E\left[u(V_{iIjJ})\right] \begin{cases} = \frac{1}{2} & \text{if } \alpha_Z = \alpha_Y \\ > \frac{1}{2} & \text{if } \alpha_Z > \alpha_Y \\ < \frac{1}{2} & \text{if } \alpha_Z < \alpha_Y \end{cases}$$

for all (i, I, j, J). If we take the expectation of both sides of (1.4) and apply (2.2), we end up with (1.7).

Observe that

$$(2.3) \qquad \operatorname{var}(V_{iIjJ}) \leq \sigma_e^{\,2} + \sigma_f^{\,2}$$

for all (i, I, j, J) in Γ. If $\alpha_Z > \alpha_Y$, it follows from (2.3) that

$$E\left[u(V_{iIjJ})\right] \geq \frac{1}{2} + d$$

for all (i, I, j, J) in Γ , where

$$(2.4) \qquad d = (2\pi)^{-\frac{1}{2}} \int_0^{(\alpha_Z - \alpha_Y)/(\sigma_e^{\,2} + \sigma_f^{\,2})^{\frac{1}{2}}} e^{-z^2/2}\, dz \quad .$$

Hence

$$(2.5) \qquad E(w) \geq \frac{1}{2} + d \qquad \text{if } \alpha_Z > \alpha_Y \quad .$$

An analogous relation holds when $\alpha_Z < \alpha_Y$.

We may remark in passing that the relation (1.7) holds even for certain non-normal distributions of the e_i's and f_j's. Let us use $F_Y(e)$ and $F_Z(f)$ to denote the cumulative distribution functions of the e's and f's respectively. Suppose that $F_Y(e)$ and $F_Z(f)$ are both symmetric (about 0) and continuous. If two independent random variables are each distributed symmetrically about 0, then their sum or difference is also distributed symmetrically about 0. Hence (f - e) is symmetrically distributed about 0, and V_{iIjJ} is symmetrically distributed about $(\alpha_Z - \alpha_Y)$, so that V_{iIjJ} has median $(\alpha_Z - \alpha_Y)$. We thus conclude that (1.7) holds whenever $F_Y(e)$ and $F_Z(f)$ are both symmetric about 0 and continuous [with certain trivial exceptions which could occur if there is no density for V_{iIjJ} in the interval between 0 and $(\alpha_Z - \alpha_Y)$].

3. The variance of w. Next we establish the least upper bound for var(w) when H_o is true. In the previous section we found that we could relax the normality assumption somewhat and still be able to prove (1.7); in this section, however, we give a proof of (1.8) which is valid only for the case of normally distributed e_i's and f_j's.

It is easily shown {refer to [4, equations (4.2-4.3)]} that, if V and V' follow a bivariate normal distribution with correlation coefficient $\rho(V, V')$ and with $E(V) = E(V') = 0$, then

(3.1) $$\text{cov}\,[u(V),\, u(V')] = (1/2\pi)\,\sin^{-1}\rho(V,V') \quad .$$

Now when H_o is true, (2.1) becomes

(3.2) $$V_{iIjJ} = p_{iIjJ}(f_J - e_i) + q_{iIjJ}(f_j - e_I) \quad ,$$

where $q_{iIjJ} = 1 - p_{iIjJ}$. Thus V_{iIjJ} is normal with mean 0 and variance $(p^2_{iIjJ} + q^2_{iIjJ})(\sigma_e^{\,2} + \sigma_f^{\,2})$. From (1.4) and (3.1) we obtain

(3.3) $$\text{var}(w) = v = v(\tau)$$

$$= (1/T^2)(1/2\pi) \sum_{\Omega} \sum_{\Omega} \sin^{-1}\rho(V_{iIjJ}, V_{i'I'j'J'}) \quad ,$$

where the first summation is over (i, I, j, J) and the second summation is over (i', I', j', J'). We have

$$(3.4)\qquad \rho(V_{iIjJ}, V_{i'I'j'J'}) = \frac{\mathrm{cov}(V_{iIjJ}, V_{i'I'j'J'})}{[\mathrm{var}(V_{iIjJ})]^{\frac{1}{2}}\,[\mathrm{var}(V_{i'I'j'J'})]^{\frac{1}{2}}}$$

$$= \rho^{(1)}_{iIjJ,i'I'j'J'}\tau + \rho^{(2)}_{iIjJ,i'I'j'J'}(1-\tau) \qquad ,$$

where

$$(3.5)\quad \rho^{(1)}_{iIjJ,i'I'j'J'} = (\delta_{ii'}p_{iIjJ}p_{i'I'j'J'} + \delta_{iI'}p_{iIjJ}q_{i'I'j'J'} + \delta_{Ii'}q_{iIjJ}p_{i'I'j'J'} + \delta_{II'}q_{iIjJ}q_{i'I'j'J'})$$

$$(p^2_{iIjJ} + q^2_{iIjJ})^{-\frac{1}{2}}\,(p^2_{i'I'j'J'} + q^2_{i'I'j'J'})^{-\frac{1}{2}}$$

$$= [\delta_{ii'}(X_I - W_j)(X_{I'} - W_{j'}) + \delta_{iI'}(X_I - W_j)(W_{J'} - X_{i'}) + \delta_{Ii'}(W_J - X_i)(X_{I'} - W_{j'}) + \delta_{II'}(W_J - X_i)(W_{J'} - X_{i'})]$$

$$[(X_I - W_j)^2 + (W_J - X_i)^2]^{-\frac{1}{2}}\,[(X_{I'} - W_{j'})^2 + (W_{J'} - X_{i'})]^{-\frac{1}{2}}$$

and

$$(3.6)\quad \rho^{(2)}_{iIjJ,i'I'j'J'} = (\delta_{jj'}q_{iIjJ}q_{i'I'j'J'} + \delta_{jJ'}q_{iIjJ}p_{i'I'j'J'} + \delta_{Jj'}p_{iIjJ}q_{i'I'j'J'} + \delta_{JJ'}p_{iIjJ}p_{i'I'j'J'})(p^2_{iIjJ} + q^2_{iIjJ})^{-\frac{1}{2}}(p^2_{i'I'j'J'} + q^2_{i'I'j'J'})^{-\frac{1}{2}}$$

$$= [\delta_{jj'}(W_J - X_i)(W_{J'} - X_{i'}) + \delta_{jJ'}(W_J - X_i)(X_{I'} - W_{j'}) + \delta_{Jj'}(X_I - W_j)(W_{J'} - X_{i'}) + \delta_{JJ'}(X_I - W_j)(X_{I'} - W_{j'})]$$

$$[(X_I - W_j)^2 + (W_J - X_i)^2]^{-\frac{1}{2}}[(X_{I'} - W_{j'})^2 + (W_{J'} - X_{i'})^2]^{-\frac{1}{2}} .$$

The δ's in (3.5) and (3.6) are Kronecker deltas.

Let us define

$$(3.7) \qquad Q_k = (1/2\pi T^2) \sum_{\Omega} \sum_{\Omega} \sin^{-1} \rho^{(k)}_{iIjJ,i'I'j'J'}$$

for k = 1, 2. Then

$$(3.8a) \qquad \lim_{\tau \to 1} v(\tau) = Q_1$$

and

$$(3.8b) \qquad \lim_{\tau \to 0} v(\tau) = Q_2 \quad .$$

The bound Q in (1.8) we define by

$$(3.9) \qquad Q = \max(Q_1, Q_2) \quad .$$

If we can demonstrate that

$$(3.10) \qquad v(\tau) \leq Q \quad , \qquad 0 < \tau < 1 \quad ,$$

then (3.10) together with (3.8) will be sufficient to establish (1.8).

To prove (3.10), observe first that, if we use the simplified notation

$$(3.11) \qquad \rho = \rho(\tau) = \rho^{(1)}\tau + \rho^{(2)}(1 - \tau) \quad ,$$

then (for $\rho^2 \neq 1$) we have

$$(3.12) \qquad \frac{d^2 \sin^{-1} \rho}{d\tau^2} = [\rho^{(1)} - \rho^{(2)}]^2 (1 - \rho^2)^{-3/2} \rho \quad .$$

We now show that (3.12) is ≥ 0. Because of (1.6b), we know that p_{iIjJ}, q_{iIjJ}, $p_{i'I'j'J'}$, $q_{i'I'j'J'}$ are all ≥ 0 (for all quadruples in Ω). Therefore $\rho^{(1)}$

(3.5) and $\rho^{(2)}$ (3.6) must each be ≥ 0, from which it follows that ρ (3.11) is ≥ 0 (for all τ, $0 < \tau < 1$). Since $\rho \geq 0$, the right-hand side of (3.12) is ≥ 0, so that

$$(3.13) \qquad \frac{d^2 \sin^{-1} \rho(V_{iIjJ}, V_{i'I'j'J'})}{d\tau^2} \geq 0$$

for all τ $(0 < \tau < 1)$ and for all $(i,I,j,J),(i',I',j',J')$ in Γ .

Taking the second derivatives of both sides of (3.3) with respect to τ and then applying (3.13), we conclude that

$$(3.14) \qquad v''(\tau) \geq 0 \qquad (0 < \tau < 1) \quad .$$

Finally, (3.14) taken together with (3.8) establishes (3.10).

The proof of the bound (1.8) is thus complete.

4. Asymptotic normality. This section establishes that w (1.4) is asymptotically normal under H_o (1.3), provided that certain mild restrictions are obeyed. To prove this result, we will appeal to a theorem of Hoeffding [1, Theorem 8.1] concerning the asymptotic normality of U-statistics for random variables independently but not necessarily identically distributed.

A U-statistic must be of the form [1, equation (5.1)] ,

$$(4.1) \qquad U = \binom{n}{m}^{-1} \Sigma' \Phi(x_{\alpha_1}, x_{\alpha_2}, \ldots, x_{\alpha_m}) \quad ,$$

where the summation Σ' is over all $\binom{n}{m}$ sets satisfying $1 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_m \leq n$, the function Φ is symmetric in its m (vector) arguments, and the x_α's (α = 1, 2,..., n) are mutually independent (but not necessarily identically distributed) r-dimensional random variables of the form $x_\alpha = (x_\alpha^{(1)}, x_\alpha^{(2)}, \ldots, x_\alpha^{(r)})$. In (4.1), suppose we set $n = M + N$, $m = 4$, $r = 3$,

$$x_\alpha = (x_\alpha^{(1)}, x_\alpha^{(2)}, x_\alpha^{(3)}) = (e_\alpha, X_\alpha, 1) \quad \text{for } 1 \leq \alpha \leq M \tag{4.2}$$

$$= (f_{\alpha-M}, W_{\alpha-M}, 2) \quad \text{for } M + 1 \leq \alpha \leq M + N \quad ,$$

and

$$\Phi(x_{\alpha_1}, x_{\alpha_2}, x_{\alpha_3}, x_{\alpha_4}) = \Sigma\, \zeta(x_h, x_H, x_{h'}, x_{H'}) \tag{4.3}$$

$$u\left(\frac{\left[x_H^{(2)} - x_{h'}^{(2)}\right]\left[x_{H'}^{(1)} - x_h^{(1)}\right] - \left[x_{H'}^{(2)} - x_h^{(2)}\right]\left[x_H^{(1)} - x_{h'}^{(1)}\right]}{x_H^{(2)} + x_{H'}^{(2)} - x_h^{(2)} - x_{h'}^{(2)}}\right) ,$$

where the summation in (4.3) is over all $4! = 24$ permutations (h, H, h', H') of $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ (thereby ensuring that Φ is symmetric), and where

$$\zeta(x_h, x_H, x_{h'}, x_{H'}) = 1 \text{ if } x_h^{(3)} = x_H^{(3)} = 1,\ x_{h'}^{(3)} = x_{H'}^{(3)} = 2,\ x_H^{(2)} > x_h^{(2)},\ x_{H'}^{(2)} > x_{h'}^{(2)},$$

$$x_{H'}^{(2)} \geq x_h^{(2)},\ x_H^{(2)} \geq x_{h'}^{(2)}$$

$$= 0 \quad \text{otherwise} \quad .$$

Although the $x_\alpha^{(2)}$'s and $x_\alpha^{(3)}$'s are fixed constants, we consider them for our present purposes to be random variables for which all the density is concentrated at a single point; thus the x_α's are not identically distributed. If H_o (1.3) is true, then the statistic U determined by (4.1), (4.2-4.3) will be related to w (1.4) by the equation

$$U = k_{M,N}\, k^o_{M,N}\, w \tag{4.4}$$

where

(4.5) $$k_{M,N} = \binom{M}{2}\binom{N}{2}\binom{M+N}{4}^{-1}$$

and

(4.6) $$k^{o}_{M,N} = T\binom{M}{2}^{-1}\binom{N}{2}^{-1} .$$

As we already indicated, we will introduce some mild assumptions in order to prove that, under H_o, $[\mathrm{var}(w)]^{-1/2}(w - \frac{1}{2})$ is asymptotically $N(0,1)$.

Assumption 4A. We assume that $n \longrightarrow \infty$ in such a way that N/M approaches some constant $c > 0$.

Assumption 4B. We assume that there exists a fraction $P_1 > 0$ such that, for all (M, N), the number of quadruples (i,I,j,J) satisfying (1.6) is $\geq P_1\binom{M}{2}\binom{N}{2}$; i.e., the proportion of potential quadruples belonging to Γ is $\geq P_1$ for all n.

Assumption 4C. We assume that there exists a fraction P_2, $0 < P_2 < 1$, and a number ϵ, $0 < \epsilon < 1$, such that, for all n, the fraction of the T quadruples (i, I, j, J) belonging to Γ for which the relations

(4.7a) $$\epsilon \leq |X_I - W_j| \leq 1/\epsilon$$

and

(4.7b) $$\epsilon \leq |W_J - X_i| \leq 1/\epsilon$$

hold is $\geq P_2$; i.e., (4.7) is satisfied by at least P_2 T quadruples in Γ .

It is not claimed that these three assumptions are necessarily indispensable to an asymptotic normality proof. However, none of the three appears to be particularly restrictive.

From (4.5) and Assumption 4A we obtain

(4.8) $$\lim_{n \to \infty} k_{M,N} = 6c^2/(1 + c)^4 \quad .$$

From (4.6) and Assumption 4B we have

(4.9) $$k^o_{M,N} \geq P_1 \qquad \text{for all } n \quad .$$

Looking at (4.4), we see that, if we can prove that $[\mathrm{var}(U)]^{-1/2}\ [U - E(U)]$ is asymptotically N(0, 1), then it will follow immediately that $[\mathrm{var}(w)]^{-1/2}$ $(w - \frac{1}{2})$ is likewise asymptotically N(0, 1). Hoeffding's Theorem 8.1 [1], which is concerned with non-identically distributed x_α's, presents a set of conditions [1, (8.2 - 8.4)] under which $[\mathrm{var}(U)]^{-1/2} [U - E(U)]$ is asymptotically N(0, 1). In [4, Section 7] it was shown that the conditions [1, (8.2 - 8.4)] are satisfied if $0 \leq \Phi \leq 1$ and

(4.10) $$n^2\ \mathrm{var}(U) \longrightarrow \infty \qquad \text{as } n \longrightarrow \infty \quad .$$

Now clearly $0 \leq \Phi \leq 1$ [see (4.3)]. Hence all that remains in proving the asymptotic normality of w is to prove (4.10). Note that, by virtue of (4.4) and (4.8 - 4.9), the relation (4.10) is equivalent to

(4.11) $$n^2\ \mathrm{var}\ (w) \longrightarrow \infty \qquad \text{as } n \longrightarrow \infty \quad .$$

Thus we want to prove (4.11) for all τ, $0 < \tau < 1$. Let us use Γ' to denote that subset of Γ consisting of all quadruples (i,I,j,J) which belong to Γ and which satisfy (4.7). Then, since ρ (3.4) is always ≥ 0, it follows from (3.3) that

(4.12) $$\mathrm{var}(w) \geq (1/2\pi T^2) \sum_{\Gamma'} \sum_{\Gamma'} \sin^{-1} \rho\ (V_{iIjJ}, V_{i'I'j'J'}) \quad .$$

Since $\sin^{-1} \rho \geq \rho$ $(0 \leq \rho \leq 1)$, we obtain from (4.12) and (3.4) the inequality

$$(4.13)\qquad \mathrm{var}(w) \geq (1/2\pi T^2)\Big[\tau \sum_{\Omega'}\sum_{\Omega'} \rho^{(1)}_{iIjJ,i'I'j'J'} + (1-\tau)\sum_{\Omega'}\sum_{\Omega'} \rho^{(2)}_{iIjJ,i'I'j'J'}\Big] .$$

Thus (4.13) tells us that (4.11) will be established (for all τ) if we can show that

$$(4.14)\qquad n^2T^{-2} \sum_{\Omega'}\sum_{\Omega'} \rho^{(k)}_{iIjJ,i'I'j'J'} \longrightarrow \infty \qquad \text{as } n \longrightarrow \infty$$

for $k = 1, 2$. It will suffice to prove (4.14) for $k = 1$, since the proof for $k = 2$ is analogous.

Let Ω_2^* denote the set of all octuples $(i, I, j, J;\ i', I', j', J')$ such that both (i, I, j, J) and (i', I', j', J') are in Ω', and such that $i = i'$ and $I = I'$; let Ω_1^* denote the set of all octuples $(i, I, j, J;\ i', I', j', J')$ such that (i, I, j, J) and (i', I', j', J') are both in Ω', and such that $\delta_{ii'} + \delta_{iI'} + \delta_{Ii'} + \delta_{II'} = 1$. Then we can write

$$(4.15)\qquad \sum_{\Omega'}\sum_{\Omega'} \rho^{(1)}_{iIjJ,i'I'j'J'} = \sum_{\Omega_1^*} \rho^{(1)}_{iIjJ,i'I'j'J'} + \sum_{\Omega_2^*} \rho^{(1)}_{iIjJ,i'I'j'J'} .$$

Now it follows from (3.5) and (4.7) that any $\rho^{(1)}$ associated with an octuple in Ω_1^* will be $\geq \frac{1}{2}\epsilon^4$, and any $\rho^{(1)}$ associated with an octuple in Ω_2^* will be $\geq \epsilon^4$. Hence (4.15) implies that

$$(4.16)\qquad \sum_{\Omega'}\sum_{\Omega'} \rho^{(1)}_{iIjJ,i'I'j'J'} \geq (\mu_1 + 2\mu_2)\left(\tfrac{1}{2}\epsilon^4\right) ,$$

where μ_1 and μ_2 are the numbers of (distinct) octuples belonging to Ω_1^* and Ω_2^* respectively.

Let ν denote the number of quadruples which belong to $\underline{\bigcap}'$. Assumptions 4B and 4C tell us that

$$\nu \geq P \binom{M}{2}\binom{N}{2} \quad , \tag{4.17}$$

where $P = P_1 P_2 (>0)$. Let ν_g ($g = 1, 2, \ldots, M$) denote the number of quadruples (i, I, j, J) in $\underline{\bigcap}'$ such that either $i = g$ or $I = g$. Then

$$\sum_{g=1}^{M} \nu_g = 2\,\nu \quad , \tag{4.18}$$

and

$$\sum_{g=1}^{M} \nu_g^{\,2} = \mu_1 + 2\,\mu_2 \quad . \tag{4.19}$$

The expression on the left-hand side of (4.19) attains its minimum value subject to the restriction (4.18) if $\nu_g = 2\,\nu/M$ for every g. Hence

$$\mu_1 + 2\,\mu_2 \geq \sum_{g=1}^{M} (2\,\nu/M)^2 = 4\,\nu^2/M \quad . \tag{4.20}$$

From (4.16), (4.20), and (4.17) it follows that

$$n^2\, T^{-2} \sum_{\underline{\bigcap}'} \sum_{\underline{\bigcap}'} \rho^{(1)}_{iIjJ,i'I'j'J'} \geq n^2\, T^{-2}\, P^2 (M - 1) \binom{M}{2} \binom{N}{2}^2 \epsilon^4 \quad . \tag{4.21}$$

Since T is $\leq \binom{M}{2}\binom{N}{2}$, (4.21) implies that

$$n^2\, T^{-2} \sum_{\underline{\bigcap}'} \sum_{\underline{\bigcap}'} \rho^{(1)}_{iIjJ,i'I'j'J'} \geq n(M + N) \binom{M}{2}^{-2} \binom{N}{2}^{-2} P^2 (M-1) \binom{M}{2} \binom{N}{2}^2 \epsilon^4$$

$$\geq 2nP^2 \epsilon^4 \quad ,$$

which is sufficient to prove (4.14) (for k = 1). This completes the proof that $[\mathrm{var}(w)]^{-1/2} (w - \frac{1}{2})$ is asymptotically N(0, 1) under H_o.

5. Consistency. In this section we prove that, under Assumptions 4A and 4B, the test (1.9) is consistent against all alternatives $\alpha_Y \neq \alpha_Z$. Our proof will use an argument similar to that given in [2, pp. 58-59].

As a preliminary, we establish a bound for var(w). From (1.4) we have

$$\operatorname{var}(w) = T^{-2} \sum_{\Omega} \sum_{\Omega} \operatorname{cov}[u(V_{iIjJ}), u(V_{i'I'j'J'})] .$$

Hence

$$(5.1) \quad \operatorname{var}(w) \leq T^{-2} \sum_{i<I} \sum_{j<J} \sum_{i'<I'} \sum_{j'<J'} |\operatorname{cov}[u(V_{iIjJ}), u(V_{i'I'j'J'})]|$$

$$\leq T^{-2} \sum_{i<I} \sum_{j<J} \left\{ \left[\binom{M}{2}\binom{N}{2} - \binom{M-2}{2}\binom{N-2}{2} \right] (1) + \left[\binom{M-2}{2}\binom{N-2}{2} \right] (0) \right\}$$

$$= T^{-2} \binom{M}{2}\binom{N}{2} \left[\binom{M}{2}\binom{N}{2} - \binom{M-2}{2}\binom{N-2}{2} \right] .$$

At this point we apply Assumption 4B to (5.1) and obtain

$$(5.2) \quad \operatorname{var}(w) \leq P_1^{-2} \binom{M}{2}^{-2} \binom{N}{2}^{-2} \binom{M}{2}\binom{N}{2} \left[\binom{M}{2}\binom{N}{2} - \binom{M-2}{2}\binom{N-2}{2} \right]$$

$$\leq P_1^{-2} \left[1 - \left(1 - \frac{2}{M-1}\right)^2 \left(1 - \frac{2}{N-1}\right)^2 \right]$$

$$\leq 6\, P_1^{-2} (M^{-1} + N^{-1}) \qquad \text{for } M > 3,\ N > 3 .$$

This bound (5.2) is of course valid regardless of the values of α_Y, α_Z, τ, or other parameters.

From (5.2) and (1.8) we obtain immediately the inequality

$$Q \leq 6\, P_1^{-2}(M^{-1} + N^{-1}) \qquad (5.3)$$

We now proceed with the consistency proof. The proof is virtually the same whether $\alpha_Z > \alpha_Y$ or $\alpha_Z < \alpha_Y$; we give the proof only for $\alpha_Z > \alpha_Y$. If $\alpha_Z > \alpha_Y$, then we use (2.5) to write

$$\begin{aligned} P\left\{\text{rejecting } H_o\right\} &= P\left\{ Q^{-\frac{1}{2}} \left| w - \tfrac{1}{2} \right| > z_{\alpha/2}\right\} \\ &\geq P\left\{ Q^{-\frac{1}{2}}(w - \tfrac{1}{2}) > z_{\alpha/2}\right\} \\ &\geq P\left\{ w - E(w) > Q^{\frac{1}{2}} z_{\alpha/2} - d \right\} \end{aligned} \qquad (5.4)$$

where d is defined by (2.4). Because of (5.3) and Assumption 4A, Q will become arbitrarily small as n (= M + N) becomes large. Hence we can apply Tchebycheff's inequality to (5.4) to obtain the relation

$$P\left\{\text{rejecting } H_o\right\} \geq 1 - \frac{\text{var}(w)}{(d - Q^{\frac{1}{2}} z_{\alpha/2})^2} \qquad (5.5)$$

for n sufficiently large that d exceeds $Q^{1/2} z_{\alpha/2}$. Finally, from (5.2) and Assumption 4A it follows that the right-hand side of (5.5) approaches 1 as $n \longrightarrow \infty$.

6. Concluding remarks. We make some final observations:

(i) We might point out that V_{iIjJ} (1.5) is just one member of an infinite class of linear functions of (Y_i, Y_I, Z_j, Z_J) having medians (means) equal to $(\alpha_Z - \alpha_Y)$. The original reason why V_{iIjJ} was chosen was that V_{iIjJ} is the only linear function in the class whose variance is some (known) constant times

$(\sigma_e{}^2 + \sigma_f{}^2)$: this resulted in ρ (3.4, 3.11) being a linear function of τ, which in turn enabled us to prove (3.14) with no difficulty. It appeared that a bound analogous to (1.8) would not be easy to prove if any function other than V_{iIjJ} (1.5) had been used.

The choice of the function V_{iIjJ}, however, can also be considered in a different light. Let V^o represent any member of the class (which we will call C) of linear functions of (Y_i, Y_I, Z_j, Z_J) with medians $(\alpha_Z - \alpha_Y)$. Then if the quantity $\text{var}(V^o)/(\sigma_e{}^2 + \sigma_f{}^2)$ is considered as a sort of loss function, it is easily shown that the minimax value of this "loss function" will be achieved for $V^o = V_{iIjJ}$ (1.5):

$$\min_{V^o \epsilon C} \max_{\tau} \frac{\text{var}(V^o)}{\sigma_e{}^2 + \sigma_f{}^2} = \frac{\text{var}(V_{iIjJ})}{\sigma_e{}^2 + \sigma_f{}^2} .$$

Thus, in a certain sense, the choice of the linear function V_{iIjJ} (1.5) constitutes a "minimax variance" choice.

(ii) The original reason for excluding from the summation in (1.4) those (i, I, j, J) not belonging to $\underline{(\bar{\ })}$ was that, if (1.6b) is not satisfied, then the coefficients of Y_i, Y_I in (1.5) will not both be negative and the coefficients of Z_j, Z_J in (1.5) will not both be positive. On the other hand, for any (i,I,j,J) in $\underline{(\bar{\ })}$, V_{iIjJ} (1.5) will have its coefficients of Y_i, Y_I both negative and its coefficients of Z_j, Z_J both positive, thereby ensuring that $\rho^{(1)}$ (3.5) and $\rho^{(2)}$ (3.6), and hence ρ (3.4, 3.11), are ≥ 0; the fact that $\rho \geq 0$ was used in proving (3.14).

Restricting the summation in (1.4) only to those (i, I, j, J) which belong to $\underline{(\bar{\ })}$ has certain other effects in addition to providing an important link in the proof of (3.14). It is not known whether a relation analogous to (1.8) could be established if the summation in (1.4) were taken over all (i, I, j, J) instead of just over those (i, I, j, J)belonging to Ω, although such a relation would probably be difficult to prove even if it were true. Even if such a relation could be proved, however, it is not certain that the associated test would necessarily have better power than our test (1.9). Although this hypothetical test based on all V_{iIjJ}'s would obviously have better power in the extreme case where T is near 0, it would not necessarily also have better power in all other cases. One effect of taking the summation (1.4) just over Ω is that all V_{iIjJ}'s with variance $\leq \sigma_e^2 + \sigma_f^2$ are included in the summation while all V_{iIjJ}'s with variance $> \sigma_e^2 + \sigma_f^2$ are excluded from the summation, so that w (1.4) is (in a certain sense) based only on those V_{iIjJ}'s with the best discriminating capabilities, so to speak. For this reason, it seems possible that the test (1.9) could in some situations be more powerful than the hypothetical test based on all V_{iIjJ}'s.

The test (1.9) obviously is somewhat more advantageous with respect to calculations than the hypothetical test, since the summations Q_k (3.7) contain fewer terms than would their analogues under the hypothetical test based on all V_{iIjJ}'s.

We also note that the (i, I, j,J) in Ω constitute all the (i, I, j, J) such that the intervals (X_i, X_I) and (W_j, W_J) have point(s) in common. However, it does not appear that this effect as such has any consequences.

(iii) It was pointed out that (1.7) holds even for certain non-normal distributions, whereas (1.8) was proved only under the assumption of normality. It is not known to what extent (3.10) is satisfied for non-normal distributions, but we might conjecture that (3.10) is satisfied for some non-normal distributions but not others.

If this is the case, it might be possible to obtain some number $> Q$ which would be an upper bound on var(w) for a large class of distributions. This provides an area for further investigation.

(iv) It appears that the test (1.9) is not unbiased: note that the probability of rejection when H_o is true (although always $\leq \alpha$ approximately) will not be constant, but rather will vary with τ.

(v) Only the two-tailed test was discussed in this paper. However, the extension to one-tailed tests is immediate.

(vi) We can obtain confidence bounds on $(\alpha_Z - \alpha_Y)$ associated with the test (1.9). The technique for getting the bounds is similar to the one often used with the ordinary Wilcoxon statistic: we find that value of $\Delta = (\alpha_Z - \alpha_Y)$ which, when subtracted from every V_{iIjJ} in (1.4), will cause the resulting new w to be on the threshhold of significance.

Note that, if it is only desired to test the hypothesis (1.3) and no confidence bounds on Δ are needed, then it is only necessary to compute the numerator of each V_{iIjJ} (1.5) rather than the entire fraction (1.5), inasmuch as

$$u(V_{iIjJ}) = u\left[(X_I - W_j)(Z_J - Y_i) - (W_J - X_i)(Y_I - Z_j)\right] \quad .$$

However, if confidence bounds on Δ are to be obtained, then the full fractions V_{iIjJ} (1.5) all need to be calculated.

(vii) The performance of the numerical calculations for the test (1.9) is covered in a different report [5] in more detail than has been given here. For example, [5] points out that, if the formula

$$2 \sin^{-1} \rho = \cos^{-1}(1 - 2\rho^2)$$

is utilized in calculating the terms in the summations (3.7), this obviates the need for obtaining any square roots.

The calculation of Q rather than of w of course presents the principal computational problem associated with the test (1.9). The calculation of w is simple compared with the calculation of Q.

Since Q is a function only of the X_i's and W_j's and not of the Y_i's and Z_j's, Q can be calculated before the Y_i's and Z_j's are ever even available. This fact may be of advantage in experimental situations where the Y_i's and Z_j's represent the outcome of the experiment and the X_i's and W_j's are available before the experiment begins.

(viii) Calculating Q by hand is virtually out of the question: a high-speed computer is required. Moreover, it appears that even a high-speed computer is not fast enough to calculate Q economically except when M and N are rather small, since the sums Q_1, Q_2 (3.7) each involve huge numbers of terms even for moderate M, N.

Fortunately, however, there is a way to circumvent this problem of evaluating Q. Instead of determining Q exactly, we can calculate an adequate estimate of Q by taking a sample of the elements which make up the sums (3.7). It appears that a close estimate of Q can be obtained in this manner without costing an excessive amount of computer time. Different kinds of sampling are possible, including stratified sampling. A more detailed discussion of possible sampling techniques is given in [5].

A different approach to circumventing the evaluation of Q might be to try to obtain a general method for finding some number which is slightly greater than Q but which is much easier to calculate than Q. Such a number could then be used in lieu of Q. To implement this approach would probably require considerable investigation, and it is hard to tell whether the approach could work at all well.

7. Acknowledgments. The author wishes to thank Dr. Frederic M. Lord of Educational Testing Service for suggesting the problem of developing a test of the hypothesis that two parallel regression lines are identical when the variances are not necessarily equal.

REFERENCES

[1] HOEFFDING, WASSILY (1948). A class of statistics with asymptotically normal distributions. Ann. Math. Statist., vol. 19, pp. 293-325.

[2] MANN, H. B., and WHITNEY, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Statist., vol. 18, pp. 50-60.

[3] POTTHOFF, RICHARD F. (1962). Use of the Wilcoxon statistic for a generalized Behrens-Fisher problem. Institute of Statistics Mimeo Series No. 315, Department of Statistics, University of North Carolina.

[4] POTTHOFF, RICHARD F. (1962). A test of whether two regression lines are parallel when the variances may be unequal. Institute of Statistics Mimeo Series No. 331, Department of Statistics, University of North Carolina.

[5] POTTHOFF, RICHARD F. (1962). Illustration of a test which compares two parallel regression lines when the variances are unequal. Institute of Statistics Mimeo Series No. 323, Department of Statistics, University of North Carolina.

[6] WILCOXON, FRANK (1945). Individual comparison by ranking methods. Biometrics, vol. 1, pp. 80-83.