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

ILLUSTRATION OF A TECHNIQUE WHICH TESTS WHETHER TWO REGRESSION LINES ARE PARALLEL WHEN THE VARIANCES ARE UNEQUAL

by

Richard F. Potthoff

February, 1962

Contract No. AF 49(638)-213

It may be desired to test the hypothesis that two regression lines are parallel without assuming that the variances of the two sets of error terms are necessarily equal. This paper presents a relatively non-technical discussion of a test which can be used for this situation. A numerical illustration is included. The test statistic used is analogous to the well-known Wilcoxon statistic. This paper is intended for the practitioner rather than for the theoretician; the more technical aspects of the test are covered in a separate paper.

This research was supported partially by the Mathematics Division of the Air Force Office of Scientific Research and partially by Educational Testing Service.

Institute of Statistics
Mimeo Series No. 320

ILLUSTRATION OF A TECHNIQUE WHICH TESTS WHETHER TWO REGRESSION LINES ARE PARALLEL WHEN THE VARIANCES ARE UNEQUAL[1]

by

Richard F. Potthoff

1. Introduction. Suppose we have m pairs $(Y_1, X_1), (Y_2, X_2), \ldots, (Y_m, X_m)$, such that, for each i, Y_i observes the relation

$$(1.1) \qquad Y_i = \alpha_Y + \beta_Y X_i + e_i \quad ,$$

where α_Y and β_Y are unknown parameters (regression coefficients), the X_i's are specified constants, and the e_i's are normal and independent with mean 0 and unknown variance σ_e^2. Suppose also that we have n pairs (Z_1, W_1), $(Z_2, W_2), \ldots, (Z_n, W_n)$, such that, for each j, Z_j observes the relation

$$(1.2) \qquad Z_j = \alpha_Z + \beta_Z W_j + f_j \quad ,$$

where α_Z and β_Z are unknown parameters (regression coefficients), the W_j's are specified constants, and the f_j's are normal and independent with mean 0 and unknown variance σ_f^2. We will assume (with no loss of generality) that $m \leq n$. Note that we are allowing for the possibility that the variance of the e_i's may be different from the variance of the f_j's. Suppose we desire to test the hypothesis that the two β-coefficients are equal, i.e., the hypothesis that $\beta_Y = \beta_Z$. In other words, we want to test the hypothesis that the two regression lines associated with (1.1) and (1.2) are parallel, and of course we want to have a test which will be valid regardless of what the values of σ_e^2 and σ_f^2 are. In this paper we shall present such a test, along with a numerical example which will illustrate how the computations are made. Not only can we test the hypothesis

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

$\beta_Y = \beta_Z$, but we can also obtain confidence bounds on $(\beta_Z - \beta_Y)$.

A practical situation in which the problem treated by this paper could arise might be as follows. Suppose that we have $(m + n)$ classes of students, m of which receive curriculum number 1 and the remaining n of which receive curriculum number 2. Suppose also that, for each of the $(m + n)$ classes, we have available (i) some standard measure of the achievement of the class obtained after completion of the course, and (ii) a concomitant variate representing some standard measure of the ability of the class obtained before the course started. Then, for the i-th class receiving curriculum number 1, Y_i would be the achievement measure of this class obtained after the course, and X_i would be the ability measure of the class obtained before the course started; for the j-th class receiving curriculum number 2, Z_j would be the achievement measure of the class obtained after the course, and W_j would be the ability measure of the class obtained before the course. It is assumed that the Y_i's are connected with the X_i's by a linear relation of the form (1.1), and that the Z_j's are associated with the W_j's via a linear regression of the form (1.2); the variances associated with the two curriculums may be different. We suppose that we desire to compare curriculum number 1 with curriculum number 2 to determine which one is better. As a first step in this comparison, we will probably want to test the hypothesis that $\beta_Y = \beta_Z$; this is essentially the hypothesis that the difference between the effects of the two curriculums (if any) is the same irrespective of ability level. (If β_Y is different from β_Z, then that means that the difference between the effects of the two curriculums is not constant but rather depends on the ability level, and might even mean that one curriculum is better for brighter students while the other is better for duller students.)

The discussion in this paper will be relatively non-technical. A separate paper covers the technical aspects of the topic of the present paper.

2. The test. We now present the formula for the test statistic. For every pair (i, I) such that $1 \leq i < I \leq m$ [there are $\frac{1}{2}m(m-1)$ such pairs altogether], let us define

$$(2.1) \qquad C_{iI} = \frac{Y_I - Y_i}{X_I - X_i} .$$

For every pair (j,J) such that $1 \leq j < J \leq n$ [there are $\frac{1}{2}n(n-1)$ such pairs altogether], let us define

$$(2.2) \qquad D_{jJ} = \frac{Z_J - Z_j}{W_J - W_j} .$$

Now C_{iI} (2.1) will have expected value and median β_Y, while the expectation and median of D_{jJ} (2.2) will be β_Z. Thus the expected value and median of

$$(2.3) \qquad V_{iIjJ} = D_{jJ} - C_{iI} = \frac{Z_J - Z_j}{W_J - W_j} - \frac{Y_I - Y_i}{X_I - X_i}$$

will be $\beta_Z - \beta_Y$. Hence, if the null hypothesis $\beta_Y = \beta_Z$ is true, we would expect about half the V_{iIjJ}'s to be positive and half to be negative.

Altogether there will be a total of

$$\left[\frac{1}{2}m(m-1)\right] \times \left[\frac{1}{2}n(n-1)\right] = \frac{1}{4}mn(m-1)(n-1)$$

different V_{iIjJ}'s. Let S be the number of V_{ijIJ}'s which are positive, and let w be the proportion which are positive. Then

$$(2.4) \qquad w = \frac{4S}{mn(m-1)(n-1)} .$$

This statistic w will have expected value $\frac{1}{2}$ if the null hypothesis $\beta_Y = \beta_Z$ is true, and will be approximately normally distributed. The variance of w when the hypothesis $\beta_Y = \beta_Z$ is true will depend on σ_e^2 and σ_f^2, but it can be shown that, regardless of what σ_e^2 and σ_f^2 are, (2.5) $$\text{var}(w) \leq \frac{2m+5}{18m(m-1)} \quad .$$

(Remember that m is the smaller of m,n if m and n are different).

The test is as follows: if we want (e.g.) to use a two-tailed test at the 5°/o level to test the hypothesis $\beta_Y = \beta_Z$, we can reject the hypothesis if

(2.6) $$\sqrt{\frac{18m(m-1)}{2m+5}} \left| w - \frac{1}{2} \right| > 1.96$$

and accept otherwise.

This test (2.6) will be conservative in the sense that the probability of rejecting the null hypothesis when it is true will not be exactly equal to 5°/o, but will generally be somewhat less than 5°/o . The reason for this is that var (w) (which is unknown) will generally be somewhat below the bound (2.5).

It is possible that some of the V_{iIjJ}'s (2.3) may be undefined because of a condition $X_i = X_I$ and/or $W_j = W_J$. It is also possible that some of the V_{iIjJ}'s, although well-defined, may turn out to be exactly zero, so that they are neither positive nor negative. In either of these cases, we can handle the situation by tallying $\frac{1}{2}$ for each such V_{iIjJ} in the determination of S. For example, suppose that m = 8, n = 10. Then there will be 28 x 45 = 1260 potential V_{iIjJ}'s. Suppose that all the X_i's are different, but that there are two W_j's which are alike (with the rest all being different). Then 28 of the V_{iIjJ}'s will be undefined. Of the remaining 1232, suppose that 5 are zero, 711 are positive, and 516 are negative. Then we can calculate

$$S = 711 + \frac{1}{2}(28 + 5) = 727.5 \quad .$$

It is most desirable to have all the X_i's different and all the W_j's different, if possible.

The statistic w (2.4) used here bears a certain similarity to the well-known Wilcoxon statistic, whose use with respect to a problem simpler than but somewhat similar to the one considered in this paper is discussed in [1, 2]. It might be noted that, in the problem considered in [1] and [2], it was necessary to assume only that both sets of error terms followed symmetrical distributions, whereas in this paper we have assumed that both sets of error terms are normal (although with possibly different variances). The reason for the assumption of normality here is that this assumption was used in proving the bound (2.5). It is not known whether there are symmetrical distributions besides the normal for which the bound (2.5) would still hold, but, in view of the results of [1], which are concerned with an analogous situation, it would not be too surprising if (2.5) turned out to be valid for a wide class of symmetrical distributions rather than just for the normal. If such proved to be the case, we could of course relax the assumption of normality.

3. Numerical example. We now give a numerical example to illustrate the computation of the test statistic presented in the previous section. Suppose that $m = 6$ and $n = 7$, and suppose that our two samples are

$$Y_1 = 482.9,\ Y_2 = 538.7,\ Y_3 = 557.1,\ Y_4 = 591.2,\ Y_5 = 597.1,\ Y_6 = 650.6$$
$$X_1 = 92,\ X_2 = 102,\ X_3 = 108,\ X_4 = 112,\ X_5 = 117,\ X_6 = 126$$

and

$$Z_1 = 514.0,\ Z_2 = 527.7,\ Z_3 = 530.0,\ Z_4 = 537.3,\ Z_5 = 538.8,\ Z_6 = 550.1,\ Z_7 = 553.3$$
$$W_1 = 92,\ W_2 = 99,\ W_3 = 100,\ W_4 = 103,\ W_5 = 105,\ W_6 = 109,\ W_7 = 114\ .$$

For the sake of orderliness, we arranged the first sample so that it is in order of increasing magnitude of the X_i's, and we arranged the second sample so that it is in order of increasing magnitude of the W_j's.

Our first step is to compute the C_{iI}'s (2.1) and the D_{jJ}'s (2.2). There will be $\frac{1}{2} \cdot 6(6 - 1) = 15$ C_{iI}'s and $\frac{1}{2} \cdot 7(7 - 1) = 21$ D_{jJ}'s. We obtain

$$C_{12} = \frac{Y_2 - Y_1}{X_2 - X_1} = \frac{538.7 - 482.9}{102 - 92} = \frac{55.8}{10} = 5.58$$

$$C_{13} = \frac{Y_3 - Y_1}{X_3 - X_1} = \frac{557.1 - 482.9}{108 - 92} = 4.64$$

$$C_{14} = \frac{591.2 - 482.9}{112 - 92} = 5.42$$

$C_{15} = 4.57$, $C_{16} = 4.93$, $C_{23} = 3.07$, $C_{24} = 5.25$, $C_{25} = 3.89$, $C_{26} = 4.66$,

$C_{34} = 8.52$, $C_{35} = 4.44$, $C_{36} = 5.19$, $C_{45} = 1.18$, $C_{46} = 4.24$, $C_{56} = 5.94$.

Also

$$D_{12} = \frac{Z_2 - Z_1}{W_2 - W_1} = \frac{527.7 - 514.0}{99 - 92} = 1.96$$

$$D_{13} = \frac{530.0 - 514.0}{100 - 92} = 2.00$$

$D_{14} = 2.12$, $D_{15} = 1.91$, $D_{16} = 2.12$, $D_{17} = 1.79$,

$D_{23} = 2.30$, $D_{24} = 2.40$, $D_{25} = 1.85$, $D_{26} = 2.24$, $D_{27} = 1.71$,

$D_{34} = 2.43$, $D_{35} = 1.76$, $D_{36} = 2.23$, $D_{37} = 1.66$, $D_{45} = 0.75$,

$D_{46} = 2.13$, $D_{47} = 1.45$, $D_{56} = 2.82$, $D_{57} = 1.61$, $D_{67} = 0.64$.

From this point on, the computation is closely similar to the computation of the Wilcoxon statistic, which was discussed in [2]. We have 15 C_{iI}'s and 21 D_{jJ}'s. Let us form a 21 x 15 chart with 315 entries (see Table I.). Each row in the chart pertains to a D_{jJ}, and each column to a C_{iI}. For convenience, both the C_{iI}'s and the D_{jJ}'s are arranged in increasing order of magnitude. The 315 entries in the chart represent the 315 V_{iIjJ}'s (2.3).

We count up and find that, of the 315 V_{iIjJ}'s, 19 are positive and the remaining 296 are negative. Hence

$$w = \frac{S}{315} = \frac{19}{315} = .060$$

Thus our test statistic [see (2.6)] is

$$\sqrt{\frac{18m(m-1)}{2m+5}}\,(w-\tfrac{1}{2}) = \sqrt{\frac{18x6x5}{2x6+5}}\,(.060-.500) = \sqrt{\frac{540}{17}}(-.440)=5.636(-.440)= -2.48 \quad .$$

Since -2.48 is greater in absolute value than 1.96, we reject the null hypothesis $\beta_Y = \beta_Z$ if we are making a two-tailed test at the 5°/o level.

It is of course also possible to make a one-tailed test of the hypothesis $\beta_Y = \beta_Z$ (against the alternative $\beta_Z > \beta_Y$ or the alternative $\beta_Z < \beta_Y$) if this is what is desired.

If we are interested only in testing the hypothesis $\beta_Y = \beta_Z$ and are not interested in getting confidence bounds on $(\beta_Z - \beta_Y)$, it is not necessary to compute the 315 entries of Table I. It will suffice simply to find out how many of these extries are positive and how many negative.

To find this out, an alternative computing technique, which operates by ranking the combined group of $\frac{1}{2}$ m(m-1) C_{iI}'s and $\frac{1}{2}$ n(n-1) D_{jJ}'s, is available. This alternative technique, which is described in [2], provides perhaps the easiest way of determining S for testing the hypothesis $\beta_Y = \beta_Z$, assuming that no confidence bounds on $(\beta_Z - \beta_Y)$ are wanted.

If confidence bounds are wanted, however, we can obtain them by utilizing Table I.

4. Confidence bounds. The method of obtaining confidence bounds here is closely analogous to the method described and illustrated in [2]. Suppose, for example, that we want to obtain a two-sided 95°/o confidence interval for $(\beta_Z-\beta_Y)$.

To do this, we find that value of $\delta = \beta_Z - \beta_Y$ which, when subtracted from every entry in Table I., would cause (the resulting new) w to be on the threshhold of significance. Now w will be on the threshhold of being significantly large if 267 of the 315 entries are positive and 1 entry is zero, since

$$(1.960 \times \sqrt{\frac{17}{540}} + \frac{1}{2}) \times 315 = (\frac{1.960}{5.636} + .500) \times 315 = .848 \times 315 = 267.1 ;$$

and w will be on the threshhold of being significantly small if 47 of the 315 entries in the table are positive and 1 entry is zero, since

$$(-\frac{1.960}{5.636} + \frac{1}{2}) \times 315 = .152 \times 315 = 47.9 .$$

Upon looking at Table I., we can determine that 267 of its 315 entries are larger than -3.94 (with 1 entry being exactly equal to -3.94). Thus -3.94 is the lower end of our confidence interval. We also find that 47 of the entries in the table exceed -1.66 (one entry is equal to -1.66 and the remaining 267 entries are smaller than -1.66). Hence -1.66 is the upper end of the confidence interval, and we can state that

$$-3.94 \leq (\beta_Z - \beta_Y) \leq -1.66$$

with confidence coefficient $\geq 95^{\circ}/o$.

In most practical situations the experimenter would never find out what the true values of the parameters were. However, the example in this paper was constructed artificially, so that all the parameters are known. The values used were $\alpha_Y = 20$, $\beta_Y = 5$, $\alpha_Z = 330$, $\beta_Z = 2$, $\sigma_e^2 = 100$, and $\sigma_f^2 = 4$. Thus $(\beta_Z - \beta_Y) = -3$.

When the values of m and n are small, the computations which are required for obtaining the test statistic and confidence bounds discussed in this paper are not too lengthy and can be performed with a desk calculator. However, for a sufficiently heavy volume of data it would probably pay to utilize more automatic methods to do the calculations.

REFERENCES

[1] Richard F. Potthoff, "Use of the Wilcoxon Statistic for a Generalized Behrens-Fisher Problem," Institute of Statistics Mimeograph Series No. 315, University of North Carolina.

[2] Richard F. Potthoff, "Comparing the Means of Two Symmetrical Populations," Institute of Statistics Mimeograph Series No. 316, University of North Carolina.

TABLE I.

	C_{45}= 1.18	C_{23}= 3.07	C_{25}= 3.89	C_{46}= 4.24	C_{35}= 4.44	C_{15}= 4.57	C_{13}= 4.64	C_{26}= 4.66	C_{16}= 4.93	C_{36}= 5.19	C_{24}= 5.25	C_{14}= 5.42	C_{12}= 5.58	C_{56}= 5.94	C_{34}= 8.52
D_{67}=0.64	-0.54	-2.43	-3.25	-3.60	-3.80	-3.93	-4.00	-4.02	-4.29	-4.55	-4.61	-4.78	-4.94	-5.30	-7.88
D_{45}=0.75	-0.43	-2.32	-3.14	-3.49	-3.69	-3.82	-3.89	-3.91	-4.18	-4.44	-4.50	-4.67	-4.83	-5.19	-7.77
D_{47}=1.45	0.27	-1.62	-2.44	-2.79	-2.99	-3.12	-3.19	-3.21	-3.48	-3.74	-3.80	-3.97	-4.13	-4.49	-7.07
D_{57}=1.61	0.43	-1.46	-2.28	-2.63	-2.83	-2.96	-3.03	-3.05	-3.32	-3.58	-3.64	-3.81	-3.97	-4.33	-6.91
D_{37}=1.66	0.48	-1.41	-2.23	-2.58	-2.78	-2.91	-2.98	-3.00	-3.27	-3.53	-3.59	-3.76	-3.92	-4.28	-6.86
D_{27}=1.71	0.53	-1.36	-2.18	-2.53	-2.73	-2.86	-2.93	-2.95	-3.22	-3.48	-3.54	-3.71	-3.87	-4.23	-6.81
D_{35}=1.76	0.58	-1.31	-2.13	-2.48	-2.68	-2.81	-2.88	-2.90	-3.17	-3.43	-3.49	-3.66	-3.82	-4.18	-6.76
D_{17}=1.79	0.61	-1.28	-2.10	-2.45	-2.65	-2.78	-2.85	-2.87	-3.14	-3.40	-3.46	-3.63	-3.79	-4.15	-6.73
D_{25}=1.85	0.67	-1.22	-2.04	-2.39	-2.59	-2.72	-2.79	-2.81	-3.08	-3.34	-3.40	-3.57	-3.73	-4.09	-6.67
D_{15}=1.91	0.73	-1.16	-1.98	-2.33	-2.53	-2.66	-2.73	-2.75	-3.02	-3.28	-3.34	-3.51	-3.67	-4.03	-6.61
D_{12}=1.96	0.78	-1.11	-1.93	-2.28	-2.48	-2.61	-2.68	-2.70	-2.97	-3.23	-3.29	-3.46	-3.62	-3.98	-6.56
D_{13}=2.00	0.82	-1.07	-1.89	-2.24	-2.44	-2.57	-2.64	-2.66	-2.93	-3.19	-3.25	-3.42	-3.58	-3.94	-6.52
D_{14}=2.12	0.94	-0.95	-1.77	-2.12	-2.32	-2.45	-2.52	-2.54	-2.81	-3.07	-3.13	-3.30	-3.46	-3.82	-6.40
D_{16}=2.12	0.94	-0.95	-1.77	-2.12	-2.32	-2.45	-2.52	-2.54	-2.81	-3.07	-3.13	-3.30	-3.46	-3.82	-6.40
D_{46}=2.13	0.95	-0.94	-1.76	-2.11	-2.31	-2.44	-2.51	-2.53	-2.80	-3.06	-3.12	-3.29	-3.45	-3.81	-6.39
D_{36}=2.23	1.05	-0.84	-1.66	-2.01	-2.21	-2.34	-2.41	-2.43	-2.70	-2.96	-3.02	-3.19	-3.35	-3.71	-6.29
D_{26}=2.24	1.06	-0.83	-1.65	-2.00	-2.20	-2.33	-2.40	-2.42	-2.69	-2.95	-3.01	-3.18	-3.34	-3.70	-6.28
D_{23}=2.30	1.12	-0.77	-1.59	-1.94	-2.14	-2.27	-2.34	-2.36	-2.63	-2.89	-2.95	-3.12	-3.28	-3.64	-6.22
D_{24}=2.40	1.22	-0.67	-1.49	-1.84	-2.04	-2.17	-2.24	-2.26	-2.53	-2.79	-2.85	-3.02	-3.18	-3.54	-6.12
D_{34}=2.43	1.25	-0.64	-1.46	-1.81	-2.01	-2.14	-2.21	-2.23	-2.50	-2.76	-2.82	-2.99	-3.15	-3.51	-6.09
D_{56}=2.82	1.64	-0.25	-1.07	-1.42	-1.62	-1.75	-1.82	-1.84	-2.11	-2.37	-2.43	-2.60	-2.76	-3.12	-5.70