RANK CORRELATION AND REGRESSION IN A NONNORMAL SURFACE

F. N. DAVID[1] AND EVELYN FIX[2]
UNIVERSITY OF CALIFORNIA, BERKELEY

1. The problem

The problem, seemingly simple, which led to the work reported in this paper, was this: to find a "usable" test for a difference in position of two regression lines if it is known that the two lines have the same slope but it is also known that the marginal distributions are extremely skewed and one is loath to use transformations and normal theory may be inadequate. More precisely, we consider the following situation. Suppose that there are two sets of bivariate correlated observations, n_1 in the first set and n_2 in the second set, with $n_1 + n_2 = n$. Neither the functional form of the bivariate population (or populations) nor any of the parameters descriptive of the population (populations) from which the observations have been drawn is known. However, the number of observations (between 20 and 50) is sufficient to indicate that the marginal distributions may be extremely skew Type III and the regression may be linear. Furthermore, from the description of the experiments generating the bivariate observations, it seems reasonable to assume that, if indeed they do come from different populations, the regression lines of the parent populations have the same slope with the difference occurring in the intercepts or position of the lines. The data may be analyzed in a variety of ways, including the orthodox one of transforming the data and using normal theory tests. Here we discuss the possibility of using tests based on the rank correlation coefficient attributed to Spearman. We propose the use of these tests in their conditional form. In order that the criteria proposed may be compared with others, we suggest a bivariate functional form for the population, which we have been unable to find discussed elsewhere and which enables us to throw some light on the behavior of the mean value of Spearman's rho when the population is nonnormal.

2. Rank correlation in the normal correlation distribution

Before proceeding with the discussion of the problem outlined, it is useful to recall what is known about Spearman's rho in the case of the bivariate normal

[1] This paper was prepared with the partial support of the Office of Naval Research (Nonr-222-43).
[2] This investigation was supported (in part) by a research grant (No. RG-3666) from the National Institutes of Health, Public Health Service.

distribution when the correlation is not zero. Suppose that n pairs of observations $\{x_i, y_i\}$ are randomly and independently drawn from a bivariate normal population with correlation ρ. The $\{x_i\}$ and $\{y_i\}$ are each ranked in order of magnitude. The product moment correlation between the ranks, r_s, is, of course, Spearman's ρ. Sundrum [6] derived the exact distribution of r_s for $n = 3$. In general for all n, it is known (Moran [3]) that

$$E(r_s) = \frac{6}{\pi(n+1)}\left[\sin^{-1}\rho + (n-2)\sin^{-1}\frac{\rho}{2}\right]. \tag{1}$$

Kendall [2], modified by Fieller, Hartley, and Pearson [1], gave

$$\sigma_{r_s}^2 \doteqdot \frac{1}{n-1}[1 - 1.563465\rho^2 + 0.304743\rho^4 + 0.155286\rho^6 + 0.061552\rho^8 + 0.022099\rho^{10} + 0.019785\rho^{12}]. \tag{2}$$

David, Fix, and Mallows, in an unpublished manuscript, have shown that, to order ρ^2, this variance is

$$\sigma_{r_s}^2 \doteqdot \frac{1}{n-1} - \frac{3\rho^2}{\pi^2 n^{(2)}(n+1)^2}[(19n^3 - 89n^2 + 168n - 108) - 8\sqrt{3}(n-1)^{(3)}]. \tag{3}$$

Fieller, Hartley, and Pearson suggested from an empirical investigation that, if

$$z = \frac{1}{2}\log\frac{1+r_s}{1-r_s}, \tag{4}$$

then z is approximately normally distributed. If we write $E(r_s) = R$, then

$$E(z) \doteqdot \frac{1}{2}\log\frac{1+R}{1-R} + \sigma_{r_s}^2\frac{R}{(1-R^2)^2} \tag{5}$$

and

$$\operatorname{Var} z \doteqdot \sigma_{r_s}^2\frac{1}{(1-R^2)^2}. \tag{6}$$

This may be applied in the two sample problem to deriving a test for the equivalence of the slopes of the two regression lines. Thus, if r_1 and r_2 are the rank correlation coefficients of each sample ranked separately and the corresponding transforms z_1 and z_2 are calculated, the criterion under the null hypothesis of no difference of slope,

$$\frac{(z_1 - z_2)(1 - R^2)}{(\sigma_{r_1}^2 + \sigma_{r_2}^2)^{1/2}}, \tag{7}$$

is approximately a unit normal variable.

3. Criteria for position

The equivalence of position of two regression lines in terms of ranks does not appear to have been discussed. If one desires to analyze data in this way, it is difficult to choose criteria except by very general considerations, for there is no really workable procedure such as the likelihood ratio which will indicate the appropriate test to use against specified alternatives. Four criteria are put forward here, the first two by analogy with orthodox normal theory for the original observations, the last two on common-sense grounds. We consider the problem in the following terms.

Under the null hypothesis H_0, there will be one population Π. Let

$$E(x) = \xi, \qquad E(y) = \eta. \tag{8}$$

Under the alternative hypothesis H_1, there will be two populations Π_1 and Π_2. Let

$$\begin{aligned} E(x|\Pi_1) &= \xi_1, & E(y|\Pi_1) &= \eta_1, \\ E(x|\Pi_2) &= \xi_2, & E(y|\Pi_2) &= \eta_2. \end{aligned} \tag{9}$$

It should be recognized that, although the problem seems very like that of testing the difference between two means, there is a difference in that one has the additional information that the slopes of the regression lines under H_1 are the same and this information should not be ignored.

In terms of ranks, write R_{x_i} for the rank of x_i and R_{y_i} for the rank of y_i. For all four criteria designed to test a variant of H_0 and sensitive to a variant of H_1, it is assumed that the two samples are ranked together but that calculations are carried out on these ranks relating to one sample only.

(i) H_0: $\xi_1 = \xi_2 = \xi$. $\quad H_1$: $\xi_1 \neq \xi_2$.

Recall for a moment the situation in which there is a bivariate normal distribution with marginal means ξ and η, marginal standard deviations unity and correlation coefficient ρ, assumed known. Given a bivariate normal sample $x_i, y_i, i = 1, 2, \cdots, n$, from this distribution, we wish to test the hypothesis that $\xi = \xi_0$ (some specified value) against the alternative that $\xi \neq \xi_0$. The likelihood ratio criterion is

$$\bar{x} - \rho\bar{y}. \tag{10}$$

When ρ is not known, the criterion

$$\bar{x} - r\bar{y} \tag{11}$$

might be used where r is the sample product moment correlation coefficient.

Let R_{x_i} and R_{y_i} be the ranks of the pair x_i, y_i belonging to the first set of observations, n_1 in number, the ranks, however, as given by the combined ranked sample. By analogy with the procedure for the bivariate normal surface, we choose the criterion

$$T_1 = \frac{1}{n_1\sigma} \sum_{\mathbf{n}_1} \left[R_{x_i} - \frac{n+1}{2} - r_s \left(R_{y_i} - \frac{n+1}{2} \right) \right], \tag{12}$$

where the sum is over only those observations which belong to the first set of n_1 observations indicated by $\mathbf{n}_1$. Further, r_s is the Spearman product moment correlation coefficient for the combined set of n observations and $\sigma^2 = (n^2 - 1)/12$.

A similar criterion to test

$$H_0\colon \eta_1 = \eta_2 = \eta, \qquad H_1\colon \eta_1 \neq \eta_2,$$

will be

$$T_1^* = \frac{1}{n_1\sigma} \sum_{n_1} \left[R_{y_i} - \frac{n+1}{2} - r_s\left(R_{x_i} - \frac{n+1}{2}\right)\right]. \tag{13}$$

(ii) $H_0\colon \xi_1 = \xi_2 = \xi, \qquad \eta_1 = \eta_2 = \eta.$

$H_1\colon \xi_1 \neq \xi_2, \qquad \eta_1 \neq \eta_2.$

Again, one recalls that, given the same normal distribution as in (i) with ρ known, the likelihood ratio test procedure yields the criterion,

$$(\bar{x} - \xi)^2 - 2\rho(\bar{x} - \xi)(\bar{y} - \eta) + (\bar{y} - \eta)^2. \tag{14}$$

This suggests that the appropriate rank criterion would be

$$T_2 = \frac{1}{n_1^2\sigma^2}\left\{\left[\sum_{n_1}\left(R_{x_i} - \frac{n+1}{2}\right)\right]^2 - 2r_s\left[\sum_{n_1}\left(R_{x_i} - \frac{n+1}{2}\right)\right]\left[\sum_{n_1}\left(R_{y_i} - \frac{n+1}{2}\right)\right] + \left[\sum_{n_1}\left(R_{y_i} - \frac{n+1}{2}\right)\right]^2\right\}. \tag{15}$$

It is certain that the criteria T_1 and T_2 are not the optimum possible since the original variables $\{x_i, y_i\}$ are not normally distributed whereas the criteria are based on analogy with criteria obtained from underlying normality. Two other criteria may be suggested on common-sense grounds. These are the average of the sum of the distances from the line $R_x = R_y$ for one set of observations as one possibility and the average of the sum of the distances from the line $R_y = -R_x + (n + 1)$ as the other. Thus,

(iii) $H_0\colon \xi_1 = \xi_2 = \xi, \qquad \eta_1 = \eta_2 = \eta;$

$H_1\colon \xi_1 > \xi_2, \qquad \eta_1 < \eta_2; \qquad \xi_1 < \xi_2, \qquad \eta_1 > \eta_2.$

The criterion is

$$\varphi = \frac{1}{n_1\sigma} \sum_{n_1} (R_{x_i} - R_{y_i}). \tag{16}$$

Similarly, given that

(iv) $H_0\colon \xi_1 = \xi_2 = \xi, \qquad \eta_1 = \eta_2 = \eta;$

$H_1\colon \xi_1 < \xi_2, \qquad \eta_1 < \eta_2; \qquad \xi_1 > \xi_2, \qquad \eta_1 > \eta_2,$

the criterion is

$$\theta = \frac{1}{n_1\sigma} \sum_{n_1} [R_{x_i} + R_{y_i} - (n + 1)] \tag{17}$$

4. Conditional tests

Consider any one particular combined ranked sample and, for the sake of an example, the criterion T_1. There will be n pairs of ranks and we may construct n quantities

$$d_j = \left(R_{x_j} - \frac{n+1}{2}\right) - r_s\left(R_{y_j} - \frac{n+1}{2}\right). \tag{18}$$

This set of $\{d_j\}$ can be looked on as a finite population of n with polykays,

$$\begin{aligned} K_1 &= \frac{1}{n}\sum_{j=1}^{n} d_j = 0, \\ -n\,K_{11} = K_2 &= \frac{1}{n-1}\sum_{j=1}^{n} d_j^2 = \frac{n\sigma^2}{n-1}(1 - r_s^2). \end{aligned} \tag{19}$$

The second and higher polykays of this set are conditional on the r, that is, on the particular combined sample, as is expected. We treat the $\{d_j\}$ as a randomization set and consider all possible $\binom{n}{n_1}$ samples of n_1 which may be generated from it. T_1 is the mean of a sample of n_1 which under the null hypothesis is randomly drawn without replacement from the finite population of the $\{d_j\}$. Accordingly, from known results,

$$\begin{aligned} \sigma E(T_1) &= 0, \\ \sigma^2 \operatorname{Var}(T_1) &= K_2\left(\frac{1}{n_1} - \frac{1}{n}\right) = \frac{\sigma^2(n - n_1)}{n_1(n-1)}(1 - r_s^2) \end{aligned} \tag{20}$$

or

$$\operatorname{Var}\left[\frac{T_1}{(1 - r_s^2)^{1/2}}\right] = \frac{n - n_1}{n_1(n-1)}. \tag{21}$$

The third and fourth polykays are

(22)

$$K_3 = \frac{n^2}{n^{(3)}}\sum_{j=1}^{n} d_j^3, \qquad K_4 = \frac{n}{n^{(4)}}\left\{(n+1)^{(2)}\sum_{j=1}^{n} d_j^4 - 3(n-1)\left(\sum_{j=1}^{n} d_j^2\right)^2\right\},$$

whence

$$\sigma^2\mu_3(T_1) = \frac{n^2}{n^{(3)}}\left(\frac{1}{n_1} - \frac{1}{n}\right)\left(\frac{1}{n_1} - \frac{2}{n}\right)\left(\sum_{j=1}^{n} d_j^3\right) \tag{23}$$

and

$$\begin{aligned} \sigma^2\mu_4(T_1) = K_4&\left[\left(\frac{1}{n_1} - \frac{1}{n}\right)^3 - \frac{1}{n}\left(\frac{1}{n_1} - \frac{1}{n}\right)^2 + \frac{1}{n^2}\left(\frac{1}{n_1} - \frac{1}{n}\right)\right] \\ &+ \frac{3(n-1)}{n+1}\left(\frac{1}{n_1} - \frac{1}{n}\right)^2\left(K_2^2 - \frac{K_4}{n}\right). \end{aligned} \tag{24}$$

When $n_1 = n_2$ the third moment is zero and, in general, provided n_1 is not very different from n_2 the distribution of T_1 will not be far off normal. The distribution exhibits the expected saw-tooth effect characteristic of rank distributions and the β_2 is usually less than three. The significance levels for T_1 given β_1 and β_2 may be found directly from table 42 (Pearson and Merrington) in [5]. Usually, however, it will be enough to assume normality. The conditional test which we

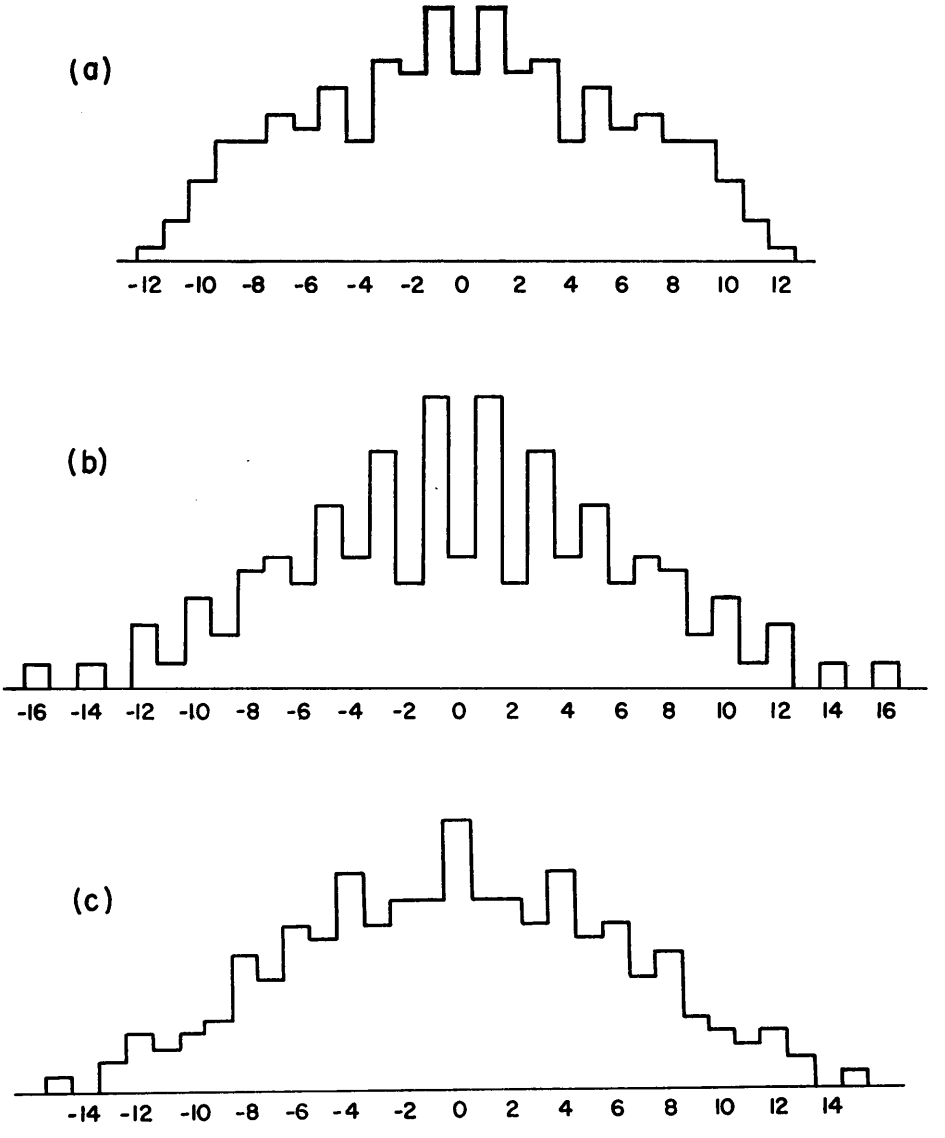


FIGURE 1

Randomization distribution of $n_1\sigma\varphi$ for a sample of $n_1 = 5$, for three particular samples of 10. Sample before ranking is bivariate normal, $\rho = 0.5$ and $\binom{10}{5} = 252$. Finite population.

Sample (a): $d = -3, -1, 0, 0, 0, 8, 4, -2, -4, -2$; $\beta_2 = 2.15$.

Sample (b): $d = 2, -2, 4, -4, 6, 0, -5, -5, 4, 0$; $\beta_2 = 2.66$.

Sample (c): $d = 3, 7, -5, -5, -1, -1, 2, 1, -3, 2$; $\beta_2 = 2.48$.

propose is that $T_1/(1 - r_s^2)^{1/2}$ be considered as a normal variable with variance $n_2/n_1(n - 1)$. The effect of $\beta_2 < 3$ will be to make the first kind of error less than that nominally given by the assumption of normality.

The same procedure may be used for φ and θ. We have that

$$E(\varphi) = 0 = E(\theta),$$

(25)
$$\operatorname{Var}\left[\frac{\varphi}{(1 - r_s)^{1/2}}\right] = \frac{2(n - n_1)}{n_1(n - 1)} = \operatorname{Var}\left[\frac{\theta}{(1 + r_s)^{1/2}}\right].$$

Then $\varphi/(1 - r_s)^{1/2}$ and $\theta/(1 + r_s)^{1/2}$ may each be assumed to be normally distributed. For purposes of illustration, the randomization distributions for $n_1\sigma\varphi$, $n = 10$, $n_1 = 5$ are given in figure 1 for three samples drawn from a bivariate normal with $\rho = 0.5$. It will be noted that even with such small numbers the assumption of normality for the distribution is not unseemly. Numerical comparisons of the critical levels are given in table I.

TABLE I

COMPARISONS OF THE CRITICAL VALUES

Sample	(a)			(b)			(c)			
Critical level	≧10	≧11	≧12	≧12	≧14	≧16	≧11	≧12	≧13	≧15
True	.040	.016	.004	.036	.016	.008	.040	.028	.012	.004
Normal	.046	.031	.021	.034	.016	.007	.039	.027	.018	.008
Pearson-Merrington	.042	.020	.006	.032	.014	....	.039	.024	.013	....

The criterion T_2 is the mean of a sum of squares. In the randomization set it may be assumed that T_2 is distributed (approximately) proportionally as χ^2 with two degrees of freedom. We have that

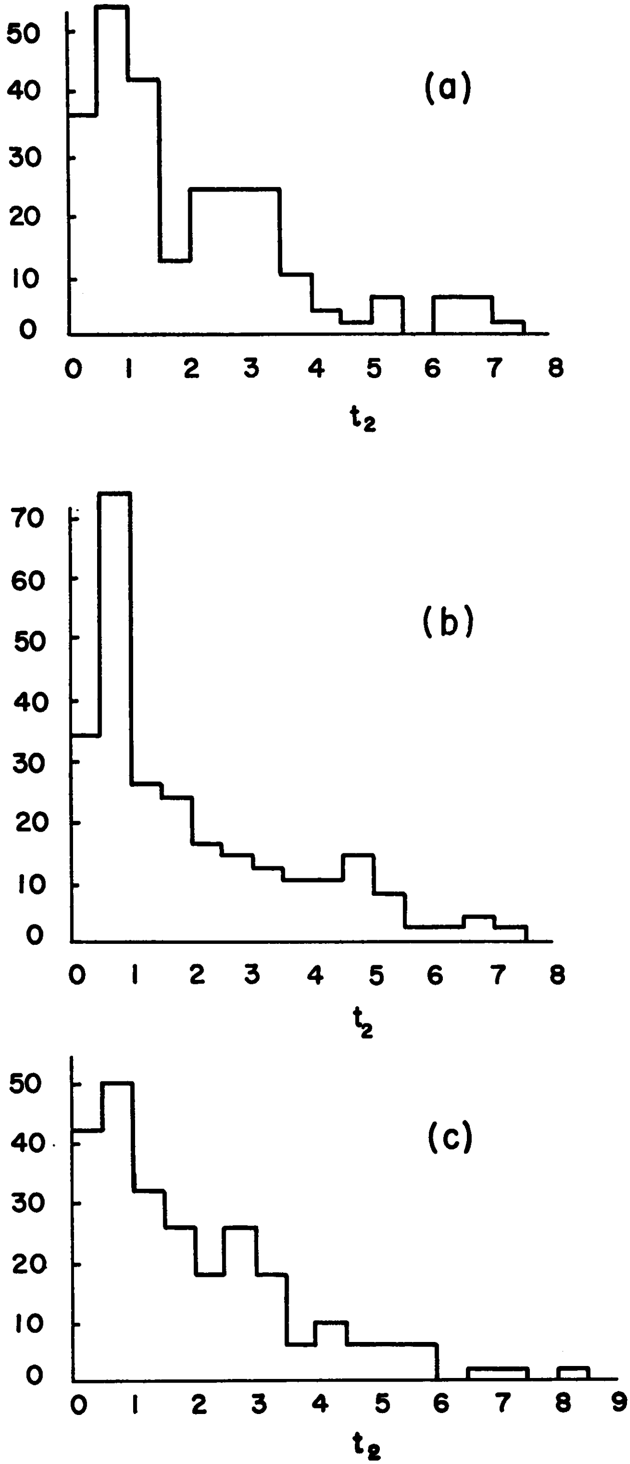
(26)
$$E(T_2) = \frac{2(n - n_1)}{n_1(n - 1)}(1 - r_s^2)$$

and the test of significance may be carried out from reference to the χ^2 tables. The agreement between the distribution of the $\binom{n}{n_1}$ values of T_2 and the modified χ^2 distribution is not very close for small samples and its use is not recommended for samples less than $10 + 10 = 20$.

If we write

(27)
$$t_2 = \frac{n_1(n - 1)}{n_2(1 - r_s^2)} T_2,$$

then t_2 is distributed (approximately) directly as χ^2. Table II gives the mean and standard deviation of t_2 in the randomization set when the original sample before ranking is randomly and independently drawn from the normal bivariate distribution with correlation $\rho = 0.5$.



(a)
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8
t_2
(b)
70
60
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8
t_2
(c)
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9
t_2

FIGURE 2

TABLE II

CRITERION t_2 (OR T_2) WHEN THE ORIGINAL SAMPLE IS BIVARIATE NORMAL ($\rho = 0.5$)

n	n_1	r_s	Mean t_2	Standard Deviation of t_2
10	5	0.636	2	1.716
10	5	0.091	2	1.720
10	5	0.467	2	1.703
10	4	0.636	2	1.702
9	4	0.733	2	1.566
7	3	0.679	2	1.501
6	3	−0.257	2	1.481

Three of the randomization distributions of t_2 are given in figure 2. As n, the combined sample size, increases and if n_1 is not very different from $n/2$, it is to be expected that the randomization distribution will be more closely approximated by the χ_2^2 or the modified χ_2^2 distribution.

5. Mean value of r_s for any surface

For all four criteria, the variances given in the preceding section are dependent on r_s, the rank product moment correlation coefficient of the particular sample. If it is required to use the criteria in an unconditional test, then it becomes necessary to discuss $E(r_s)$ for T_1, φ, and θ, and $E(r_s^2)$ for T_2. Such a discussion will need to be referred to the particular surface which it is thought might describe the original unranked observations. We begin with a few general considerations.

Define a function $H(t)$ such that

$$H(t) \begin{cases} = 0, & t \leqq 0 \\ = 1, & t > 0. \end{cases} \tag{28}$$

Then

$$R_{x_i} - 1 = \sum_{j=1}^{n} H(x_i - x_j), \qquad R_{y_i} - 1 = \sum_{m=1}^{n} H(y_i - y_m), \tag{29}$$

FIGURE 2

Randomization distribution of t_2.
Finite population; $n = 10 = 5 + 5$.

Sample (a): $R_{x_i} = 8, 7, 3, 5, 4, 2, 6, 9, 1, 10;$
$R_{y_i} = 8, 4, 1, 3, 5, 6, 10, 9, 2, 7.$

Sample (b): $R_{x_i} = 4, 1, 3, 7, 8, 6, 2, 10, 9, 5;$
$R_{y_i} = 7, 3, 6, 10, 1, 8, 2, 5, 4, 9.$

Sample (c): $R_{x_i} = 1, 8, 9, 4, 5, 2, 6, 7, 10, 3;$
$R_{y_i} = 5, 10, 7, 6, 2, 3, 9, 8, 4, 1.$

and since

$$r_s = \frac{\sum_{i=1}^{n} \left(R_{x_i} - \frac{n+1}{2}\right)\left(R_{y_i} - \frac{n+1}{2}\right)}{n\sigma^2}, \tag{30}$$

we have that

$$r_s = \frac{S - \left(\frac{n-1}{2}\right)^2}{n\sigma^2} \tag{31}$$

where

$$S = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{m=1}^{n} H(x_i - x_j)H(y_i - y_m). \tag{32}$$

Consequently the expectation of r_s reduces to the expectation of S which will be

(33)

$$\begin{aligned} E(S) &= n^{(3)}P\{x_i > y_j, y_i > y_m | i \neq j \neq m\} + n^{(2)}P\{x_i > x_j, y_i > y_j | i \neq j\} \\ &= n^{(3)}P_1 + n^{(2)}P_2, \end{aligned}$$

say, and

$$E(r_s) = \frac{12}{n+1}\left[(n-2)P_1 + P_2 - \frac{1}{4}(n-1)\right]. \tag{34}$$

This is the technique which Moran applied for evaluating $E(r_s)$ for the bivariate normal distribution and for which he found

$$P_1 = \frac{1}{2} - \frac{1}{2\pi}\cos^{-1}\frac{\rho}{2}, \qquad P_2 = \frac{1}{2} - \frac{1}{2\pi}\cos^{-1}\rho. \tag{35}$$

It is clear that the functional form assumed for the bivariate population distribution will affect the mean value of r_s.

6. Double Gamma distributions

In the first section we stated that the set of bivariate observations was such that the margins might be described by extremely skew Type III distributions and the regressions could be linear. This suggests that a surface with Type III margins might be a suitable functional form to describe the parent population generating the samples. Rhodes [4] proposed a surface of this kind which also had linear regressions. He was followed by Van Uven [7], who did not seem to be aware of Rhodes' work and who proposed essentially the same surface. Rhodes' distribution is however unsuitable for the present purposes since it is constrained to lie within a wedge-shaped area whereas we require the bivariate surface to take all values between zero and infinity. To meet the conditions of the problem we devised a distribution which must be known although it is new to us.

It was Weldon, in his well-known dice problem, who first suggested taking three independent variables x, y, z and studying the surface formed by X, Y, where

$$X = x + y, \qquad Y = x + z. \tag{36}$$

In Weldon's case, x, y, and z were each binomially distributed but it is clear that his distribution is only one of a broad class of bivariate populations resulting from the addition of independent random variables. This will be discussed elsewhere by C. L. Mallows. For present purposes, we propose the addition of independent Gamma variables. The resulting surface has the merit of simplicity but has a discontinuity of functional form which is not entirely satisfactory.

Let A, B, and C be independent random variables with

$$\begin{aligned} p(A) &= \frac{1}{\Gamma(a)} A^{a-1}e^{-A}, \qquad p(B) = \frac{1}{\Gamma(b)} B^{b-1}e^{-B}, \\ p(C) &= \frac{1}{\Gamma(c)} C^{c-1}e^{-C}, \qquad a, b, c > 0; \qquad A, B, C > 0. \end{aligned} \tag{37}$$

Define

$$U = A + B, \qquad V = A + C, \tag{38}$$

whence

$$p(UV) = \frac{e^{-U-V}}{\Gamma(a)\Gamma(b)\Gamma(c)} \int_0^{\min U,V} A^{a-1}(U - A)^{b-1}(V - A)^{c-1}e^{A}\, dA. \tag{39}$$

In order to display the properties of the surface, we may also build it up in the following way. The cumulants of A are $\kappa_r = a(r - 1)!$ and similarly for B and C. The center of the distribution is at $(a + b, a + c)$, the correlation between U and V is

$$\rho = \frac{a}{[(a + b)(a + c)]^{1/2}} \tag{40}$$

and

$$p(U) = \frac{U^{a+b-1}e^{-U}}{\Gamma(a + b)}. \tag{41}$$

For the regression we consider the moments of V for U fixed and vice versa. We have

$$p(AB) = \frac{1}{\Gamma(a)\Gamma(b)} A^{a-1}B^{b-1}e^{-(A+B)} \tag{42}$$

so that if we write

$$g = \frac{A}{A + B} = \frac{A}{U}, \qquad G = A + B = U, \tag{43}$$

then, integrating out G,

$$p(g) = p\left(\frac{A}{U}\right) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} g^{a-1}(1 - g)^{b-1}. \tag{44}$$

Accordingly for U fixed

$$\kappa_r(V|U) = \kappa_r(c) + U^r\kappa_r(g), \qquad E(V|U) = c + U\kappa_1(g). \tag{45}$$

The cumulants of C are $c(r-1)! = \kappa_r$ for $r = 1, 2, \cdots$ and the cumulants of g are those of a Type I variable of which the first four are

$$\begin{aligned}
\kappa_1(g) &= \frac{a}{\Delta}, \qquad \Delta = a + b,\\
\kappa_2(g) &= \frac{ab}{\Delta^2(\Delta+1)},\\
\kappa_3(g) &= \frac{2ab(b-a)}{\Delta^3(\Delta+1)(\Delta+2)},\\
\kappa_4(g) &= \frac{3ab}{\Delta^2(\Delta+1)(\Delta+2)(\Delta+3)}\left[\frac{(b-a)^2}{\Delta^2} - \frac{ab(\Delta+2)}{\Delta^2(\Delta+1)}\right].
\end{aligned} \tag{46}$$

The regression of V on U, that is,

$$E(V|U) = c + U\frac{a}{a+b}, \tag{47}$$

is linear no matter what a, b, and c and thus by symmetry so is the regression of U on V. We thus have a distribution with Type III margins and linear regressions which extends over the whole plane $0 < V, U < +\infty$.

Since C and A are assumed independent,

$$p(AC|U) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)\Gamma(c)}\left(\frac{A}{U}\right)^{A-1}\left(1-\frac{A}{U}\right)^{b-1}C^{c-1}\frac{e^{-C}}{U} \tag{48}$$

or

$$p(AV|U) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)\Gamma(c)}\left(\frac{A}{U}\right)^{a-1}\left(1-\frac{A}{U}\right)^{b-1}(V-A)^{c-1}\frac{e^{-V+A}}{U} \tag{49}$$

or

(50)

$$p(V|U) = \frac{\Gamma(a+b)e^{-V}}{\Gamma(a)\Gamma(b)\Gamma(c)U^{a+b-1}}\int_0^{\min U,V} A^{a-1}(U-A)^{b-1}(V-A)^{c-1}e^A\,dA,$$

while

$$p(UV) = \frac{e^{-(U+V)}}{\Gamma(a)\Gamma(b)\Gamma(c)}\int_0^{\min U,V} A^{a-1}(U-A)^{b-1}(V-A)^{c-1}e^A\,dA. \tag{51}$$

In general, it does not seem possible to carry out an integration for A which will result in a distribution which will itself be integrable. The conditions of the original problem, however, required the margins to be skew and this implies that a, b, and c should not be large in order to meet these required conditions. We have evaluated some special cases.

(i) $a = b = c = 1$ ($\rho = 1/2$, center at 2, 2),

$$
(52) \quad
\begin{aligned}
p(VU) &= e^{-U}(1 - e^{-V}), && V < U,\\
p(VU) &= e^{-V}(1 - e^{-U}), && V > U,\\
p(V|U) &= \frac{1}{U}(1 - e^{-V}), && V < U,\\
p(V|U) &= \frac{e^{-V}}{U}(e^{U} - 1), && V > U.
\end{aligned}
$$

(ii) $a = b = 1, c = 2$ ($\rho = 1/\sqrt{6}$, center at 2, 3),

$$
(53) \quad
\begin{aligned}
p(VU) &= e^{-U}[1 - (V + 1)e^{-V}], && V < U,\\
p(VU) &= e^{-V}[(V - U + 1) - (V + 1)e^{-U}], && V > U,\\
p(V|U) &= \frac{1}{U}[1 - (V + 1)e^{-V}], && V < U,\\
p(V|U) &= \frac{e^{-U}}{U}[(V - U + 1)e^{U} - (V + 1)], && V > U.
\end{aligned}
$$

(iii) $a = b = 1, c = 3$ ($\rho = 1/2\sqrt{2}$, center at 2, 4)

$$
(54) \quad
\begin{aligned}
p(VU) &= e^{-U}\left\{1 - \frac{1}{2}e^{-V}[(V + 1)^2 + 1]\right\}, && V < U,\\
p(VU) &= \frac{e^{-V}}{2}\{[(V - U + 1)^2 + 1] - e^{-U}[(V + 1)^2 + 1]\}, && V > U,\\
p(V|U) &= \frac{1}{2U}\{2 - e^{-V}[(V + 1)^2 + 1]\}, && V < U,\\
p(V|U) &= \frac{e^{-V}}{2U}\{e^{U}[(V - U + 1)^2 + 1] - [(V + 1)^2 + 1]\}, && V > U.
\end{aligned}
$$

(iv) $a = 1, b = 2, c = 2$ ($\rho = 1/3$, center at 3, 3),

$$
(55) \quad
\begin{aligned}
p(VU) &= e^{-U}\{2 + (U - V) - e^{-V}[UV + (U + V) + 2]\}, && V < U,\\
p(VU) &= e^{-V}\{2 + (V - U) - e^{-U}[UV + (U + V) + 2]\}, && V > U,\\
p(V|U) &= \frac{2}{U^2}\{2 + (U - V) - e^{-V}[UV + (U + V) + 2]\}, && V < U,\\
p(V|U) &= \frac{2e^{-V}}{U^2}\{e^{U}[2 + (V - U)] - [UV + (U + V) + 2]\}, && V > U.
\end{aligned}
$$

(v) $a = 2 = b = c$ ($\rho = 1/2$, center at 4, 4),

$$
\begin{aligned}
p(VU) &= e^{-U}\{[V(U - V) + 2(2V - U) - 6]\\
&\qquad + e^{-V}[UV + 2(U + V) + 6]\}, && V < U,\\
p(VU) &= e^{-V}\{[U(U - V) + 2(2U - V) - 6]\\
&\qquad + e^{-U}[UV + 2(U + V) + 6]\}, && V > U,
\end{aligned}
$$

(56)

$$p(V|U) = \frac{6}{U^3}\{[V(U - V) + 2(2V - U) - 6] + e^{-V}[UV + 2(U + V) + 6]\}, \qquad V < U,$$

$$p(V|U) = \frac{6e^{-V}}{U^3}\{e^{U}[U(U - V) + 2(2U - V) - 6] + [UV + 2(U + V) + 6]\}, \qquad V > U.$$

7. Expectation of Spearman's rho for the double Gamma distribution

For those Gamma distributions which are integrable, the $E(r_s)$ is easily obtained. We give some values of P_1 and P_2 in table III for three such distributions from which one can calculate

$$E(r_s) = 12[(n - 2)P_1 + P_2 - 0.25(n - 1)]/(n + 1)$$

TABLE III

VALUES OF P_1 AND P_2 FOR THREE GAMMA SURFACES

	$\rho = 1/2$ $a = b = c = 1$	$\rho = 1/\sqrt{6}$ $a = b = 1; c = 2$	$\rho = 1/3$ $a = 1; b = c = 2$
$P_1 = P\{x_i > x_j, y_i > y_m \mid j \neq m\}$	125/432	487/1728	5723/20736
$P_2 = P\{x_i > x_j, y_i > y_j\}$	1/3	91/288	131/432

as given in section 5. The result $P_2 = 1/3$ will be true for all Gamma surfaces for which $\rho = 1/2$. This follows from the definition

(57) $$P_2 = P\{x_i > x_j, y_i > y_j\} = P\{(A_i - A_j) + (B_i - B_j) > 0, (A_i - A_j) + (C_i - C_j) > 0\},$$

and the assumption that A, B, and C are all independent and have identical distributions. We may, therefore, define a new set of variables

(58) $$z_1 = A_i - A_j, \qquad -z_2 = B_i - B_j, \qquad -z_3 = C_i - C_j,$$

whence

(59) $$P_2 = P\{z_1 > z_2, z_1 > z_3\} = P\{z_1 > z_2 \text{ and } z_3\} = \frac{1}{3}.$$

This is also the value of P_2 for the normal surface when $\rho = 0.5$. For purposes of illustration we give some values of $E(r_s)$ for different n and ρ in table IV. The values for the normal surface (from Moran's formula) are also given for comparison. It is clear that for these surfaces $E(r_s)$ is nearly the same function of ρ as for the normal case. Since this is the quantity which enters in the over-all variance of T_1, θ, and φ, it would seem unlikely that these variances would be greatly affected by differences from the bivariate normal of the type envisaged. It appears likely, although further research is necessary to establish this point,

TABLE IV

VALUES OF $E(r_s)$ FOR $n = 5, 10, 20, 50$ AND $\rho = 1/2, 1/\sqrt{6}, 1/3$

n	$\rho = 1/2$ Gamma $a = b = c = 1$	$\rho = 1/2$ Normal	$\rho = 1/\sqrt{6}$ Gamma $a = b = 1; c = 2$	$\rho = 1/\sqrt{6}$ Normal	$\rho = 1/3$ Gamma $a = 1; b = c = 2$	$\rho = 1/3$ Normal
5	0.4028	0.4080	0.3229	0.3301	0.2624	0.2681
10	0.4343	0.4419	0.3497	0.3585	0.2849	0.2916
20	0.4524	0.4613	0.3651	0.3747	0.2978	0.3051
50	0.4641	0.4739	0.3750	0.3851	0.3061	0.3138
∞	0.4722	0.4826	0.3819	0.3925	0.3119	0.3199

that, as with the ordinary product moment correlation coefficient, linearity of regression is a crucial factor. Provided this is retained, the surface may then be distorted from the bivariate normal to a remarkable extent without unduly upsetting the expected value.

To study the over-all variance of T_2 it will be necessary to investigate $E(r_s^2)$ for both the normal and the double Gamma distributions. At present this is only known to order $1/n$ for the normal case and not at all for the double Gamma distribution.

8. Mean of T_1, θ, and φ under the alternative hypothesis

The over-all tests using T_1, θ, and φ may be carried out by assuming that the criteria are normally distributed with known variance. Since each criterion may be looked on as the mean of a sample of n_1 drawn from a finite population of n, the assumption of normality will be justifiable for reasonable n and n_1 not too different from $n/2$. The variance of the criteria in each case will depend on $E(r_s)$ which itself depends on the correlation in the population. If this is not known in practice, possibly the best thing to do is to substitute r_s for its expected value, which will mean that again n must be reasonably large.

For the mean value of the criteria under alternate hypotheses of the double Gamma type, it is enough to consider the marginal distributions. Under the null hypothesis, let f_1 be the p.d.f. of the variable U and

(60) $$F_1 = \int_h^U f_1\, dU,$$

where h is conventional for the start of the distribution. Let f_2 and F_2 have similar meanings under the alternate hypothesis. Since only a margin is considered, we have a Wilcoxon situation and it is well known that for the jth rank

(61) $$p(j) = \int_h^\infty \sum_r F_1^r(1 - F_1)^{n_1-r}F_2^{j-r-1}(1 - F_2)^{n_2-j+r}\, dF_2\binom{n_1}{r}\binom{n_2 - 1}{j - r - 1},$$

whence

$$E(j) = \int_h^\infty \{F_2(n_2 - 1) + n_1 F_1 + 1\}\, dF_2. \tag{62}$$

We take first a positive change in the mean of U and of V, the correlation in the surface remaining unchanged.

(i) $f_1 = e^{-U}$, $\quad F_1 = 1 - e^{-U}$; $\quad f_2 = Ue^{-U}$, $\quad F_2 = 1 - Ue^{-U} - e^{-U}$;

$$h = 0, \qquad \rho = 0.5, \qquad E(j) = \frac{1}{4} n_1 + \frac{n+1}{2}.$$

In order to keep the correlation unchanged, a similar set of hypotheses is specified for V and

$$\begin{aligned} E(T_1) &= \frac{n_2}{4\sigma}\,[1 - E(r_s)], \\ E(\varphi) &= 0, \\ E(\theta) &= \frac{n_2}{2\sigma}. \end{aligned} \tag{63}$$

To a first approximation, we shall suppose the variance of each criterion under H_1 is the same as the variance under H_0 and that for T_1 we have $E(r_s^2)$ the same as for the normal surface. This will mean that the variance is underestimated in each case with a consequent magnification of the power of the test. Table V shows the approximate power for two of the criteria, namely T_1 and θ. In spite of the fact that the powers calculated are almost certainly too large, two points emerge. The first point is that θ is superior to T_1 in detecting change in both $E(U)$ and $E(V)$. This is as expected since θ was designed to be sensitive to precisely this kind of change, whereas T_1 is really sensitive only to a change in $E(U)$. The second point, which is a little unexpected, is that both tests appear very insensitive to changes in the ratio of $n_1 : n_2$, although the maximum power is achieved when $n_1 = n_2$.

TABLE V

APPROXIMATE POWER OF THE CRITERIA T_1 AND θ TO DETECT CHANGES OF UNITY IN $E(U)$ AND $E(V)$ IN THE DOUBLE GAMMA SURFACE. $\rho = 0.5$.
Probability of first kind of error = 0.05. One-tailed test

			Power	
n	n_1	n_2	T_1	θ
10	5	5	0.21	0.46
	4	6	0.21	0.45
20	10	10	0.32	0.72
	8	12	0.31	0.70
50	25	25	0.58	0.97
	20	30	0.57	0.97

To set up a realistic situation in which the φ criterion might be expected to be the most powerful is difficult for the double Gamma surfaces proposed. Under the null hypothesis, the surfaces range from zero to $+\infty$. Under the alternate hypothesis, it will be supposed $\alpha \leqq U \leqq +\infty$ and $-\beta \leqq V \leqq +\infty$. Thus we have

(ii) $f_1 = e^{-U}, \quad F_1 = 1 - e^{-U}; \quad f_2 = e^{-U+\alpha}, \quad F_2 = 1 - e^{-U+\alpha}$

and

$$E(j) = \frac{n+1}{2} + \frac{n_2}{2}(1 - e^{-\alpha}) \tag{64}$$

for the U margin. For the V margin

$$E(j) = \frac{n+1}{2} + \frac{n_2}{2}(1 - e^{\beta}) \tag{65}$$

so that

$$\begin{aligned} E(T_1) &= \frac{n_2}{2\sigma}\{1 - e^{-\alpha} - r_s(1 - e^{\beta})\}, \\ E(\varphi) &= \frac{n_2}{2\sigma}\{e^{\beta} - e^{-\alpha}\}, \\ E(\theta) &= \frac{n_2}{\sigma}\left\{1 - \frac{1}{2}(e^{\beta} + e^{-\alpha})\right\}. \end{aligned} \tag{66}$$

Calculations show that the approximate power of these test criteria is largest for φ and smallest for θ for α and β positive.

9. Other tests

Apart from the obvious method of transforming the data, if we move away from the idea of a rank test criterion, there are various other tests which might be applied. Perhaps the most pertinent of these is a slight variant of a test due to Mood, but it is only one of many. Consider the original observations $\{x_i, y_i\}$ of the combined samples $n_1 + n_2 = n$ and order the x in magnitude, making a dichotomy at the median point x_M. Assuming linearity of regression, let the regression lines of the first and second samples be

$$\begin{aligned} Y_1 &= \alpha_1 + \beta x, \\ Y_2 &= \alpha_2 + \beta x. \end{aligned} \tag{67}$$

The hypothesis to be tested is $\alpha_1 = \alpha_2 = \alpha$, unspecified. For the combined data, a line is chosen so that the median of the deviations from this line in the $n/2$ observations [$(n - 1)/2$, if n is odd] to the left of x_M is zero and similarly for the $n/2$ observations to the right of x_M. Counting the observations in the sample of n_1 above and below the line to the left of x_M and above the line to the right, a 2×2 table may be formed in each cell of which the expected frequency is $n_1/4$. Since this test is not dependent on the functional form of the bivariate distribution of the population from which $n_1 + n_2 = n$ observations have been drawn, there seems to be no advantage in applying the same idea to the ranks of the x and the y. The only difficulty in application will lie in the choice of the line. This can be formidable if the number of observations is large.

From recent work on the power function of test criteria based on ranked and ordered variables, it would appear possible that tests more powerful than those discussed here could be obtained if, instead of the ranks, the equivalent normal deviate of the rank is used. It is proposed to discuss tests based on the equivalent normal deviate in a subsequent paper after further investigation of the power functions of the tests proposed here.

We would like to acknowledge stimulating criticism from colleagues, in particular from C. L. Mallows. Barbara Snow constructed the randomization distributions.

10. Numerical appendix

As an illustration of the type of data which originally suggested the problem we give here the precipitation in inches for two areas in southern California for the two years 1957 and 1958. The experiments to determine the efficacy of cloud seeding operations of which this is part of the numerical data have been described elsewhere, and also the statistical analysis used. Table VI shows the precipitation for one target area and one comparison area, seeded and not seeded. The years 1957 and 1958 are kept separate because in 1958 seeding operations were

TABLE VI

PRECIPITATION IN INCHES

1957				1958			
(Not Seeded in Ventura)				(Seeded in Ventura)			
Seeded in S.B.		Not Seeded in S.B.		Seeded in S.B.		Not Seeded in S.B.	
Target	Comparison	Target	Comparison	Target	Comparison	Target	Comparison
1.035	0.500	0.212	0.120	0.775	0.730	0.325	0.060
0.000	0.000	0.265	0.220	1.232	0.250	1.635	1.400
0.198	0.190	0.100	0.110	0.000	0.000	1.128	0.690
0.235	0.470	0.152	0.090	1.428	0.090	0.335	0.140
0.005	0.000	0.180	0.100	0.558	0.240	0.785	0.030
0.445	0.140	0.015	0.000	2.740	2.990	0.482	0.000
0.312	0.100	1.682	1.500	0.010	0.000	0.258	0.040
0.070	0.000	0.002	0.000	0.055	0.030	0.972	2.510
0.148	0.060	0.688	0.560	0.948	0.010	0.785	2.210
0.000	0.000	0.072	0.000	0.358	0.020	0.162	0.040
0.062	0.430	0.302	0.440	0.320	0.150	0.772	0.860
1.728	0.440	0.008	0.060	0.142	0.060	0.022	0.000
0.008	0.000					2.655	0.430
						0.358	0.290

TABLE VII

RANKED PRECIPITATIONS

1957				1958			
(Not Seeded in Ventura)				(Seeded in Ventura)			
Seeded in S.B.		Not Seeded in S.B.		Seeded in S.B.		Not Seeded in S.B.	
Target	Comparison	Target	Comparison	Target	Comparison	Target	Comparison
23	23	16	15	16	21	9	11.5
1.5	4.5	18	18	22	17	24	23
15	17	11	14	1	2.5	21	20
17	22	13	11	23	13	10	14
4	4.5	14	12.5	14	16	17.5	7.5
21	16	7	4.5	26	26	13	2.5
20	12.5	24	25	2	2.5	7	9.5
9	4.5	3	4.5	4	7.5	20	25
12	9.5	22	24	19	5	17.5	24
1.5	4.5	10	4.5	11.5	6	6	9.5
8	19	19	20.5	8	15	15	22
25	20.5	5.5	9.5	5	11.5	3	2.5
5.5	4.5					25	19
						11.5	18

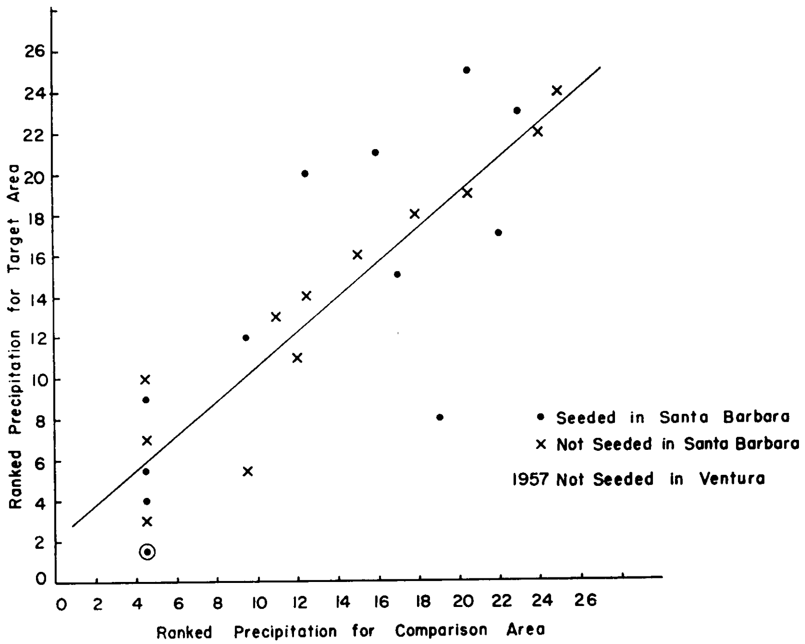


FIGURE 3(a)

Ranked precipitation for target and comparison areas.
1957—Not seeded in Ventura.

also taking place in other areas. For each year the target precipitations (seeded and not seeded) are ranked and similarly the comparison precipitations. The ties were ranked by giving each observation in the tie the midrank. Subsequent calculations showed it to be immaterial whether this procedure was adopted or

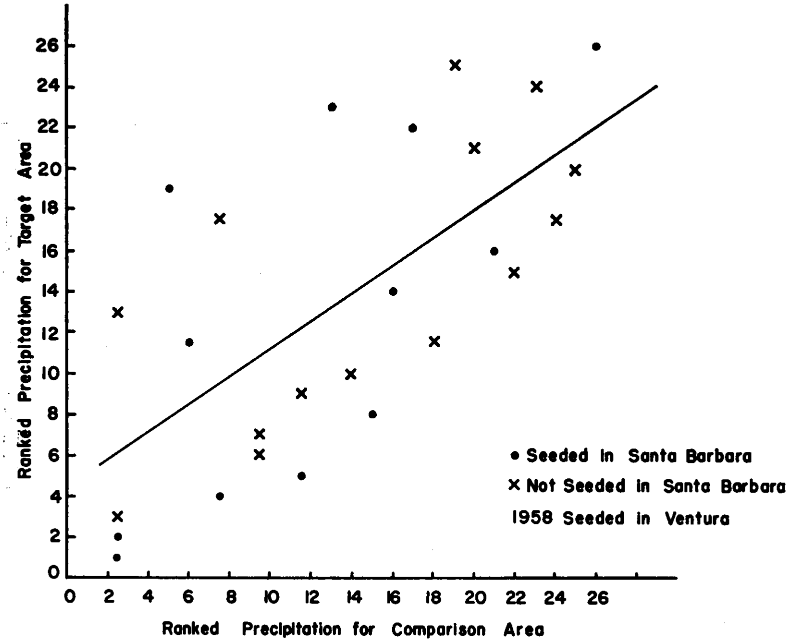


FIGURE 3(b)

Ranked precipitation for target and comparison areas.
1958—Seeded in Ventura.

the procedure of assigning a random order to the observations within the tie. The ranked observations are given in table VII. These results are shown graphically in figures 3(a) and 3(b). The rank regression lines of target area (y) on comparison area (x) are

$$1957: \quad y - 13 = 0.8532\,(x - 13),$$

$$1958: \quad y - 13.5 = 0.6799\,(x - 13.5).$$

The null hypothesis is that there is no effect due to seeding so that one regression line represents the true state of affairs. The alternate hypothesis is that seeding is effective so that there are two regression lines, one for seeded and one for nonseeded, with the seeded line lying parallel to but above that of the nonseeded. It is clear that either T_1^* or φ will be the appropriate criterion to use.

If R_{y_i} and R_{x_i} are the ranks of the target and comparison precipitation for the seeded area, then

$$T_1^* = \frac{1}{n_1\sigma} \sum_{n_1} \left[\left(R_{y_i} - \frac{n+1}{2}\right) - r_s\left(R_{x_i} - \frac{n+1}{2}\right)\right],$$

(68)
$$\varphi = \frac{1}{n_1\sigma} \sum_{n_1} (R_{y_i} - R_{x_i}),$$

$$\sigma^2_{T_1^*} = \frac{(n - n_1)}{n_1(n-1)}(1 - r_s^2), \qquad \sigma^2_\varphi = \frac{2(n - n_1)}{n_1(n-1)}(1 - r_s).$$

Substitution from the tabulated ranks gives the results shown in table VIII. The conclusion is that the tests have failed to detect increase of precipitation due to seeding operations.

TABLE VIII

Year	T_1^*	$\sigma^2_{T_1^*}$	φ	σ_φ	$T_1^*/\sigma^2_{T_1^*}$	φ/σ_φ
1957	−0.0056	0.1023	0.0053	0.1063	−0.055	0.050
1958	0.0269	0.1584	0.094	0.1728	0.170	0.546

REFERENCES

[1] E. C. Fieller, H. O. Hartley, and E. S. Pearson, "Tests for rank correlation coefficients. I," *Biometrika*, Vol. 44 (1957), pp. 470–481.

[2] M. G. Kendall, "Rank and product moment correlation," *Biometrika*, Vol. 36 (1949), pp. 177–193.

[3] P. A. P. Moran, "Rank correlation and product moment correlation," *Biometrika*, Vol. 35 (1948), pp. 203–206.

[4] E. C. Rhodes, "On a certain skew correlation surface," *Biometrika*, Vol. 14 (1922), pp. 355–377.

[5] E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Vol. 1, Cambridge, Cambridge University Press, 1954.

[6] R. M. Sundrum, "A method of systematic sampling based on order properties," *Biometrika*, Vol. 40 (1953), pp. 452–456.

[7] M. J. Van Uven, "Extension of Pearson's probability distributions to two variables, I, II, III, IV," *Nederl. Akad. Wetensch. Proc., Ser. A*, Vol. 50 (1947), pp. 1063–1070, 1252–1264. Vol. 51 (1948), pp. 41–52, 191–196.