SOME THEOREMS ON QUADRATIC FORMS APPLIED IN THE STUDY OF ANALYSIS OF VARIANCE PROBLEMS, II. EFFECTS OF INEQUALITY OF VARIANCE AND OF CORRELATION BETWEEN ERRORS IN THE TWO-WAY CLASSIFICATION

By G. E. P. Box

Imperial Chemical Industries, Manchester, England and North Carolina State College

1. Summary and Introduction. Theorems already enunciated in a previous paper on quadratic forms are used to determine the effects of inequality of variance and first order serial correlation of errors in the two-way classification on the analysis of variance. It is found that when the appropriate null hypothesis is true, inequality of variance from column to column results in an increased chance of exceeding the significance point for the test on homogeneity of column means, and a decreased chance for the corresponding test on row means. For moderate differences in variance neither effect is large. First order serial correlation within rows produces a large effect on the "between rows" comparisons, but little effect on the "between columns" comparisons.

2. The two-way analysis of variance classification. Consider the analysis of variance for a two-way table with k columns and n rows, with one observation in each cell. Experiments in which k treatments are tested in n blocks are an important source of data classified in this way. In such tables the variance might change from treatment to treatment due to the influence of the treatments themselves. Changes in variance might also occur from block to block, for in some circumstances where experimental material was inhomogeneous in mean from block to block it might well be inhomogeneous in variance also.

A further source of departure from the assumptions usually made in the analysis of variance concerns possible lack of independence between the "error" components of the observations. In many types of experiments this difficulty is met by the introduction of randomisation. Data occur, however, in circumstances where there is no possibility of using this device, usually because the factor which is to be studied is the effect of time or position, which itself gives rise to the correlation.

For instance, the first example of analysis of variance of a two-way table in R. A. Fisher's *Statistical Methods for Research Workers* [1] concerns data quoted from Shaw [2] on the frequency of rainfall classified by hour of the day and month of the year. As Fisher himself points out, strong serial correlation between errors within months occurs because showers of rain which last more than one hour are recorded in successive hours. No question of randomisation arises in

---

Received 6/15/53.

this example. In discussing the analysis of variance table, Fisher remarks that the serial correlation between hours within months entirely invalidates the "between months" comparisons, but that the "between hours" comparisons may still be made (as an approximate test). The truth of the latter part of this statement is perhaps not immediately obvious, and it is of interest to make a closer study of such examples.

Other instances of two-way tables in which serial correlation between errors might be expected are quoted by Daniels [3] in experiments in wool research where, for example, the variation in weight of slubbing coming from adjacent positions on the wool card is considered. Daniels recognised that correlation effects might invalidate the analysis of variance procedure and carried out some theoretical investigation of the problem [4]. He considered the effects of small inequalities of variance and small correlations between errors, using an approximate method. Tests for the existence of departures from assumptions in the two-way table were discussed by Box in 1950 [5], when reference was given to the results now published.

In what follows we retain the assumption of normality, but allow the variance to differ from column to column and correlation to occur within rows. By substituting columns for rows we can also study the effect of differences in variance from row to row and the effect of correlation within columns.

3. Distribution of items in the analysis of variance table. We need to refer to theorems, equations and sections of a previous paper [6] with the same general title. We indicate such reference by the addition of a prime to the number of the theorem, etc. Thus Theorem 2.1′ and Section 5′ refer to Theorem 2.1 and Section 5, respectively, of the previous paper.

Suppose we have a two-way classification of observations with k columns and n rows and y_{ti} is the observation in the t^{th} column and i^{th} row. Then we can perform an analysis of variance corresponding to the entries in the first three columns of Table 1. We make the usual assumptions that y_{ti} may be represented by a linear model

$$(3.1) \qquad y_{ti} = \alpha + \beta_i + \gamma_t + z_{ti}, \qquad \sum_{i=1}^{n} \beta_i = 0; \qquad \sum_{t=1}^{k} \gamma_t = 0.$$

Alternatively we can denote the model for all the elements of the t^{th} column of the table ($t = 1, 2, \cdots, k$) by

$$(3.2) \qquad \mathbf{y}_{t.} = \alpha \mathbf{1}_n + \boldsymbol{\beta} + \gamma_t \mathbf{1}_n + \mathbf{z}_{t.}$$

where $\mathbf{y}_{t.}$ is the $n \times 1$ vector of entries in the t^{th} column, $\mathbf{z}_{t.}$ is the corresponding vector of errors, $\mathbf{1}_n$ is an $n \times 1$ vector of unit elements, and $\boldsymbol{\beta}$ an $n \times 1$ vector of row constants $\beta_1, \beta_2, \cdots, \beta_n$. We shall also need the notation $\mathbf{y}_{.i}$ $\mathbf{z}_{.i}$ to denote $k \times 1$ vectors of observations and errors in the i^{th} row of the table.

We do not make the usual assumption that the z_{ti} have the same variance and are uncorrelated. Instead we assume that $\mathbf{z}_{.i}$ follows the normal multivariate

TABLE 1

Analysis of variance for a two-way table with column variances unequal and correlation of errors within rows

Source	D/F	Sums of Squares, Q	*Expectation of Q	*Null distribution of Q
Rows	$n - 1$	$Q_R = k \sum_{i=1}^{n} (\bar{y}_{.i} - \bar{y}_{..})^2$	$k \sum_{i=1}^{n} \beta_i^2 + (n-1)(\bar{v}_{tt} + (k-1)\bar{v}_{ts})$	† $(\bar{v}_{tt} + (k-1)\bar{v}_{ts})\chi^2(n-1)$
			$k \sum_{i=1}^{n} \beta_i^2 + (n-1)\bar{\sigma}^2$	† $\bar{\sigma}^2\chi^2(n-1)$
Columns	$k - 1$	$Q_C = n \sum_{t=1}^{k} (\bar{y}_{t.} - \bar{y}_{..})^2$	$n \sum_{t=1}^{k} \gamma_t^2 + (k-1)(\bar{v}_{tt} - \bar{v}_{ts})$	‡ $\sum_{j=1}^{k-1} \lambda_j \chi^2(1)$
			$n \sum_{t=1}^{k} \gamma_t^2 + (k-1)\bar{\sigma}^2$	
Residual (Error)	$(n-1) \times (k-1)$	$Q_E = \sum_{t=1}^{k} \sum_{i=1}^{n} (y_{ti} - \bar{y}_{t.} - \bar{y}_{.i} + \bar{y}_{..})^2$	$(n-1)(k-1)(\bar{v}_{tt} - \bar{v}_{ts})$	$\sum_{j=1}^{k-1} \lambda_j \chi^2(n-1)$
			$(n-1)(k-1)\bar{\sigma}^2$	

* The upper expressions hold in the general case, the lower ones when the correlations are zero, that is, when only differences in variance from column to column occur. In this case $\{u_{ts}\} = \{(\delta_{ts} - k^{-1})\sigma_t^2\}$. In both cases, $\bar{v}_{tt} = \bar{\sigma}^2$ is the average variance $\Sigma_t v_{tt}/k$, and $\bar{v}_{ts}$ is the average covariance $\Sigma \Sigma_{t \neq s} v_{ts}/\{k(k-1)\}$, while $\lambda_1, \lambda_2, \cdots, \lambda_{k-1}$ are the $k - 1$ nonzero latent roots of $\{u_{ts}\} = \{v_{ts} - \bar{v}_{t.}\}$, where $\bar{v}_{t.}$ is the average value of the elements in the t^{th} row or column of v.

† Distribution not independent of Q_E.

‡ Distribution independent of Q_E.

law with variance-covariance matrix $\mathcal{E}(\mathbf{z}_{.i}\mathbf{z}'_{.i}) = \mathbf{v} = \{v_{ts}\}$. We further assume that $\mathbf{z}_{.j}$ ($j = 1, 2, \cdots, i-1, i+1, \cdots, n$) follows the same law independently of $\mathbf{z}_{.i}$. Thus $v_{11}, \cdots, v_{tt}, \cdots, v_{kk}$ are the k variances and $v_{12}, v_{13}, \cdots, v_{ts}, \cdots, v_{k-1\,k}$ the $\frac{1}{2}k(k-1)$ covariances, the same for every row. This enables us to study the effects of column to column heterogeneity of variance and/or "within rows" correlation of errors. The expected values and null distributions of the sums of squares, when the observations are so represented, are shown in Table 1. They are derived below.

Let $\mathbf{Y}_{t.}$ be an $n \times 1$ vector of elements $Y_{t1}, Y_{t2}, \cdots, Y_{tn}$ obtained from $\mathbf{y}_{t.}$ by orthogonal transformation $\mathbf{Y}_{t.} = \mathbf{p}\mathbf{y}_{t.}$, and let the $n \times n$ orthogonal matrix $\mathbf{p}$ have all the elements of its last row equal to $n^{-1/2}$, thus ensuring that $Y_{tn} = n^{1/2}\bar{y}_{t.}$. Then

$$\mathbf{Y}_{t.} = \mathbf{p}\mathbf{y}_{t.} = \alpha\boldsymbol{\delta} + \mathbf{B} + \gamma_t\boldsymbol{\delta} + \mathbf{Z}_{t.} \tag{3.3}$$

where $\boldsymbol{\delta} = \mathbf{p}\mathbf{1}_n$, $\mathbf{B} = \mathbf{p}\boldsymbol{\beta}$ and $\mathbf{Z}_{t.} = \mathbf{p}\mathbf{z}_{t.}$.

Due to the nature of $\mathbf{p}$, in the vector $\boldsymbol{\delta}$ the last element is $n^{1/2}$ and the remaining elements are zeros, and in $\mathbf{B}$ the last element is zero, since $\sum_1^n \beta_i = 0$. The transformed columns of the original two-way table and the transformed column of row means may now be written out as follows:

						Row Means
$B_1 + Z_{11}$	$\cdots$	$B_1 + Z_{t1}$	$\cdots$	$B_1 + Z_{k1}$		$B_1 + \bar{Z}_{.1}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		$\cdots$
$B_i + Z_{1i}$	$\cdots$	$B_i + Z_{ti}$	$\cdots$	$B_i + Z_{ki}$		$B_i + \bar{Z}_{.i}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		$\cdots$
$B_{n-1} + Z_{1n-1}$	$\cdots$	$B_{n-1} + Z_{tn-1}$	$\cdots$	$B_n + Z_{kn-1}$		$B_{n-1} + \bar{Z}_{.n-1}$
$n^{1/2}(\alpha + \gamma_1) + Z_{1n}$	$\cdots$	$n^{1/2}(\alpha + \gamma_t) + Z_{tn}$	$\cdots$	$n^{1/2}(\alpha + \gamma_k) + Z_{kn}$		$n^{1/2}\alpha + \bar{Z}_{.n}$

Now consider the $nk \times 1$ partitioned vector $\mathbf{z}$ and the $nk \times nk$ partitioned matrices $\mathbf{P}$ and $\mathbf{V}$ defined by

$$\mathbf{z}' = [\mathbf{z}'_{1.} \mid \mathbf{z}'_{2.} \mid \cdots \mid \mathbf{z}'_{t.} \mid \cdots \mid \mathbf{z}'_{k.}]$$

$$\mathbf{P} = \begin{bmatrix} \mathbf{p} & \mathbf{0}_n & \cdots & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{p} & \cdots & \mathbf{0}_n \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0}_n & \mathbf{0}_n & \cdots & \mathbf{p} \end{bmatrix} \qquad \mathbf{V} = \begin{bmatrix} v_{11}\mathbf{I}_n & v_{12}\mathbf{I}_n & \cdots & v_{1k}\mathbf{I}_n \\ v_{12}\mathbf{I}_n & v_{22}\mathbf{I}_n & \cdots & v_{2k}\mathbf{I}_n \\ \cdots & \cdots & \cdots & \cdots \\ v_{1k}\mathbf{I}_n & v_{2k}\mathbf{I}_n & \cdots & v_{kk}\mathbf{I}_n \end{bmatrix} \tag{3.4}$$

where $\mathbf{I}_n$ is the $n \times n$ unit matrix and $\mathbf{0}_n$ is the $n \times n$ null matrix. For $\mathbf{Z}$, denoting the fector $\mathbf{Pz}$ of transformed variables, the matrix of variances and covariances is

$$\mathcal{E}(\mathbf{ZZ}') = \mathcal{E}(\mathbf{Pzz}'\mathbf{P}') = \mathbf{P}\mathcal{E}(\mathbf{zz}')\mathbf{P}' = \mathbf{P}'\mathbf{VP} = \mathbf{V}. \tag{3.5}$$

Since we are concerned with normal variates, it follows that the Z_{ti} are distributed in precisely the same manner as are the z_{ti}, that is the vector $\mathbf{Z}_{.i}$ of transformed errors in the i^{th} row follows the normal multivariate law, with variance-covariance matrix $\mathbf{v}$, independently of the errors in the other rows.

Between columns sum of squares.

$$Q_C = n \sum_{t=1}^{k} (\bar{y}_{t.} - \bar{y}_{..})^2 = \sum_{t=1}^{k} (n^{1/2}\gamma_t + Z_{tn} - \bar{Z}_{.n})^2 \tag{3.6}$$

and the matrix of the quadratic form $q_n = \sum_{t=1}^{k} (Z_{tn} - \bar{Z}_{.n})^2$ is

$$\mathbf{m} = \mathbf{I}_k - k^{-1}\mathbf{1}_k\mathbf{1}_k' \tag{3.7}$$

while the variance-covariance matrix for the vector of errors $\mathbf{Z}_{.n}$ is $\mathbf{v}$. We have therefore

$$\mathbf{u} = \{u_{ts}\} = \mathbf{vm} = \{v_{ts} - \bar{v}_{t.}\} \tag{3.8}$$

where u_{ts} is the element of the t^{th} row and s^{th} column of the matrix $\mathbf{u}$ and $\bar{v}_{t.}$ is the arithmetic mean of the entries in the t^{th} row (or column) of $\mathbf{v}$. It follows from equation (2.5′) that the expectation and null distribution of Q_C are those shown in Table 1 where the λ's are the latent roots of $\mathbf{u}$.

Residual (error) sum of squares.

$$\begin{aligned} Q_E &= \sum_{i=1}^{n} \sum_{t=1}^{k} (y_{ti} - \bar{y}_{t.} - \bar{y}_{.i} + \bar{y}_{..})^2 \\ &= \sum_{i=1}^{n-1} \sum_{t=1}^{k} (Y_{ti} - \bar{Y}_{.i})^2 = \sum_{i=1}^{n-1} \sum_{t=1}^{k} (Z_{ti} - \bar{Z}_{.i})^2. \end{aligned} \tag{3.9}$$

Denote $\sum_{t=1}^{k} (Z_{ti} - \bar{Z}_{.i})^2$ by q_i $(i = 1, 2, \cdots, n)$. Then q_i follows the same distribution as q_j $(j = 1, 2, \cdots, n)$ independently of q_j. In particular it follows the same distribution as q_n discussed above. Also, $\sum_{i=1}^{n-1} q_i = Q_E$ is distributed independently of Q_C, in the form indicated in Table 1.

Between rows sum of squares.

$$Q_R = k \sum_{i=1}^{n} (\bar{y}_{.i} - \bar{y}_{..})^2 = k \sum_{i=1}^{n-1} Y_{.i}^2 = k \sum_{i=1}^{n-1} (B_i + \bar{Z}_{.i})^2. \tag{3.10}$$

Remembering that $\sum_{i=1}^{n-1} B_i^2 = \sum_{i=1}^{n} \beta_i^2$ we have

$$\mathcal{E}(Q_R) = k \sum_{i=1}^{n} \beta_i^2 + (n - 1)\{\bar{v}_{tt} + (k - 1)\bar{v}_{ts}\}, \tag{3.11}$$

where $\bar{v}_{tt}$ is the average variance $\sum_t v_{tt}/k$ and $\bar{v}_{tt}$ is the average covariance $\sum\sum_{t\neq s} v_{ts}/\{k(k - 1)\}$. Now $\bar{Z}_{.i}$ is distributed normally and independently of $\bar{Z}_{.j}$ $(i \neq j = 1, 2, \cdots, n)$. Hence, when the null hypothesis that $\sum_1^n \beta_1^2 = 0$ is true, Q_R is distributed like $X_R = \{\bar{v}_{tt} + (k - 1)\bar{v}_{ts}\}\chi^2(n - 1)$.

Since $\mathbf{Z}_{.n}$ is distributed independently of $\mathbf{Z}_{.i}$ $(i = 1, \cdots, n - 1)$, Q_R and Q_C

are distributed independently. Usually Q_R will not be distributed independently of Q_E, however, as will now be shown.

Dependence of Q_R and Q_E. To investigate the dependence of Q_R and Q_E we transform the $k \times 1$ vector $\mathbf{Z}_{.i}$ of the transformed variates in the i^{th} row of the two-way table to the vector $\mathbf{W}_{.i}$ by means of the orthogonal transformation $\mathbf{W}_{.i} = \mathbf{RZ}_{.i}$, where the elements of the last (k^{th}) row of $\mathbf{R} = \{r_{ts}\}$ are all equal to $k^{1/2}$ so that $W_{ki} = k^{1/2}\bar{Z}_{.i}$. The variance-covariance matrix for the new variates is now given by

$$\mathcal{E}(\mathbf{W}_{.i}\mathbf{W}_{.i}') = \mathcal{E}(\mathbf{RZ}_{.i}\mathbf{Z}_{.i}'\mathbf{R}') = \mathbf{R}\mathcal{E}(\mathbf{Z}_{.i}\mathbf{Z}_{.i}')\mathbf{R}' = \mathbf{RvR}' \tag{3.12}$$

and therefore $\mathcal{E}(W_{ti}W_{ki}) = k^{1/2} \sum_{s=1}^{k} \bar{v}_{s.}r_{ts}$. Now $\sum_{s=1}^{k} r_{ts} = 0$ for $t = 1, 2, \cdots k - 1$. The covariances between W_{ki} and $W_{1i}, W_{2i}, \cdots, W_{k-1i}$ cannot therefore *all* be zero unless $\bar{v}_{s.}$, the mean of entries in the s^{th} row or column of $\mathbf{v}$, is constant for all s, since in a k-space only one vector can be simultaneously at right angles to $k - 1$ other linearly independent vectors. In particular the condition that $\bar{v}_{s.}$ is constant for all s is satisfied when the observations are independent and the variances are equal (when $\mathbf{v} = \sigma^2\mathbf{I}_k$) and also when the observations are circularly correlated. This condition usually will not be satisfied, however. In particular it will not be satisfied when the observations are independent but the variances are unequal, or when the variances are equal and the observations are serially but not circularly correlated.

If W_{ki} is not distributed independently of W_{ti} ($t = 1, 2, \cdots, k - 1$), then W_{ki}^2 will not be distributed independently of $\sum_{t=1}^{k-1} W_{ti}^2$ and $Q_R = \sum_{i=1}^{n-1} W_{ki}^2$ will not be distributed independently of $Q_E = \sum_{i=1}^{n-1} \sum_{t=1}^{k-1} W_{ti}^2$.

4. Distribution of test criteria.

Between columns test. When the appropriate null-hypothesis is true, the ratio of mean squares $(n - 1)Q_C/Q_E$ is distributed like

$$X_C/X_E = \left\{\sum_{j=1}^{k-1} \lambda_j \chi^2(1)\right\} \Big/ \left\{\sum_{j=1}^{k-1} \lambda_j \chi^2(n - 1)\right\}, \tag{4.1}$$

where the λ's, which are the same for both numerator and denominator, are the $k - 1$ nonzero latent roots of the matrix $\mathbf{u} = \{v_{ts} - \bar{v}_{t.}\}$, and the numerator and denominator are distributed independently. We may use the exact series of Theorem 4.1′ to find the value of Pr $(Q_C/Q_E > Y_0)$ and so provide a check on the F approximation, provided we choose examples in which n is odd so that $n - 1$ is even.

To use the approximation of Theorem 6.1′ we require the first two cumulants of Q_C and Q_E when the null hypothesis is true. Using equations (2.5′) and (2.6′) we have

$$K_1(Q_C) = k(\bar{v}_{tt} - \bar{v}_{..}) = (k - 1)(\bar{v}_{tt} - \bar{v}_{ts}) \tag{4.2}$$

$$K_2(Q_C) = 2 \sum_{t=1}^{k} \sum_{s=1}^{k} (v_{ts} - \bar{v}_{t.})(v_{ts} - \bar{v}_{s.})$$

$$(4.3) \qquad = 2\left\{\sum_{t=1}^{k}\sum_{s=1}^{k} v_{ts}^2 - 2k\sum_{t=1}^{k} \bar{v}_{t.}^2 + k^2\bar{v}_{..}^2\right\}$$

$$= 2\sum_{t=1}^{k}\sum_{s=1}^{k}(v_{ts} - \bar{v}_{t.} - \bar{v}_{s.} + \bar{v}_{..})^2$$

where $\bar{v}_{tt} = \sum_{t=1}^{k} v_{tt}/k$, while $\bar{v}_{..} = \sum_{t=1}^{k}\sum_{s=1}^{k} v_{ts}/k^2$ and $\bar{v}_{t.} = \sum_{s=1}^{k} v_{ts}/k$. Now $K(Q_E) = (n-1)K_1(Q_c)$ and $K_2(Q_E) = (n-1)K_2(Q_C)$. Hence the null distribution of the ratio of mean squares $(n-1)\,Q_C/Q_E$ is approximately that of $F\{(k-1)\epsilon, (k-1)(n-1)\epsilon\}$ where

$$(4.4) \qquad \epsilon = k^2(\bar{v}_{tt} - \bar{v}_{..})^2/(k-1)\left\{\sum_{t=1}^{k}\sum_{s=1}^{k} v_{ts}^2 - 2k\sum_{t=1}^{k}\bar{v}_{t.}^2 - k^2\bar{v}_{..}^2\right\}.$$

We notice that the comparison of column and residual mean squares is without bias, whatever the nature of the matrix $\mathbf{v}$. The discrepancy that arises is represented in the approximation as a reduction by the same fraction ϵ of both degrees of freedom in the F ratio.

Between rows test. For testing row means the appropriate ratio of mean squares is $(k-1)Q_R/Q_E$. As we have seen, Q_R and Q_E are not distributed independently and the comparison is biassed unless the average covariance $\bar{v}_{ts}$ is zero.

To obtain under the null hypothesis the exact probability

$$P = \Pr\{Q_R/Q_E > \phi_\alpha\}$$

where $(k-1)\phi_\alpha = F_\alpha\{n-1, (n-1)(k-1)\}$ is the α probability point of the F distribution with $n-1$ and $(n-1)(k-1)$ degrees of freedom, we rewrite the probability in the form $\Pr\{(Q_R - \phi Q_E) > 0\}$ and employ Theorem 4.3′ as explained in Section 5′.

Let $\dot{\mathbf{Z}}$ be a $k(n-1) \times 1$ vector of the Z_{ti} arranged in the order $Z_{11}, \cdots, Z_{k1}$; $Z_{12}, \cdots, Z_{k2}$; $Z_{1(n-1)}, \cdots, Z_{k(n-1)}$. Let $\dot{\mathbf{V}}$ be the variance-covariance matrix for the Z_{ti} arranged in this order; thus $\dot{\mathbf{V}} = \mathcal{E}(\dot{\mathbf{Z}}\dot{\mathbf{Z}}')$. Then under the null hypothesis $\sum_{i=1}^{n}\beta_i^2 = \sum_{i=1}^{n-1} B_i^2 = 0$, the quadratic forms Q_R and Q_E are each functions of $\dot{\mathbf{Z}}$,

$$(4.5) \qquad Q_R = k\sum_{i=1}^{n-1}\bar{Z}_{.i}^2 = \dot{\mathbf{Z}}'\mathbf{M}_R\dot{\mathbf{Z}}$$

$$(4.6) \qquad Q_E = \sum_{i=1}^{n-1}\sum_{t=1}^{k}(Z_{ti} - \bar{Z}_{.i})^2 = \dot{\mathbf{Z}}'\mathbf{M}_E\dot{\mathbf{Z}}.$$

We require the probability that $\dot{\mathbf{Z}}'\mathbf{M}\dot{\mathbf{Z}}$ exceeds zero, where $\mathbf{M} = (\mathbf{M}_R - \phi\mathbf{M}_E)$. Now $\mathbf{M}_R$ is a $k(n-1) \times k(n-1)$ matrix partitioned after every k^{th} row and column, with each of its $n-1$ diagonal positions occupied by a $k \times k$ matrix $\mathbf{m}_R = k^{-1}\mathbf{1}_k\mathbf{1}_k'$ and zeros elsewhere:

(4.7)
$$\mathbf{M}_R = \begin{bmatrix} \mathbf{m}_R & \mathbf{0}_k & \cdots & \mathbf{0}_k \\ \mathbf{0}_k & \mathbf{m}_R & \cdots & \mathbf{0}_k \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0}_k & \mathbf{0}_k & \cdots & \mathbf{m}_R \end{bmatrix}.$$

Also, $\mathbf{M}_E$ and $\dot{\mathbf{V}}$, and hence $\dot{\mathbf{V}}(\mathbf{M}_R - \phi\mathbf{M}_E)$, are of this same form with the $k \times k$ matrices in the diagonal positions equal respectively to $\mathbf{m}_E = \mathbf{I}_k - k^{-1}\mathbf{1}_k\mathbf{1}_k'$, to $\mathbf{v}$, and to $\mathbf{v}(m_R - \phi\mathbf{m}_E)$. Hence the $(n - 1)k$ roots of the determinental equation

(4.8)
$$|\dot{\mathbf{V}}(\mathbf{M}_R - \phi\mathbf{M}_E) - \lambda\mathbf{I}_{k(n-1)}| = 0$$

are the k roots of the equation

(4.9)
$$\Delta_k = |\mathbf{v}(\mathbf{m}_R - \phi\mathbf{m}_E) - \lambda\mathbf{I}_k| = |\{\bar{v}_{t.} - \phi(v_{ts} - \bar{v}_{t.}) - \lambda\delta_{ts}\}| = 0,$$

each repeated $n - 1$ times where δ_{ts} is the Kronecker delta. Thus

(4.10)
$$\Pr\{Q_R/Q_E > \phi\} = \Pr\left\{\sum_{i=1}^{r'} \lambda_i\chi^2(n-1) + \sum_{j=r'+1}^{k} \lambda_j\chi^2(n-1) > 0\right\},$$

where λ_i and λ_j are respectively positive and negative roots of equation (4.9). No serious lack of generality in conclusions will be introduced if, in the examples we consider, we make the number of rows n odd so that $n - 1$ is even. Then we can apply Theorem 4.3′ and the required probability is

(4.11)
$$\Pr\{Q_R/Q_E > \phi\} = \sum_{i=1}^{r'} \sum_{s=1}^{(n-1)/2} \alpha_{is},$$

where the α's are obtained from equations (2.24′) (2.25′), and (2.26′).

The theory above may be used to study the distributions of the test criteria for any matrix $\mathbf{v}$. We use it here to consider the effect upon the significance test when

(i) the errors are independent but inequality of variance from column to column occurs,

(ii) the errors have equal variance but are serially correlated within rows.

5. Effect of inequality of column variances in two-way table. If we assume that the variance-covariance matrix $\mathbf{v}$ of "errors within rows" is diagonal, with elements $v_{11} = \sigma_1^2$, $v_{22} = \sigma_2^2$, $\cdots$, $v_{kk} = \sigma_{k'}^2$, we have the case in which the variance changes from column to column but the errors are distributed independently.

Between columns test. The matrix $\mathbf{u}$ of equation 3.8 reduces to

(5.1)
$$\mathbf{u} = \{(\delta_{ts} - k^{-1})\sigma_t^2\}.$$

Taking $n - 1$ even we can obtain the exact distribution of Q_c/Q_E under the null hypothesis, using Theorem 4.1′.

On simplification of equation (4.4) we find that $(n - 1)Q_c/Q_E$ is distributed approximately as $F\{(k - 1)\epsilon, (k - 1)(n - 1)\epsilon\}$ where

$$\epsilon = \{1 + c^2(k - 2)/(k - 1)\}^{-1}$$

and c is the coefficient of variation of the variances, given in equation (8.2′). The calculated values in Table 2 indicate that, as would be expected, the divergencies are similar to those for equal groups with the one-way classification.

Between rows test. Since the covariances v_{ts} are all zero, the comparison of row and error mean squares is not biassed. However, the row and column mean squares are not distributed independently. After substituting σ_t^2 for v_{tt} and zero for $v_{ts}(t \neq s)$ in Δ_k of equation (4.9), the resulting determinant may be simplified still further.

Here and in what follows we shall refer to the columns of a determinant, counting from left to right, as $c_1, c_2, \cdots$, etc., and the rows, counting from top to bottom, as $r_1, r_2, \cdots$, etc. By adding $c_2, c_3, \cdots, c_k$ to c_1, then subtracting $r_1 \times \sigma_j^2/\sigma_1^2$ from r_j $(j = 2, 3, \cdots, k)$, and finally dividing each row by k and changing signs in the last $k - 1$ rows, we find that the required k values of λ are the solutions of the determinantal equation

$$(5.2)\qquad (-1)^{k-1}\Delta_k = \begin{vmatrix} \sigma_1^2 - \lambda & \sigma_1^2(1+\phi)/k & \sigma_1^2(1+\phi)/k & \cdots & \sigma_1^2(1+\phi)/k \\ \lambda(1 - \sigma_2^2/\sigma_1^2) & \phi\sigma_2^2 + \lambda & 0 & \cdots & 0 \\ \lambda(1 - \sigma_3^2/\sigma_1^2) & 0 & \phi\sigma_3^2 + \lambda & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda(1 - \sigma_k^2/\sigma_1^2) & 0 & 0 & \cdots & \phi\sigma_k^2 + \lambda \end{vmatrix} = 0$$

In the one-way classification it appeared that for a given range of variances the greatest discrepancies might be expected when $k - 1$ of the variances were equal, while the k^{th} was larger (say a times as large as the others). Suppose the variances are $1, 1, \cdots, 1, a$. Then (5.2) reduces to

$$(5.3)\qquad (\phi + \lambda)^{(k-2)}[k\lambda^2 - \{(k - 1)(1 - \phi a) + (a - \phi)\}\lambda - ka\phi] = 0$$

from which all the λ's are readily obtained.

The results of a number of calculations using methods described above are shown in Table 2. It appears that the discrepancies in probability both for the test on rows as well as for the test on columns are not very large.

As was the case for the one-way classification, the effect of column-to-column differences in variance is to cause the significance of column differences in mean to be overestimated, although the differences in variance would have to be large for the effect to become serious. In the row comparisons, discrepancies of similar order but in the opposite direction occur, leading to underestimation of significance. Comparison of the first and third lines with the second and fourth

TABLE 2

Probabilities of exceeding 5 % point when column variances are unequal in the two-way analysis of variance table

Number of Rows n	Number of Columns k	Column Variances	True Chance (per cent) of Exceeding 5% point		Values in approximating distribution F (h', h) of ratio of mean squares*	
			Row Test, Exact	Columns Test, Approx.	h'	h
11	3	1 2 3	4.25	5.49	1.85(2)	18.46(20)
5	3	1 2 3	4.27	5.59	1.85(2)	7.38(8)
11	3	1 1 3	3.76	5.93	1.72(2)	17.24(20)
5	3	1 1 3	3.91	6.12†	1.72(2)	6.90(8)
3	5	1 1 1 1 3	4.47	6.92‡	3.21(4)	6.43(8)
3	11	1 1···1 3	4.86	7.09	7.90(10)	15.79(20)

* Bracketted values show appropriate degrees of freedom when variances are equal.

† 5.98 by the exact method.

‡ 6.75 by the exact method.

lines in Table 2 shows that the effects are worst when all the variances but one are at the lower end of the range. Comparison of the last four lines in the table suggests that the between-rows discrepancy is worse when the number of rows exceeds the number of columns, while the between-columns discrepancy is worse when the number of columns exceeds the number of rows.

6. Effect of serial correlation of errors within rows. Suppose that the normally distributed errors z_{1i}, z_{2i}, $\cdots$, z_{ki} in the i^{th} row of the analysis of variance table all have equal variance σ^2 but are not distributed independently. Thus $\mathbf{V} = \sigma^2\boldsymbol{\varrho}$, where $\boldsymbol{\varrho} = \{\rho_{ts}\}$ is a $k \times k$ matrix with diagonal elements all unity and the element ρ_{ts} of the t^{th} column and s^{th} row is the coefficient of correlation between z_{ti} and z_{si}, the same for all i. The theory described above enables us to examine the effect of any such correlation we choose.

A type of correlation of particular interest in practice is serial correlation which might be expected to arise when the observations within rows or columns were made at equally spaced intervals of time or space. This occurs when the rows of the two-way table are associated with a time factor, as in Fisher's example [1] and in the growth and wear curve examples of [5], or with a space factor as in Daniel's examples [3].

Normally the first order coefficient ρ_1, or ρ as we shall denote it, will be the largest of the serial correlations. We shall study the case where this first order serial correlation is taken into account but the effect of other correlations is ignored. Thus we shall assume

$$(6.1)\qquad \boldsymbol{\varrho} = \begin{bmatrix} 1 & \rho & 0 & \cdots & 0 & 0 \\ \rho & 1 & \rho & \cdots & 0 & 0 \\ 0 & \rho & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & \rho \\ 0 & 0 & 0 & \cdots & \rho & 1 \end{bmatrix}.$$

To ensure positive definiteness we also assume

$$(6.2)\qquad |\rho| < [2 \cos \{\pi/(k+1)\}]^{-1}.$$

"Between columns" test. In order to determine the exact probability $\Pr \{Q_C/Q_E > \phi\}$, we require the latent roots of $\mathbf{u}$ of equation (3.8). Making the substitution $\mathbf{v} = \boldsymbol{\varrho}\sigma^2$, where $\boldsymbol{\varrho}$ is defined in equation (6.1), and writing $\lambda = \lambda'\sigma^2$, the determinantal equation multiplied by k/σ^2 is

$$(6.3)\qquad | k\rho\mathbf{m} - k\lambda'\mathbf{I}_k | = \begin{vmatrix} k-(1+\rho)-k\lambda' & k\rho-(1+\rho) & -(1+\rho) & \cdots & -(1+\rho) \\ k\rho-(1+2\rho) & k-(1+2\rho)-k\lambda' & k\rho-(1+2\rho) & \cdots & -(1+2\rho) \\ -(1+2\rho) & k\rho-(1+2\rho) & k-(1+2\rho)-k\lambda' & \cdots & -(1+2\rho) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -(1+2\rho) & -(1+2\rho) & -(1+2\rho) & \cdots & k\rho-(1+2\rho) \\ -(1+\rho) & -(1+\rho) & -(1+\rho) & \cdots & k-(1+\rho)-k\lambda' \end{vmatrix} = 0.$$

To solve the equation, the determinant is first reduced to a more tractable form by a series of elementary transformations as follows:

(i) Add $c_2 + c_3 + \cdots + c_k$ to c_1.

(ii) Divide c_1 by $-k\lambda'$, and add $(2\rho + 1) \times c_1$ to $c_2, c_3, \cdots, c_k$, in turn.

(iii) Substitute $\lambda' = 1 + \rho\vartheta$ and divide $c_2, \cdots, c_k$ by $k\rho$.

(iv) Add c_3 to c_2, c_4 to c_3, $\cdots$, c_k to c_{k-1}.

(v) Add r_1 to r_2, $r_1 + r_2$ to r_3, $\cdots$, $r_1 + r_2 \cdots + r_{k-1}$ to r_k, and multiply c_k by k.

(vi) Add $(\vartheta - 2)c_1 + 2c_2 + 3c_3 \cdots + (k-1)c_{k-1}$ to c_k, change the sign of c_k, and interchange c_2 and c_k.

(vii) Add a new first row $r_0 = 10100 \cdots 0$ and a new first column $c_0 = 100 \cdots 0$, which leaves the value of the determinant unchanged.

(viii) Add r_0 to $r_1, r_2, \cdots, r_k$ in turn, and interchange rows and columns.

We now have the equation in the more manageable form

$$\text{(6.4)} \qquad \begin{vmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & \cdots & k-2 & k-1 & k \\ 1 & -\vartheta & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -\vartheta & 1 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & -\vartheta & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & -\vartheta & 1 \end{vmatrix} = 0,$$

the nonzero solutions of which give the required λ's via the relation $\lambda_t = (1 + \rho\vartheta_t)\sigma^2$. Denote the matrix of the determinant of (6.4) by $\mathbf{L}$ and a column vector of real numbers $(x_0, x_1, \cdots, x_k)$ by $\mathbf{x}$. Then the necessary and sufficient condition that a nontrivial solution exists for the equations $\mathbf{Lx} = \mathbf{0}$ is that $|\mathbf{L}| = 0$. Thus corresponding to each value of ϑ satisfying (6.4) there exists a set of solutions $x_0, x_1, \cdots, x_k$.

The equations $\mathbf{Lx} = \mathbf{0}$ may be written as

$$\text{(6.5)} \qquad \sum_{t=0}^{k} x_t = 0, \qquad \sum_{t=0}^{k} t x_t = 0,$$

$$\text{(6.6)} \qquad x_t - \vartheta x_{t+1} + x_{t+2} = 0, \qquad t = 0, 1, \cdots, k-2.$$

The difference equation (6.6) with boundary conditions given by (6.5) is readily solved by standard methods. With $\vartheta = 2\cos\phi$, a set of solutions

$$\text{(6.7)} \qquad x_t = e^{it\phi} - e^{i(k-t)\phi}$$

is obtained if $\phi = 2s\pi/k + 1$ or if $(k+2)\sin(\frac{1}{2}k\phi) = k\sin\frac{1}{2}(k+2)\phi$. In the first case,

$$\text{(6.8)} \qquad \vartheta = 2\cos\left(\frac{2s\pi}{k+1}\right), \qquad s = \begin{cases} 1, \cdots, \frac{1}{2}k; & k \text{ even}, \\ 1, \cdots, \frac{1}{2}(k-1); & k \text{ odd}. \end{cases}$$

In the second case, the remaining solutions for ϑ are most readily obtained by putting $t = \tan\frac{1}{2}\phi$, yielding

$$\text{(6.9)} \qquad \begin{cases} \displaystyle\sum_{s=1}^{k/2} (-1)^{s-1}(k-2s-1)\binom{k+2}{2s-1}(t^2)^{\frac{1}{2}k-s} = 0 & k \text{ even}, \\ \displaystyle\sum_{s=0}^{(k-1)/2} (-1)^{s}(k-2s-1)\binom{k+2}{2s}(t^2)^{\frac{1}{2}(k-1)-s} = 0 & k \text{ odd}. \end{cases}$$

TABLE 3

Values of ϑ for first order serial correlation

$k = 2$	3	4	5	6	7	8	9	10
−1.0000	0.0000	0.6180	1.0000	1.2470	1.4142	1.5321	1.6180	1.6825
	−1.3333	−0.5000	0.1165	0.5486	0.8544	1.0760	1.2405	1.3655
		−1.6180	−1.0000	−0.4450	0.0000	0.3473	0.6180	0.8308
			−1.7165	−1.2153	−0.7258	−0.3057	0.0407	0.3229
				−1.8019	−1.4142	−1.0000	−0.6180	−0.2846
					−1.8429	−1.5203	−1.1588	−0.8113
						−1.8794	−1.6180	−1.3097
							−1.9001	−1.6772
								−1.9190

TABLE 4

Between columns test: Values of ϵ

ρ	$k = 3$	$k = 5$	$k = 10$
−0.4	0.9576	0.8862	0.8233
−0.2	0.9863	0.9640	0.9453
+0.2	0.9769	0.9507	0.9222
+0.4	0.8832	0.8033	0.7718

Since $\vartheta = 2(1 - t^2)/(1 + t^2)$ is a single valued function of t^2, the polynomial equations (6.9) in t^2 supply the $\frac{1}{2}k - 1$ and $\frac{1}{2}(k - 1)$ values of ϑ required when k is even and odd, respectively, to give with (6.8) the total $k - 1$ solutions.

Values for $k = 2, 3, \cdots, 10$ are shown in Table 3, whence values of the λ's may be obtained for any chosen values of ρ and σ^2 from the relation $\lambda_t = (1 + \rho\vartheta_t)\sigma^2$. Using these values we may obtain the required probabilities from the exact series of Theorem 4.1′.

If we use the F approximation we consider that the ratio of mean squares $(n - 1)Q_C/Q_E$ is distributed approximately as $F\{(k - 1)\epsilon, (n - 1)(k - 1)\epsilon\}$, where

$$(6.10) \qquad \epsilon = \{1 + \rho^2 2(k + 1)(k - 2)^2/(k - 1)(k - 2\rho)^2\}^{-1}.$$

Values of the constant ϵ for various values of k and ρ are shown in Table 4. Since there is no bias and the effect of moderate correlation does not greatly reduce the degrees of freedom in the approximation, no large discrepancies will be expected in the between-columns comparison of the analysis of variance. Some calculated values are given in Table 6.

"Between rows" test. As we have seen already, the expectations of the row and error mean squares are equal only if the average covariance $\bar{v}_{ts} = 0$. For

TABLE 5

Between rows test: values of bias B

ρ	$k = 3$	$k = 5$	$k = 10$
−0.4	0.3684	0.3103	0.2593
−0.2	0.6471	0.6296	0.6154
+0.2	1.4615	1.4348	1.4167
+0.4	2.0909	1.9524	1.8696

TABLE 6

Probability (per cent) of exceeding 5 per cent point when first order serial correlation between errors within rows is ρ, for analysis of variance table with 5 rows and 5 columns

First order serial correlation, ρ	−0.4	−0.2	0.0	0.2	0.4
Exact per cent probability for test on rows	0.03	1.01	5.00	13.05	24.70
Approximate per cent probability for test on columns	5.90	5.27	5.00	5.37	*6.68

* By exact method, per cent probability is 6.43.

the case of first order serial correlation considered above, the expectations, under the null hypothesis, of the row and error mean squares are, respectively,

$$(1 + 2\rho(k - 1)/k)\sigma^2 \quad \text{and} \quad (1 - 2\rho/k)\sigma^2.$$

The ratio B of these expectations is

$$B = 1 + 2\rho k/(k - 2\rho). \tag{6.11}$$

Values of B for a number of values of k and ρ are shown in Table 5. This bias coefficient can be large even with only moderate correlation, and we shall therefore expect discrepancies to arise in the between-rows comparisons. Using equations (4.9), (4.10), and (4.11), exact probabilities for the between-rows test are obtained.

The results of a number of calculations for the case of the two-way table with five rows and columns are shown in Table 6. These confirm that very large discrepancies in the between-rows test in the directions expected do in fact occur, but that the between-columns comparisons are much less seriously affected. In particular, the remarks of R. A. Fisher concerning the analysis of the rainfall data are seen to be justified.

Acknowledgement. I am indebted to Mrs. Margaret Edmondson for valuable assistance with the calculations.

REFERENCES

[1] R. A. FISHER, *Statistical Methods for Research Workers*, 8th ed., Oliver and Boyd, Edinburgh, 1941, p. 226.

[2] N. SHAW, *The Air and its Ways*, Cambridge University Press, 1922, p. 211.

[3] H. E. DANIELS "Some problems of statistical interest in wool research," *J. Roy. Stat. Soc., Suppl.*, Vol. 5 (1938), pp. 89–128.

[4] H. E. DANIELS, "The effect of departures from ideal conditions other than non-normality on the t and z tests of significance," *Proc. Cambridge Philos. Soc.*, Vol. 34 (1938), pp. 321–328.

[5] G. E. P. BOX, "Problems in the analysis of growth and wear curves," *Biometrics*, Vol. 6 (1950), pp. 362–389.

[6] G. E. P. BOX, "Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 290–302.