A TEST FOR EQUALITY OF VARIANCES

Irving W. Burr and Louis A. Foster

Purdue University and Valparaiso University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 282

April, 1972

A TEST FOR EQUALITY OF VARIANCES

Irving W. Burr and Louis A. Foster

Purdue University and Valparaiso University

Testing the hypothesis of equality of variances for a set populations is a problem of much importance, and for which quite a few tests have been developed. A test proposed by A. E. Brandt (1932) and W. L. Stevens (1936) appears to have received little attention. The test they developed relied upon an asymptotic distribution. The test proposed in this paper uses a statistic denoted by Q, which is related to the coefficient of variation of the sample variances. It is a linear function of a statistic originally used by Brandt and Stevens. The definition herein is

$$Q = \sum_{i=1}^{k} s_i^4 / \left(\sum_{i=1}^{k} s_i^2\right)^2 \quad .$$

Exact moments are obtained for normal populations and equal sample sizes, and for an analogous definition of Q for unequal sample sizes. Critical values for Q are obtained by fitting the first four moments exactly for degrees of freedom, $\nu = 1(1)\ 10(2)\ 20$, and number of samples $p = 1(1)\ 10$ plus others to 64 and for $\alpha = .05, .025, .01, .001$.

The Q test compares favorably with other tests of homogeneity of variances, as to power, simplicity, availability for small samples, and ease of interpretation. Moreover it is not disturbed by a very small or even a zero sample variance.

1. INTRODUCTION

The problem of testing for homogeneity of variances of several populations is an important one. Tests based upon sample variances, standard deviations and ranges have been devised and used. Such tests are useful to justify the assumption of homoscedasticity in analysis of variance (ANOVA) and regression, and as a test of interest in its own right.

In 1947, C. Eisenhart [1] and W. G. Cochran [2] contributed articles to Biometrics summarizing various assumptions which are necessary for ANOVA, including the homogeneity assumption and the assumption that the populations are normal. In 1951, F. N. David and N. L. Johnson [3] published an article on the effect of non normality on the power function of the F test in ANOVA. In 1954, G.E.P. Box [4], [5] examined the effect on the F-statistic of violating the homogeneity assumption in one-way and two-way classifications. In reviewing these important articles and other contributions in this area, it may be gleaned that a violation of the homogeneity assumption when using ANOVA technique can create a severe disturbance to the F-ratio. Violation of the normality assumption, however does not appear to have such severe effects.

One may classify the various tests of homogeneity of variances for more than two populations as one of the following: (1) a ratio of the geometric mean to the arithmetic mean, (2) extreme values such as maximum variance divided by the minimum, (3) a ratio of the root-mean-square to the arithmetic mean of the variances, and (4) a control chart for ranges, standard deviations or variances. The test proposed in the present paper comes under general category (3). A sketch of the development of the approach follows.

In 1932, A. E. Brandt [6], in a Ph.D. thesis at Iowa State College, proposed a test statistic for use in a preliminary test to ANOVA. It is based on the following argument, as given in a discussion of the test by G. W. Snedecor [7] in 1937. Equal degrees of freedom ν are assumed for each cell estimate of population variance. Let

$$W = \sum_{j=1}^{p} \left(S_j^2 - \overline{S^2}\right)^2/(p-1) \quad , \tag{1}$$

where $\overline{S^2}$ is the arithmetic mean of the p sample variances. Hence W is the sample variance of sample variances. It is easily shown that $\sigma_W^2 = 2\sigma^4/\nu$, if each of the p populations is normal with variance σ^2. Let $Y = (p-1)W/\sigma_W^2$. Then the distribution of Y is approximately chi-square with $p-1$ degrees of freedom. Since σ^2 is unknown, Brandt proposed the following estimate. Using the mean square within-plots $\overline{S^2}$ to replace σ^2 in σ_W^2, let

$$Z = (p-1)W/\left(2\overline{S^2}^2/\nu\right) = \sum_{j=1}^{p} \left(S_j^2 - \overline{S^2}\right)^2/(2\overline{S^2}^2/\nu) \quad . \tag{2}$$

Then Z is still approximately distributed as chi-square with $p-1$ degrees of freedom, for sufficiently large samples. The rejection region is in the upper tail of the distribution since large values of z would usually be associated with heterogeneous population variances.

In 1936, W. L. Stevens [8] published a test which used the same idea as Brandt's test, although it would appear that Steven's result was reached independently of Brandt's. Let

$$\overline{S^2} = \sum_{j=1}^{p} \nu_j S_j^2 / \sum_{j=1}^{p} \nu_j \quad , \quad \text{and} \tag{3}$$

$$z_1 = \sum_{j=1}^{p} \nu_j \left(S_j^2 - \overline{S^2}\right)^2/2\overline{S^2}^2 \quad . \tag{4}$$

Then, as before, z_1 is approximately chi-square with $p-1$ degrees of freedom for sufficiently large sample sizes. This statistic of Stevens was given in a context somewhat more general than the ANOVA context in which Snedecor described Brandt's test, and does include varying degrees of freedom.

In category (1) are the L_1 test of J. Neyman and E. S. Pearson [9], and the test by M. S. Bartlett [10]. The L_1 test was further developed by S. S. Wilks [11], C. M. Thompson [12] and U. S. Nair [13]. In category (2), are a test with equal degrees of freedom $(\max s_i^2/\sum_1^p s_i^2)$ by Cochran [14] and the maximum F test $(\max S_i^2/\min S_i^2)$ by H. O. Hartley [15].

In category (4) are control charts for ranges or standard deviations, which are frequently used in statistical quality control. They may also be used to test homogeneity of variability, W. A. Shewhart [16].

2. THE Q-STATISTIC

Consider first the case of p independent random samples of n observations each from normal populations with variances σ_j^2, $j = 1,\ldots, p$. Interest is in H_0: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_p^2 = \sigma^2 > 0$, say, vs. H_1: that the σ_j^2 are not all equal. Note that even H_0 is a composite hypothesis. By using a standardized statistic for the test, however, it can be treated like a simple hypothesis.

The sample variance of sample variances is given by (1). It measures the variability between sample variances in absolute terms. If we knew the hypothetical common σ^2 we could standardize by dividing by σ_W^2, but since it is not available we standardize as in (2). Such a statistic is then free from σ^2, and also any physical units. However,

$$Z = \left[\sum_{j=1}^p S_j^4 - p\,\overline{S^2}^2\right]/\left(2\overline{S^2}^2/\nu\right)$$

$$= \frac{p^2\nu}{2}\,\frac{\sum_{j=1}^p S_j^4}{\left(\sum_{j=1}^p S_j^2\right)^2} - \frac{p\nu}{2} \quad . \tag{5}$$

Now defining

$$Q = (\sum_{j=1}^{p} S_j^4)/(\sum_{j=1}^{p} S_j^2)^2 \tag{6}$$

one has

$$Z = \frac{p^2\nu}{2} \cdot Q - \frac{p\nu}{2} \quad . \tag{7}$$

Thus the Q statistic here proposed is a linear function of that proposed by Brandt. Moreover from (6) we easily have

$$Q = \left[\frac{\mathrm{RMS}(S_j^2)}{\mathrm{AM}(S_j^2)} \right] \cdot \frac{1}{p} \quad , \tag{8}$$

thus indicating that Q is a monotone function of $\mathrm{RMS}(S_j^2)/\mathrm{AM}(S_j^2)$, and thus a test of category (3). Bartlett's and L_1 tests are of course monotone functions of $\mathrm{AM}(S_i^2)/\mathrm{GM}(S_i^2)$ and of category (1). Either of (5) or (8) shows that the minimum value of Q is $1/p$, occurring when all s_i^2 are equal and positive.

Under the assumptions of random independent samples from populations $N(\mu_j, \sigma^2)$ we shall be able to find in closed form the first four moments. Then, using a suitable system of distributions matching exactly the first four moments for each (p,ν) combination, the desired critical values are obtained as given in Table 1. Q is more convenient for finding moments and distributions than is Z.

For varying degrees of freedom ν_j, the definition is

$$Q = \overline{\nu} \sum_{j=1}^{p} \nu_j S_j^4 / [\sum_{j=1}^{p} \nu_j S_j^2]^2 \quad , \tag{9}$$

which specializes to (6) when $\nu_j \equiv \nu$.

A recent paper by A. Cohen and W. E. Strawderman [17] contains a theorem which shows that a test of $\sigma_1^2 = \ldots = \sigma_p^2 > 0$ using (6) is unbiased, that is, the probability of "acceptance" of this hypothesis when true is at a maximum, as against alternative conditions.

3. THE Q-STATISTIC FOR TWO VARIANCES FROM NORMAL POPULATIONS.

The exact density function of Q may be found for two samples of n_1, n_2 independent observations from populations $N(\mu_i, \sigma_i)$ $i = 1,2,$ L. A. Foster [18]. If $\sigma_1 = \sigma_2 = \sigma$, say, then the distribution of Q does not contain σ, just as is also true of the F distribution. In fact one can readily find

$$Q = 1 - \frac{2}{F+2+(1/F)} \quad , \qquad \text{or} \tag{10}$$

$$F = \frac{Q \pm \sqrt{2Q-1}}{1 - Q} \quad . \tag{11}$$

The $\pm$ in the latter is an indication that while Q is relatively high when either $s_1 > s_2$ or $s_2 > s_1$, F is high in the former case and low in the latter case. Thus a single upper-tail test of $\sigma_1 = \sigma_2$ for Q is equivalent to a two-tail test for F.

It is, however, the authors' belief that the F test cannot be improved upon by the use of the Q statistic, and hence results are not given in Table 1 for $p = 2$, although they were obtained by Foster [18].

4. THE Q-STATISTIC FOR MORE THAN TWO VARIANCES FROM NORMAL POPULATIONS.

Several different approaches were tried to find the explicit distribution function or the density function for Q when $p > 2$, without noteworthy success. The necessary multiple integrals and regions of integration become unmanageable even for $p = 3$. Accordingly attention was turned toward finding the first four moments for the Q statistic. This was readily accomplished Foster [18]. See Section 5. Then a distribution function for each case of $\nu = 1,\ldots, 10$ and $p \geq 3$, was fitted, matching the first four moments precisely and desired percentile points obtained. See Section 6. The results are given in Table 1.

5. THE MOMENTS OF Q, NORMAL POPULATIONS.

In evaluating the moments of Q in the normal case, the following notations are convenient:

$$\overrightarrow{S^2} = (S_1^2,\ldots, S_p^2) \tag{12}$$

$$\vec{\nu} = (\nu_1,\ldots, \nu_p) \text{ and } \bar{\nu} = \sum_{j=1}^{p} \nu_j/p \quad . \tag{13}$$

The latter will sometimes be used as a subscript for Q, or if all ν_j's are equal to ν, we write Q_ν. When considering Q^m, $m = 1,2,3,4$, let

$$\vec{m} = (m_1,\ldots, m_p) \tag{14}$$

$$\overrightarrow{m!} = (m_1!)(m_2!) \cdots (m_p!) \tag{15}$$

$$Q^{\vec{m}} = \Pi_{j=1}^{p}(\nu_j S_j^2)^{2m_j}/(\textstyle\sum_{j=1}^{p} \nu_j S_j^2)^{2m} \quad . \tag{16}$$

Then Q^m can be expressed by a multinomial summation involving terms containing $Q^{\vec{m}}$. Thus from (9)

$$Q^{m} = \left[\bar{\nu} \;\sum_{j=1}^{p} \nu_j S_j^4 / (\sum_{j=1}^{p} \nu_j S_j^2)^2\right]^m$$

$$= \bar{\nu}^m (\sum_{j=1}^{p} \nu_j S_j^4)^m / (\sum_{j=1}^{p} \nu_j S_j^2)^{2m}$$

$$= \bar{\nu}^m \sum_{\sum_1^p m_j = m} \left[\frac{m! \;\Pi_{j=1}^{p} (\nu_j S_j^4)^{m_j}}{\vec{m}! \;(\sum_{j=1}^{p} \nu_j S_j^2)^{2m}}\right]$$

$$= \sum_{\sum_1^p m_j = m} \left[\frac{(m!)(\bar{\nu}^m) \;\Pi_{j=1}^{p} (\nu_j^2 \; S_j^4)^{m_j}}{\vec{m}! \;(\Pi_{j=1}^{p} \nu_j^{m_j})(\sum_{j=1}^{p} \nu_j S_j^2)^{2m}}\right]$$

$$= \sum_{\sum_1^p m_j = m} (m!/\vec{m}!)(\bar{\nu}^m / \Pi_{j=1}^{p} \nu_j^{m_j}) \; Q^{\vec{m}} \;. \tag{17}$$

We next find an integral expression for $E(Q^m)$ under either

$$H_0 : \quad \sigma_1 = \sigma_2 = \ldots = \sigma_p \quad , \qquad \text{or} \tag{18}$$

$$H_1 : \quad \sigma_j\text{'s not all equal} \quad , \tag{19}$$

and then derive $E(Q_{\vec{\nu}}^{\vec{m}})$ in closed form for H_0.

First we obtain a representation for $Q^{\vec{m}}$ in terms of chi-square random variables. Let $Z_j = \nu_j S_j^2/\sigma_j^2$ $j = 1,\ldots, p$. Under either H_0 or H_1, each Z_j has a chi=square distribution with ν_j degrees of freedom, and they are mutually independent. Using (16)

$$Q^{\vec{m}} = \Pi_{j=1}^{p} (\sigma_j^2 Z_j)^{2m_j} / (\sum_{j=1}^{p} \sigma_j^2 Z_j)^{2m} \,. \tag{20}$$

Since the Z_j's are independent chi-square we have

$$E(Q^{\vec{m}}) = C_1 \int_0^\infty \cdots \int_0^\infty \frac{[\Pi_{j=1}^p (\sigma_j^2 z_j)^{2m_j} z_j^{(\nu_j/2)-1}]}{(\sum_{j=1}^p \sigma_j^2 z_j)^{2m}} \cdot$$

$$\cdot \ [\exp(-\tfrac{1}{2}\textstyle\sum_{j=1}^p z_j)]\ [\Pi_{j=1}^p dz_j] \ , \qquad (21)$$

where

$$C_1 = 2^{-(p\bar{\nu}/2)} \ [\Pi_{j=1}^p \Gamma(\tfrac{\nu_j}{2})]^{-1} \ . \qquad (22)$$

Consider the following transformation $y_0 = 0$ and $y_j = \sum_{i=1}^j \sigma_i^2 z_i$ for $j = 1,\ldots, p$. Then $z_j = (1/\sigma_j^2)(y_j - y_{j-1})$ for $j = 1,\ldots, p$. Since $P(Z_j = 0) = 0$, we can assume $0 < y_1 < y_2 < \ldots < y_p$. The Jacobian of this transformation is $(\sigma_1^2 \ldots \sigma_p^2)^{-1}$. Next let

$$C_2 = C_1/\Pi_{j=1}^p \sigma_j^{\nu_j} \ . \qquad (23)$$

Then

$$E(Q^{\vec{m}}) = C_2 \int_0^\infty \int_0^{y_p} \cdots \int_0^{y_2} \left\{ \frac{[\Pi_{j=1}^p (y_j - y_{j-1})^{2m_j + (\nu_j/2)-1}]}{y_p^{2m}} \cdot \right.$$

$$\left. \cdot \ [\exp(-\tfrac{1}{2}\textstyle\sum_{j=1}^p (y_j - y_{j-1})/\sigma_j^2)] \right\} \Pi_{j=1}^p dy_j \ . \qquad (24)$$

Next consider the transformation $y_j = \Pi_{i=j}^p u_i$ for $j = 0,1,\ldots, p$. Then $u_0 = y_0 = 0$, and $(y_j - y_{j-1}) = (u_j u_{j+1} \ldots u_p)(1 - u_{j-1})$, $j = 1,\ldots, p$. The Jacobian is $u_2 u_3^2 \ldots u_p^{p-1}$. Then we have

$$E(Q^{\vec{m}}) = C_2 \ [\int_0^\infty u_p^{(p\bar{\nu}/2)-1} \exp(-u_p/2\sigma_p^2)\ du_p] \ \cdot \qquad (25)$$

$$\int_0^1 \cdots \int_0^1 \Pi_{j=1}^{p-1} u_j^{(\gamma_1 + \ldots + \gamma_j)-1} (1-u_j)^{\nu_{j+1}-1} \exp\{-\tfrac{1}{2}\textstyle\sum_{j=1}^{p-1} (\Pi_{i=j}^p u_i)\delta_j\} du_1 \ldots du_{p-1},$$

where

$$\gamma_j = 2m_j + (\nu_j/2) \qquad j = 1,\ldots, p$$

$$\delta_j = (1/\sigma_j^2) - (1/\sigma_{j+1}^2) \quad j = 1,\ldots, p-1 \quad .$$

Evaluation of (25) under H_1 will be discussed later on. For the present we assume H_0, and thus that $\delta_j \equiv 0$. Also call $\sigma_j = \sigma$, $j = 1,\ldots, p$. Then

$$E(Q^{\vec{m}}) = C_2 \int_0^\infty u^{(p\bar{\nu}/2)-1} \exp(-u_p/2\sigma^2) du_p \left\{ \prod_{j=1}^{p-1} [\int_0^1 u_j^{(\gamma_1+\ldots+\gamma_j)-1} (1-u_j)^{\gamma_{j+1}-1} du_j] \right\}$$

$$= C_2 (2\sigma^2)^{p\bar{\nu}/2}\, \Gamma(p\bar{\nu}/2)\, \Pi_{j=1}^{p-1}\, \beta(\gamma_1+\ldots+\gamma_j,\ \gamma_{j+1}) \quad , \tag{26}$$

where Γ and β indicate gamma and beta functions. Using $\beta(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$, cancellation yields

$$E(Q^{\vec{m}}) = \frac{\Gamma(p\bar{\nu}/2)}{\Gamma(p\bar{\nu}/2+2m)}\ \Pi_{j=1}^{p}\ \frac{\Gamma[(\nu_j/2)+2m_j]}{\Gamma(\nu_j/2)} \quad . \tag{27}$$

Substituting (27) into (17) yields

$$E(Q_{\vec{\nu}}^{m}) = \sum_{\sum_{j=1}^p m_j = m} \left\{ \frac{m!\ \bar{\nu}^m \Gamma(p\bar{\nu}/2)}{\vec{m}!\,[\Pi_{i=1}^p \nu_j^{m_j}]\,\Gamma(p\bar{\nu}+2m)}\ \Pi_{i=1}^p \left[\frac{\Gamma(\nu_i/2)+2m_i}{\Gamma(\nu_i/2)}\right] \right\} \quad . \tag{28}$$

Or if the ν_j's are constant at ν, this expression becomes

$$E(Q_\nu^m) = \sum_{\sum_{j=1}^p m_j = m} \left\{ \frac{m!\,\Gamma(p\nu/2)\Pi_{i=1}^p \Gamma[(\nu/2)+2m_i]}{\vec{m}!\,\Gamma[(p\nu/2)+2m]\,[\Gamma(\nu/2)]^p} \right. \quad . \tag{29}$$

Specifically when $\nu_j \equiv \nu$ we have the following for $m = 1,2,3,4$:

If $m = 1$, there are p vectors $\vec{m}$, such that one component is 1 and the other $p-1$ zero. Using $\Gamma(a+1) = a\Gamma(a)$:

$$E(Q_\nu) = \frac{\nu+2}{p\nu+2} \quad . \tag{30}$$

If $m = 2$, there are p vectors $\vec{m}$, such that one component is 2, and $p(p-1)/2$ vectors with two m_i's of one and the others zero, yielding:

$$E(Q_\nu^2) = \frac{(\nu+2)[p\nu(\nu+2)+8(\nu+3)]}{(p\nu+2)(p\nu+4)(p\nu+6)} \quad . \tag{31}$$

Then using $\mu_2 = \sigma_Q^2 = E(Q_\nu^2) - [E(Q_\nu)]^2$ we have

$$\mu_2 = \frac{8\nu(\nu+2)(p-1)}{(p\nu+2)^2(p\nu+4)(p\nu+6)} \quad . \tag{32}$$

Three distinct kinds of vectors occur in Q_ν^3, from which after much tedious calculation using $\mu_3 = E(Q_\nu^3) - 3E(Q_\nu^2)E(Q_\nu) + 2[E(Q_\nu)]^3$, and $\alpha_3 = \mu_3/\mu_2^{3/2}$, one finds

$$\alpha_3^2 = \frac{8(p\nu+4)(p\nu+6)[p\nu^2+(8p-10)\nu-8]^2}{\nu(\nu+2)(p-1)(p\nu+8)^2(p\nu+10)^2} \quad . \tag{33}$$

If $p = 2$, $\nu = 1$, then $\alpha_3 = 0$; otherwise $\alpha_3 > 0$. Furthermore

$$\alpha_4 = 3(p\nu+4)(p\nu+6)[p^2\nu^4(p+3)+2p\nu^3(p^2+68p-93) + \tag{34}$$
$$+ 4\nu^2(125p^2-245p+84) - 128\nu(2p-3)+384] \div$$
$$[\nu(\nu+2)(p-1)(p\nu+8)(p\nu+10)(p\nu+12)(p\nu+14)] \quad .$$

The following are readily verified :

$$\lim_{p\to\infty} \alpha_3^2 = 0 \qquad \lim_{p\to\infty} \alpha_4 = 3 \tag{35}$$

$$\lim_{\nu\to\infty} \alpha_3^2 = 8/(p-1) \qquad \lim_{\nu\to\infty} \alpha_4 = 3(p+3)/(p-1) \tag{36}$$

6. APPROXIMATION OF CRITICAL REGION FOR Q-TEST, NORMAL POPULATIONS UNDER H_0: $\sigma_1 = \ldots = \sigma_p$.

Since the exact distribution of Q in this case was not obtained, we resort to approximation of percentile points by means of members of a family of distributions. The family chosen is given in I. W. Burr [19] and Burr and P. J. Cislak [20]. The distribution function is

$$F(x) = 1 - (1+x^c)^{-k} \quad c,k,x > 0 \quad . \tag{37}$$

The parameters c and k determine shape characteristics α_3 and α_4, and also μ, σ for x. Formulas (33) and (34) were used to find curve-shape characteristics $\alpha_{3:Q}$ and $\alpha_{4:Q}$. Then through successive approximation, a (c,k) combination was found to yield the precisely matching $\alpha_{3:x}$ and $\alpha_{4:x}$. For some combinations of $\alpha_{3:Q}$ and $\alpha_{4:Q}$ it was necessary to use the reciprocal transformation $x = 1/y$ on (37) yielding

$$F_1(y) = (1+y^{-c})^{-k} \quad c,k,y > 0 \quad . \tag{38}$$

Then matching the first two moments via

$$(Q - \mu_Q)/\sigma_Q = (x - \mu_x)/\sigma_x \quad , \tag{39}$$

it was possible to find the percentile points of Q approximately, as shown in Table 1. In some cases there were two, or even three, combinations of c and k in (37) and (38) yielding a perfect match of $\alpha_{3:Q}$ and $\alpha_{4:Q}$. In such cases we chose the largest percentile value, of the two or three, so as to be conservative. For the most part such approximated percentile points of Q differed but little for the (c,k) pairs.

It is to be noted that E. S. Pearson [21] has shown that in fitting a distribution by moments, the error displacement of percentile points on the steep tail of a skewed distribution is much more severe than on the long tail. This is fortunate here because all of the Q distributions are strongly positively skewed, and it is in the upper tail that our interest centers for percentiles for the Q test.

7. MOMENTS FOR Q, UNEQUAL SAMPLE SIZES, NORMAL POPULATIONS, UNDER H_0: $\sigma_1 = \ldots = \sigma_p$.

Starting with (28) for varying ν_j's, Foster [18], found

$$E(Q_{\vec{\nu}}) = \frac{\bar{\nu}+2}{p\bar{\nu}+2} \quad , \tag{40}$$

which is a virtual analog of (30). Next he found

$$E(Q^2_{\vec{\nu}}) = E(Q^2_{\bar{\nu}}) + 48[\bar{\nu}/\hat{\nu}-1]/[(p\bar{\nu}+2)(p\bar{\nu}+4)(p\bar{\nu}+6)] \quad , \tag{41}$$

where $\hat{\nu}$ is the harmonic mean of the ν_j's, $p/\sum_{j=1}^{p}(1/\nu_j)$. The second term on the right is the error made in the variance of Q if $\bar{\nu}$ is substituted for the ν_j's. Since $\bar{\nu} \geq \hat{\nu}$, the correction is always non-negative, being zero only for constant ν_j's. For fixed p, this error is close to zero if $\hat{\nu}$ is close to $\bar{\nu}$, or if p is large the error approaches zero faster than the variance of Q approaches zero.

Two examples are given below for $p = 3$

	ν_1	ν_2	ν_3	$\bar{\nu}$	$\hat{\nu}$	$\sigma_{Q_{\vec{\nu}}}/\sigma_{Q_{\bar{\nu}}}$
Example 1:	1	1	10	4	1.429	1.146
Example 2:	5	5	4	4.667	4.615	1.008

Foster also obtained rather complicated expressions for the corrections to the third and fourth moments of Q about the origin. An example of the effect of using $\bar{\nu}$ for $\vec{\nu}$ is given below for $p = 3$, $\nu_1 = 5$, $\nu_2 = 4$, $\nu_3 = 3$:

Calculations based on	$\bar{\nu}$	$\vec{\nu}$
Mean Q	.4286	.4286
Standard deviation	.08248	.08569
Coefficient of skewness (α_3)	1.512	1.806
Coefficient of kurtosis (α_4)	5.84	8.12

Foster also found some further results, in special cases and a series representation. If the population is discrete, it is possible to evaluate the exact distribution of S^2, and from it to obtain the exact distribution of Q, for the null hypothesis H_0. But this very quickly becomes extremely tedious, then unmanageable.

8. THE Q-TEST FOR HOMOGENEITY OF VARIANCES.

We thus have the simple test statistic

$$Q = (\textstyle\sum_{j=1}^{p} S_j^4)/(\textstyle\sum_{j=1}^{p} S_j^2)^2 \qquad (6)$$

for the hypothesis H_0: $\sigma_1 = \ldots = \sigma_p$, vs. the alternative hypothesis H_1; σ's not all equal. Then for a given set of sample variances, $s_1^2, s_2^2, \ldots, s_p^2$ assumed to come from normal populations, we use (6) to find an observed q-value. Entering Table 1 with the appropriate ν and p, and a chosen α level we find the critical value. Then if

$$q > Q_{\nu,p,\alpha}\,, \quad \text{reject } H_0 \tag{42}$$

$$q < Q_{\nu,p,\alpha}\,, \quad \text{accept } H_0\,. \tag{43}$$

Or, if the sample sizes vary we use (9) to find the observed q value, then enter Table 1 with $\bar{\nu}$, p and α.

9. TABLE OF CRITICAL VALUES FOR THE Q STATISTIC

The percentile points for the distribution of the Q statistic were approximated as discussed in Section 6. Beginning with starting values of c and k and a desired $\alpha_{3:Q}$ and $\alpha_{4:Q}$ for a (ν,p) combination, the program iterated on c and k until both $\alpha_{3:x}$ and $\alpha_{4:x}$ agreed to the nearest .00001. Then using (39) the percentile points were calculated to the nearest six decimal places and rounded off to the nearest three.

TABLE 1. PERCENTILE POINTS FOR Q-TEST (6), FOR EQUAL DEGREES OF FREEDOM ν, AND FOR p SAMPLES.

	ν = 1				ν = 2			
p	.95	.975	.99	.999	.95	.975	.99	.999
3	.915	.957	*	*	.752	.801	.863	*
4	.799	.853	.920	*	.612	.659	.720	.898
5	.702	.757	.828	*	.511	.553	.608	.773
6	.621	.675	.744	.949	.436	.479	.539	.690
7	.555	.605	.671	.865	.379	.416	.469	.606
8	.500	.547	.609	.793	.334	.366	.412	.537
9	.454	.505	.576	.750	.298	.328	.371	.481
10	.415	.461	.528	.694	.269	.295	.333	.433
12	.352	.391	.448	.598	.223	.244	.276	.358
14	.305	.339	.391	.522	.191	.208	.234	.303
15	.285	.317	.365	.490	.177	.193	.217	.280
16	.268	.297	.343	.460	.166	.180	.202	.261
18	.238	.264	.304	.409	.146	.159	.178	.228
20	.214	.237	.273	.367	.131	.142	.158	.202
22	.194	.214	.246	.332	.118	.128	.142	.180
24	.177	.196	.224	.302	.108	.116	.129	.162
26	.163	.180	.206	.276	.099	.107	.118	.148
28	.151	.166	.190	.254	.091	.098	.108	.135
30	.140	.154	.176	.234	.085	.091	.100	.124
32	.131	.144	.163	.218	.079	.085	.093	.115
36	.115	.126	.143	.189	.070	.075	.082	.100
40	.103	.112	.127	.167	.062	.066	.072	.088
45	.091	.099	.111	.145	.055	.058	.063	.076
50	.081	.088	.098	.127	.049	.052	.056	.067
60	.066	.072	.080	.102	.040	.042	.045	.053
64	.062	.067	.074	.094	.037	.039	.042	.049

* These entries exceeded 1 using the approximating distribution. Since $Q \leq 1$, they are omitted.

TABLE 1 CONTINUED

	ν = 3				ν = 4			
p	.95	.975	.99	.999	.95	.975	.99	.999
3	.657	.701	.757	.919	.596	.634	.684	.828
4	.517	.555	.605	.754	.461	.498	.549	.675
5	.423	.460	.512	.644	.374	.402	.443	.552
6	.356	.386	.430	.546	.312	.335	.369	.461
7	.307	.334	.372	.471	.267	.288	.318	.394
8	.268	.291	.325	.411	.234	.251	.276	.342
9	.238	.258	.287	.363	.207	.222	.244	.300
10	.214	.231	.257	.324	.185	.199	.218	.267
12	.177	.191	.211	.265	.153	.164	.179	.217
14	.150	.162	.178	.222	.130	.139	.151	.181
15	.140	.150	.165	.205	.121	.129	.140	.167
16	.131	.140	.154	.190	.113	.120	.130	.155
18	.115	.124	.135	.165	.100	.106	.114	.135
20	.103	.110	.120	.146	.089	.094	.101	.119
22	.093	.099	.108	.130	.081	.085	.090	.106
24	.085	.090	.098	.117	.074	.077	.082	.096
26	.078	.083	.090	.107	.068	.071	.075	.087
28	.072	.076	.082	.098	.063	.065	.069	.080
30	.067	.070	.075	.090	.058	.061	.064	.074
32	.062	.066	.070	.083	.054	.057	.060	.068
36	.055	.058	.062	.072	.048	.050	.052	.060
40	.049	.052	.055	.064	.043	.045	.047	.053
45	.043	.045	.048	.055	.038	.039	.041	.046
50	.039	.040	.043	.049	.034	.035	.037	.041
60	.032	.033	.035	.039	.028	.029	.030	.033
64	.030	.031	.033	.037	.026	.027	.028	.031

TABLE 1 CONTINUED

	ν = 5				ν = 6			
p	.95	.975	.99	.999	.95	.975	.99	.999
3	.554	.588	.631	.760	.524	.554	.593	.708
4	.425	.454	.498	.608	.399	.424	.461	.558
5	.342	.365	.399	.490	.320	.339	.368	.446
6	.285	.305	.334	.407	.266	.283	.307	.368
7	.243	.260	.284	.345	.227	.241	.261	.311
8	.212	.226	.246	.298	.198	.210	.226	.268
9	.188	.200	.217	.261	.175	.185	.199	.235
10	.168	.179	.194	.232	.157	.166	.178	.208
12	.139	.147	.159	.188				
14	.118	.125	.134	.157				
15	.110	.115	.123	.145	.102	.107	.113	.131
16	.103	.108	.115	.134				
18	.091	.095	.101	.117				
20	.081	.085	.090	.104	.076	.079	.083	.094
22	.073	.077	.081	.093				
24	.067	.070	.074	.084				
27	.062	.064	.067	.076				
28	.057	.059	.062	.070				
30	.053	.055	.058	.065	.050	.051	.053	.059
32	.050	.051	.054	.060				
36	.044	.045	.047	.052				
40	.039	.040	.042	.047	.037	.038	.039	.043
45	.035	.036	.037	.041				
50	.031	.032	.033	.036	.029	.030	.031	.033
60	.026	.026	.027	.029	.024	.025	.025	.027
64	.024	.025	.025	.027				

TABLE 1 CONTINUED

	$\nu = 7$				$\nu = 8$			
p	.95	.975	.99	.999	.95	.975	.99	.999
3	.501	.528	.562	.666	.483	.507	.539	.633
4	.379	.401	.434	.520	.364	.384	.413	.490
5	.303	.320	.346	.413	.291	.306	.328	.388
6	.252	.267	.288	.340	.242	.254	.271	.318
7	.215	.227	.244	.287	.206	.216	.230	.268
8	.187	.197	.210	.247	.180	.188	.199	.231
9	.166	.174	.185	.216	.159	.166	.176	.202
10	.149	.155	.165	.192	.142	.148	.157	.179
15	.097	.101	.106	.121	.093	.097	.101	.113
20	.072	.075	.078	.087	.069	.071	.074	.082
30	.047	.049	.050	.055	.045	.047	.048	.052
40	.035	.036	.037	.040	.034	.034	.035	.038
50	.028	.028	.029	.031	.027	.027	.028	.030
60	.023	.023	.024	.025	.022	.023	.023	.024

TABLE 1 CONTINUED

	ν = 9				ν = 10			
p	.95	.975	.99	.999	.95	.975	.99	.999
3	.468	.493	.529	.619	.456	.479	.512	.596
4	.353	.370	.396	.465	.343	.359	.383	.446
5	.281	.295	.315	.367	.274	.286	.303	.351
6	.234	.244	.260	.301	.227	.237	.250	.288
7	.199	.208	.220	.254	.194	.201	.212	.242
8	.174	.181	.191	.219	.169	.175	.184	.209
9	.154	.160	.168	.192	.149	.155	.162	.183
10	.138	.143	.150	.170	.134	.139	.145	.163
15	.090	.093	.097	.108	.088	.091	.094	.103
20	.067	.069	.071	.078	.065	.067	.069	.075
30	.044	.045	.046	.050	.043	.044	.045	.048
40	.033	.033	.034	.036	.032	.033	.033	.035
50	.026	.026	.027	.028	.025	.026	.026	.028
60	.022	.022	.022	.023	.021	.021	.022	.023

TABLE 1 CONTINUED

	ν = 12				ν = 14			
p	.95	.975	.99	.999	.95	.975	.99	.999
3	.438	.457	.486	.558	.424	.440	.466	.530
4	.328	.342	.362	.415	.318	.329	.347	.393
5	.262	.272	.287	.326	.253	.262	.275	.308
6	.217	.225	.236	.267	.210	.217	.227	.253
7	.185	.192	.201	.225	.179	.185	.192	.213
8	.161	.167	.174	.194	.156	.161	.167	.184
9	.143	.148	.154	.170	.138	.142	.148	.162
10	.128	.132	.137	.152	.124	.127	.132	.144
15	.084	.086	.089	.097	.082	.084	.086	.092
20	.063	.064	.066	.070	.061	.062	.063	.067
30	.041	.042	.043	.045	.040	.041	.042	.043
40	.031	.031	.032	.033	.030	.030	.031	.032
50	.025	.025	.025	.026	.024	.024	.024	.025
60	.020	.021	.021	.022	.020	.020	.020	.021

TABLE 1 CONTINUED

	ν = 16				ν = 18			
p	.95	.975	.99	.999	.95	.975	.99	.999
3	.413	.428	.451	.508	.405	.418	.439	.490
4	.309	.320	.335	.375	.303	.312	.326	.362
5	.247	.254	.265	.295	.242	.248	.258	.284
6	.205	.211	.219	.242	.200	.206	.213	.233
7	.175	.180	.186	.204	.171	.175	.181	.197
8	.152	.156	.162	.176	.149	.153	.158	.170
9	.135	.138	.143	.155	.132	.135	.139	.150
10	.121	.124	.128	.138	.119	.121	.125	.134
15	.080	.081	.083	.089	.078	.080	.082	.086
20	.060	.060	.062	.065	.058	.059	.060	.063
30	.039	.040	.040	.042	.039	.039	.040	.041
40	.029	.030	.030	.031	.029	.029	.029	.030
50	.023	.024	.024	.025	.023	.023	.023	.024
60	.019	.020	.020	.020	.019	.019	.019	.020

TABLE 1 CONTINUED

	$\nu = 20$					$\nu = 20$			
p	.95	.975	.99	.999	p	.95	.975	.99	.999
3	.398	.410	.429	.476	10	.117	.119	.122	.130
4	.298	.306	.319	.351	15	.077	.078	.080	.084
5	.237	.244	.252	.276	20	.058	.058	.059	.062
6	.197	.202	.209	.226	30	.038	.039	.039	.040
7	.168	.172	.178	.191	40	.028	.029	.029	.030
8	.147	.150	.154	.166	50	.023	.023	.023	.024
9	.130	.133	.136	.146	60	.019	.019	.019	.020

10. ADVANTAGES OF THE Q TEST

As compared with other tests of the homogeneity of a set of variances the following seem to be indicated:

1. The Q statistic is more easily calculated from a set s_j^2 variances than is Bartlett's test. It is of course not as easily calculated as is Cochran's or Hartley's tests.

2. The Q test is as easily interpreted as any of the tests with the possible exception of a control chart. But for such charts it is rather difficult to find the α risk involved.

3. The Q test does not require use of an asymptotic distribution, as does Bartlett's test. Thus, the critical values of the Q test are available all the way down to variances with but a single degree of freedom.

4. The Q test does not become inapplicable if a variance should be zero, as does Bartlett's or Hartley's tests.

5. There is evidence in the literature that in ANOVA one or two cells with high σ_ε^2 are much more disturbing to the analysis than one or two cells with excessively low σ_ε^2. The Q test is more sensitive to the former and less sensitive to the latter than is Bartlett's test. A single very low s_j^2 can have a very marked effect on Bartlett's test, yet but little on the Q test. Thus the Q test would seem to be more what is needed in ANOVA analyses.

6. As to power, the variety of forms the alternate hypothesis H_1 can take, make power studies difficult; not to mention the mathematical difficulties involved. But it would seem that in many common situations those tests using all of the s_j^2's such as the Q test, Bartlett's test and the L tests would give more reliable decisions, than those depending upon one or both extreme s_j^2 such as Cochran's or Hartley's tests.

References

1. Eisenhart, C., 1947. The assumptions underlying the analysis of variance, Biometrics, 3, 1-21.

2. Cochran, W. G., 1947. Some consequences when the assumptions for the analysis of variance are not satisfied, Biometrics, 3, 22-38.

3. David, F. N. and Johnson, N. L., 1951. The effect of non-normality on the power function of the F-test in the analysis of variance, Biometrika, 38, 43-57.

4. Box, G.E.P., 1954. Some theorems on quadratic forms applied in the study of analysis of variance problems. I. Effect of inequality of variance on one-way classification, Ann. Math. Stat. 25, 290-302.

5. Box, G.E.P., 1954. 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, Ann. Math. Stat., 25, 484-498.

6. Brandt, A. E., 1932. A thesis submitted for the degree of doctor of philosophy, Iowa State College, Ames, Iowa.

7. Snedecor, G. W., 1935. Statistical Methods, Collegiate Press, Inc. Ames, Iowa.

8. Stevens, W. L., 1936. Heterogeneity of a set of variances, Jour. of Genetics, 33, 398-399. An appendix to a paper by A. C. Fabergé.

9. Neyman, J. and Pearson, E. S., 1931. On the problem of k samples, Bull. Acad. Polonaise des Sci. at des Lett., Series A. 460-481.

10. Bartlett, M. S., 1937. Properties of sufficiency and statistical tests, Proc. Roy. Soc., A, 160, 268-282.

11. Wilks, S. S., 1937. The sampling distribution of the criterion λH_1 when the hypothesis tested is not true, Biometrika, 29, 124-7.

12. Thompson, C. M., 1937. An investigation into the adequacy of Dr. Wilks' curves, Biometrika, 29, 127-132.

13. Nair, U. S., 1939. The application of the moment function in the study of distribution laws in statistics, Biometrika, 30, 274-294.

14. Cochran, W. G. 1941. The distribution of the largest of a set of variances as a fraction of their total, Annals of Eugenics, 11, 47-52.

15. Hartley, H. O., 1950. The maximum F-ratio as a short-cut test for heterogeneity of variance, Biometrika, 37, 308-312.

16. Shewhart, W. A., 1931. Economic Control of Quality of a Manufactured Product, D. Van Nostrand Co., New York.

17. Cohen, A. and Strawderman, W. E., 1971. Unbiasedness of tests for homogeneity of variances, Ann. Math. Stat., 42, 355-60.

18. Foster, Louis A., 1964. Testing for Equality of Variances, Ph.D. Thesis, Purdue University.

19. Burr, I. W., 1942. Cumulative frequency functions, Ann. Math. Stat., 13, 215-232.

20. Burr, I. W. and Cislak, P. J., 1968. On a general system of distributions, I. Its curve-shape characteristics, II. The sample median, Jour. Amer. Statl. Assn., 63, 627-635.

21. Pearson, E. S., 1963. Some problems arising in approximating to probability distributions using moments, Biometrika, 50, 95-112.