# TESTS OF SIGNIFICANCE IN FACTOR ANALYSIS

By M. S. BARTLETT

Department of Mathematics, University of Manchester

I. *Description and Illustration of a Test of Significance for Analysis into Principal Components.* II. *Remarks on Lawley's Maximum Likelihood Method.* III. *Further Theoretical Discussion.*

## I. DESCRIPTION AND ILLUSTRATION OF A TEST OF SIGNIFICANCE FOR ANALYSIS INTO PRINCIPAL COMPONENTS

*Introductory Remarks.* In a previous paper (1) in this *Journal* I distinguished between two types of statistical analysis, external canonical or factor analysis between two groups of variables and internal factor analysis of a single group. By the latter was meant the particular statistical analysis of the internal correlation structure known more familiarly as an 'analysis into principal components.' It was shown that the appropriate tests of significance of the 'factors' emerging from the statistical analysis were quite distinct in the two cases ; but while a fairly satisfactory test was available for the external case, no analogous test was given in the internal case, only a method of assessing the significance of the smallest root but one against the smallest. It is the main purpose of the present paper to remedy this deficiency, and describe a more general test analogous to the test developed for the external case.

It should be noticed that an analysis into principal components is distinct from an analysis corresponding to some postulated psychological pattern, and has a different interpretation. The most systematic attempt to cover the latter problem is perhaps that by Lawley (2, 3, 4), using the method of maximum likelihood. A critical discussion of this method is given in Part II, where it is shown further that the associated $\chi^2$ test proposed by Lawley in 1940 has a partial connexion with the $\chi^2$ technique proposed in I. For convenience the theoretical justification of the formulæ and methods described in I (and, to some extent, in II) is deferred to Part III, which may be omitted by the non-mathematical reader.

The present paper does not entirely clear up the problems which it raises for either form of analysis ; but the procedures discussed seem sufficiently justified to be worth reviewing at this stage, even although further theoretical investigation is still needed.

*The $\chi^2$ Approximation.* Let us denote by $\lambda_i$ $(i = 1, 2, \ldots, p)$ the $p$ roots, in descending order of magnitude, of the determinantal equation

$$| R - \lambda | = 0, \qquad (1)$$

where $R$ is a correlation matrix of $p$ variables. It should be noted that for the determinant $| R |$ we have the relation $| R | = \lambda_1 \lambda_2 \ldots \lambda_p$. Let the total number of degrees of freedom available from the original observations be $n$ ; we shall usually have $n = r - 1$, when deviations from the general mean are employed for each variable, where $r$ is the total number of observations for each variable, i.e., in the present context $r$ is the total number of persons tested.[1] The statistical significance

[1] Note that the number of degrees of freedom $s$ of Table I of the previous paper was defined as $r - 2$. This is the usual convention for the test of significance of a single correlation, but the degrees of freedom $n = r - 1$ for each variable is more appropriate here.

of the entire correlation structure may be tested if necessary by calculating the quantity [1]

$$\chi^2 = -\{n - \tfrac{1}{6}(2p + 5)\} \log_e |R| , \qquad (2)$$

with $\frac{1}{2}p(p - 1)$ degrees of freedom. However, the usual situation will be that after the removal of the largest roots we require to test the significance of those left. The $\chi^2$ test may still, as in an external analysis, be used for this purpose, provided that the roots removed correspond to well-determined (i.e., highly significant) factors. We now take

$$\chi^2 = -\{n - \tfrac{1}{6}(2p + 5) - \tfrac{2}{3}k\} \log_e R_{p-k} , \qquad (3)$$

with $\frac{1}{2}(p - k)(p - k - 1)$ degrees of freedom, after $k$ roots $\lambda_1\ \lambda_2 \ldots \lambda_k$ have been determined, where

$$R_{p-k} = |R| \Big/ \left\{ \lambda_1 \lambda_2 \ldots \lambda_k \left[ \frac{p - \lambda_1 - \lambda_2 \ldots - \lambda_k}{p - k} \right]^{p-k} \right\} . \qquad (4)$$

It will be noticed that the test is formulated to indicate the significance of the residual roots ; however, it is often convenient to present a complete $\chi^2$ table, provided this is properly interpreted. Since the multiplying factor in (3) changes with $k$, the complete table is not strictly additive unless a constant factor is substituted ; we might employ the smallest, corresponding to $k = p - 2$ : the factor is then $n - p + \frac{1}{2}$. But if $n$ is not large, the full advantage of the approximation is retained if the proper factor is used for each individual test.

*Numerical Illustrations.* Two examples will be given to illustrate this test. For the first I have taken the data of Truman Kelly's quoted in Hotelling's original discussion (5) of an analysis into principal components, consisting of four tests (reading speed, reading power, arithmetic speed, and arithmetic power) carried out on 140 children. For the matrix $R$ we have [2]

$$\begin{pmatrix} 1 & 0{\cdot}698 & 0{\cdot}264 & 0{\cdot}081 \\ 0{\cdot}698 & 1 & -0{\cdot}061 & 0{\cdot}092 \\ 0{\cdot}264 & -0{\cdot}061 & 1 & 0{\cdot}594 \\ 0{\cdot}081 & 0{\cdot}092 & 0{\cdot}594 & 1 \end{pmatrix}$$

The variances taken up by the four roots were

$$\begin{aligned} \lambda_1 &= 1{\cdot}846 \\ \lambda_2 &= 1{\cdot}465 \\ \lambda_3 &= 0{\cdot}521 \\ \lambda_4 &= 0{\cdot}167 \\ \hline \text{Total} &= 3{\cdot}999 \end{aligned} \qquad \begin{aligned} |R| &= \lambda_1 \lambda_2 \lambda_3 \lambda_4 \\ &= 0{\cdot}2353 \end{aligned}$$

For the quantities $R_{p-k}$ in (3) we find

$$\begin{aligned} R_2 &= 0{\cdot}521 \times 0{\cdot}167 \times (2/0{\cdot}689)^2 = 0{\cdot}7330 \\ R_3 &= (0{\cdot}2353/1{\cdot}846) \times (3/2{\cdot}154)^3 = 0{\cdot}3444 \\ R_4 &= |R| = 0{\cdot}2353 . \end{aligned}$$

In practice when the smallest roots have not been evaluated $|R|$ must of course be evaluated directly. Using the common factor $n - p + \frac{1}{2} = 139 - 4 + \frac{1}{2} = 135{\cdot}5$ , we obtain the additive $\chi^2$ analysis :

[1] As noted in Part III, the criterion $|R|$ was considered by Wilks in 1932 ; the $\chi^2$ approximation proposed in equation (2) is closely connected with the $\chi^2$ approximation I have put forward for external analysis (see also Box (11) p. 344).

[2] In order to make use of Hotelling's analysis for this first example these correlations, corrected for attenuation, have been used, though of course this correction will affect test of significance to some (unknown) extent which is here ignored.

| | D.F. | $\chi^2$ |
|---|---|---|
| $\lambda_1$ | 3 | $-135{\cdot}5\ (\log_e 0{\cdot}2353 - \log_e 0{\cdot}3444) = 51{\cdot}6$ |
| $\lambda_2$ | 2 | $-135{\cdot}5\ (\log_e 0{\cdot}3444 - \log_e 0{\cdot}7330) = 102{\cdot}3$ |
| $\lambda_3$ | 1 | $-135{\cdot}5\ (\log_e 0{\cdot}7330) = 42{\cdot}1$ |
| $\lambda_4$ | — | |
| Total .. | 6 | $-135{\cdot}5\ (\log_e 0{\cdot}2353) = 196{\cdot}0$ |

The best approximation for the total $\chi^2$ would actually be $-\{n - \frac{1}{6}(2p + 5)\} \log_e \mid R \mid = -136{\cdot}833\ (\log_e 0{\cdot}2353) = 198{\cdot}0$, but of course all the components are so highly significant that this difference is unimportant. It may be wondered why the second largest component gives an apparently more significant $\chi^2$ than the first; but it should be remembered that logically the significance of at least one root is to be judged from the total $\chi^2$, not the component corresponding to $\lambda_1$. The apparent significance of this component according to the above table represents the magnitude of the root compared with the variation still remaining, and the latter contains quite a large second component. On the other hand, the apparent significance of the latter according to the analysis of the table is judged in relation to the residual variation after this component has in turn been removed, and this residual variation is comparatively small.

As the second example I have taken the hypothetical data discussed by Burt (6), for which the number of persons is only 6. It is indicated in the theoretical discussion in III that we may expect the test presented here to work satisfactorily, at least for the total $\chi^2$, for as low as about 10 degrees of freedom.[1] This means that the test must be used with considerable caution in the present example, which is unrealistic in its very small number of persons. In view of the small value of $n$, the $\chi^2$ table for this example will be presented not in additive form, but in the form in which it would arise as each set of residual factors is tested in turn for significance, the best multiplying factor given in (3) being used in each case.

From Table XIII of Burt's paper we have

$$\begin{aligned} \lambda_1 &= 3{\cdot}1021 \\ \lambda_2 &= 0{\cdot}7328 \\ \lambda_3 &= 0{\cdot}1593 \\ \lambda_4 &= 0{\cdot}0058 \\ \hline \text{Total} &= 4{\cdot}0000 \end{aligned} \qquad \begin{aligned} \mid R \mid &= \lambda_1\,\lambda_2\,\lambda_3\,\lambda_4 \\ &= 0{\cdot}002101 \end{aligned}$$

Hence

$$\begin{aligned} R_2 &= 0{\cdot}1593 \times 0{\cdot}0058 \times (2/0{\cdot}1651)^2 = 0{\cdot}1356 \\ R_3 &= (0{\cdot}002101/3{\cdot}1021) \times (3/0{\cdot}8979)^3 = 0{\cdot}02526 \\ R_4 &= \mid R \mid = 0{\cdot}002101\ . \end{aligned}$$

The $\chi^2$ test for the residual roots is thus:

| | D.F. | $\chi^2$ | $P = 0{\cdot}05$ level |
|---|---|---|---|
| $\lambda_1\ \lambda_2\ \lambda_3$ | 3+2+1 | $-2{\cdot}833\ (\log_e 0{\cdot}002101) = 17{\cdot}47$ | 12·59 |
| $\lambda_2\ \lambda_3$ | 2+1 | $-2{\cdot}167\ (\log_e 0{\cdot}02526) = 7{\cdot}97$ | 7·82 |
| $\lambda_3$ | 1 | $-1{\cdot}5\ (\log_e 0{\cdot}1356) = 3{\cdot}00$ | 3·84 |

[1] By these degrees of freedom is meant of course the number depending on the number of persons, *not* the degrees of freedom for $\chi^2$, depending only on the number of tests.

Even with such a small sample we are thus able to conclude that the first two factors are significant ; the third, however, does not reach significance (cf. the conclusion in the final paragraph of my previous paper (1) ).

*Further Comments.* It need hardly be stressed that I am concerned here with *statistical* significance. It does not follow that all the factors which reach statistical significance in a large sample necessarily remove a very large fraction of the variance ; and hence some of them may be comparatively unimportant in practice. Again, even if they are numerically important, this has no necessary implications of psychological or other reality of the factors. Merely the correlation structure of the variables is being investigated in its relation to variance : for this reason no significance can ever be attached to the last root, for it would be equivalent to asking for the correlation structure of a single variable.

If of course the basic equation (1) is replaced by another, then the above tests become inappropriate ; for example, if we ask for the significance of the factors of an equation like

$$| R - \lambda e_{ii} | = 0 , \tag{5}$$

where $e_{ii}$ denotes error variances (assumed known), then, as I showed in (1), the analysis is a special case of an external factor analysis, and other tests become relevant. Or if in place of the empirical analysis of equation (1) we substitute a factor analysis of more orthodox psychological type, then, as noted in the introductory remarks, other tests again will be required. Some discussion is given in Part II of a factor analysis of standard type, which in vector and matrix notation is represented by the equation

$$T = M_0F_0 + M_1F_1 , \tag{6}$$

where $T$ represents the vector of standardized tests, $F_0$ the set of (standardized) general or group factors, $F_1$ the (standardized) specifics, and $(M_0 , M_1)$ the factor 'loadings.' Since the difference between the factors of equation (6) and those corresponding to the analysis (1) is not always realized, it may be advisable to illustrate this difference by a simple example. Consider three tests of which the first two are correlated with coefficient $\rho$ (positive), and the third is entirely independent. On the basis of equation (6) this might be interpreted as indicating one 'general' factor common to the first two tests, but with zero loading for the third, the remaining factors being specific to each test. On the basis of an analysis into principal components, however, it is easily seen that the roots are $1 + \rho$ , 1 and $1 - \rho$ , and it will be noticed that in order of contribution to total variance it is the second root which corresponds to the uncorrelated test and, for any reasonable size of sample, this will be significant compared with the last. Thus once significant correlation is present it may automatically cause all the roots to be distinct in a principal components analysis, which should not be assumed to lead necessarily to the *simplest* factor-pattern.

## II. REMARKS ON LAWLEY'S MAXIMUM LIKELIHOOD METHOD

*The Maximum Likelihood Method in Factor Analysis.* To determine the structural factor equation of the type represented by equation (6), Lawley has investigated two methods based on the principle of maximum likelihood. The first, which he has called 'Method I,' is based on the variance-covariance sampling distribution of the test scores ; the second, 'Method II,' is based more directly on the individual test scores (see his paper (3) ). At first sight the second method would appear the more fundamental, but it is known to lead to difficulties. In a recent paper (7) M. G. Kendall raised doubts on the suitability of the maximum likelihood method in the present context ; and Dr. Lawley has informed me that he had already entertained similar doubts, at least as regards 'Method II.' In the discussion following Kendall's paper, I pointed out that the factor equation (6) is equivalent to another set of structural equations (familiar in econometrics) for which it is definitely known that the standard maximum likelihood method breaks down.

The question remains of the validity and status of Lawley's 'Method I,' and a detailed re-examination and formulation of the most appropriate statistical procedure for dealing with the structural equation (6) seems still required. Some provisional comments on Lawley's first method may, however, be useful.

*Lawley's 'Method I,' and its Associated $\chi^2$ Test.* I was at first sceptical of the value of this method because of its obvious breakdown in elementary cases. For example, it does not work for the case of two tests and one general factor, for we should then have three observed quantities, the two variances and the covariance, and four unknown loading coefficients (or equivalently, when the tests have been standardized, one observed correlation, and two unknown loading coefficients for the general factor) ; in other fields this case often arises in the guise of fitting a straight line to the plot of two variables both of which are subject to error (see, for example, Bartlett (8) ).

In psychology the number of tests is usually large enough for this difficulty not to arise. The validity of the maximum likelihood method applied to the variance-covariance distribution then depends (i) on our deciding to consider only the information contained in the variance-covariance (or equivalent correlation) matrix, and (ii) on Lawley's procedure, given (i), providing efficient estimates. Since the general validity of the maximum likelihood approach can no longer be assumed, (ii) is no longer automatic, and appears to require demonstration. One possible argument is to note that enough factors and unknown coefficients can always be introduced to take up all the available degrees of freedom and yet avoid redundancy ; representation of the correlation matrix by still fewer parameters merely implies further linear restrictions among the unknown coefficients, and this should hardly affect the question of efficiency when the likelihood method is used.

But it might be noticed that the method depends on the correlation matrix being compatible with the postulated factor-pattern. For example, with only three tests it is not possible to have a non-redundant factor equation (6) except in the case of only one general factor, which should take up all the degrees of freedom. Yet the correlation matrix

$$\begin{pmatrix} 1 & 0{\cdot}60 & -0{\cdot}28 \\ 0{\cdot}60 & 1 & 0{\cdot}60 \\ -0{\cdot}28 & 0{\cdot}60 & 1 \end{pmatrix}$$

is only compatible with a minimum of two general factors. This would have to be revealed by equations impossible to satisfy with real coefficients, and not by Lawley's associated $\chi^2$ test. If the above matrix arose in a sample, it is not clear how we should test the goodness of fit of one general factor. If we ignore such difficulties, Lawley's method appears to provide a satisfactorily objective procedure for dealing with equation (6).

It has been stressed that this factor analysis is quite distinct from that discussed in Part I of this paper ; but it is an attractive link between the two analyses that the total $\chi^2$ corresponding to the significance of the unreduced correlation matrix is necessarily the same, and only because of the difference between the factors extracted in the two analyses does the analysis of the total $\chi^2$ into its respective components differ.

It follows (see the last section of Part III of this paper) that the form of Lawley's test which should be more exact for moderate-sized samples is the determinantal form

$$\chi^2 \sim -n' \log_e | R | / | \hat{R} | ,$$

where $| \hat{R} |$ denotes the estimated correlation matrix when $k$ general factors (and $p$ specifics) have been fitted, and the multiplying coefficient $n'$ be taken, not to be $n$, but the coefficient $n - \frac{1}{6}(2p + 5)$ used in (2), or perhaps even better, the coefficient $n - \frac{1}{6}(2p + 5) - \frac{2}{3}k$ used in (3). The number of degrees of freedom of $\chi^2$ is $\frac{1}{2}(p - k)(p - k - 1) - k$ (subject, of course, to the difficulties raised above).

## III. FURTHER THEORETICAL DISCUSSION

*Equivalent Analyses of Correlation Structure.* Any résemblance of the test described in I above to the test for external analysis described in the earlier paper is more than superficial. Let us consider the logical relations among a group of variables, say, four for definiteness. The possible correlations may be considered in more than one way. Thus we may consider the correlation between variables $x_1$ and $x_2$, then the relation between $x_3$ and the group $x_1$, $x_2$, then between $x_4$ and the group $x_1$, $x_2$, $x_3$. As another alternative, we might consider the external relation between the two groups $x_1$, $x_2$ and $x_3$, $x_4$ ; then the further internal relations between $x_1$ and $x_2$ in the first group, and between $x_3$ and $x_4$ in the second. The internal relations in a single group can thus be regarded as a number of 'external' relations between various subgroups. Hence a test of the internal relations can be built up from the tests of the equivalent external relations between sub-groups. If we carry out this theoretical programme according, say, to the first of the alternative analyses mentioned above, where the relation of each further variable is considered successively, we arrive from the theory of the $\chi^2$ test for external analysis (9) at the successive $\chi^2$ quantities :

$$\begin{array}{ll} -\{n - \frac{1}{2}(3)\} \log_e (1 - r_{12}^2) , & (1 \text{ d.f.}) , \\ -\{n - \frac{1}{2}(4)\} \log_e (1 - R_{3\,,\,12}^2) , & (2 \text{ d.f.}) , \\ \vdots & \\ -\{n - \frac{1}{2}(p + 1)\} \log_e (1 - R_{p\,,\,12 \ldots p-1}^2) , & (p - 1 \text{ d.f.}) , \end{array}$$

where $r_{12}$ is the observed correlation between $x_1$ and $x_2$, $R_{p\,,\,12 \ldots p-1}$ the multiple correlation of $x_p$ with $x_1 , x_2 , \ldots , x_{p-1}$. In order to make any test arrived at invariant to the route of arrival, it is obviously convenient to choose a mean multiplying factor, which (weighting with the degrees of freedom) becomes $n - \frac{1}{2}f$, where

$$\tfrac{1}{2}p(p - 1)f = (1 \times 3) + (2 \times 4) + (3 \times 5) + \ldots (p - 1)(p + 1) ,$$

or

$$f = \tfrac{1}{3}(2p + 5) .$$

It is well known that

$$(1 - r_{12}^2)(1 - R_{3\,,\,12}^2) \ldots (1 - R_{p,\,12 \ldots p-1}^2) = | R | ,$$

whence we obtain as an approximate $\chi^2$, with $1 + 2 + \ldots + (p - 1) = \frac{1}{2}p(p - 1)$ degrees of freedom, the expression proposed in equation (2).

*Direct Derivation of the $\chi^2$ Approximation.* The above method of approach was the one I used first to arrive at (2) ; but it is satisfactory that the test finally arrived at depends on the correlation determinant $| R |$, a criterion considered many years ago by Wilks (10). It is useful to check directly the appropriateness of the $\chi^2$ approximation by writing down the moment-generating function of $-n \log_e | R |$ , by suitably adapting Wilks' formula for the $s$th order moment of $|R|$ (cf. the derivation in (9) of the $\chi^2$ test for an external analysis). Alternatively the moment-generating function could be written down from the simultaneous distribution of the correlations $R$, or from the independence of the components listed above. We obtain

$$M(t) = \frac{\Gamma^{p-1}(\frac{1}{2}n) \prod_{i=1}^{p-1} \Gamma\{\frac{1}{2}(n-i) - nt\}}{\Gamma^{p-1}(\frac{1}{2}n - nt) \prod_{i=1}^{p-1} \Gamma\{\frac{1}{2}(n-i)\}}, \tag{7}$$

and it is a matter of straightforward algebra coupled with Stirling's approximation formula for a $\Gamma$-function to show that

$$M(t) = \sum_{i=1}^{p-1}\left\{t\left[i + \frac{i^2+2i}{2n} + \ldots\right] + t^2\left[i + \frac{i^2+2i}{n} + \ldots\right] + \frac{4}{3}t^3\left[i + \frac{3(i^2+2i)}{2n} + \ldots\right] + \ldots\right\}$$

$$= \tfrac{1}{2}p(p-1)\,t\left[1 - \frac{1}{6n}(2p+5)\right] + \tfrac{1}{2}p(p-1)\,t^2\left[1 - \frac{1}{6n}(2p+5)\right]^2$$

$$+ \tfrac{1}{2}p(p-1)\frac{8t^3}{3!}\left[1 - \frac{1}{6n}(2p+5)\right]^3 + \ldots \tag{8}$$

to $0\left(t^3, \frac{1}{n}\right)$. Hence the improved approximation is

$$\chi^2 \sim -\{n - \tfrac{1}{6}(2p+5)\}\log_e |R| ,$$

confirming the mean multiplying factor arrived at by the previous argument, and from (8) having $\frac{1}{2}p(p-1)$ degrees of freedom.[1]

*Closeness of the $\chi^2$ Approximation.* The order of closeness of this approximation will clearly be comparable to that for an external analysis, in view of their relationship, and will consequently be expected to work quite well down to a sample with about 10 degrees of freedom (9). Let us, for example, in the case of only two variables, where merely the test of a single correlation is involved, check the closeness of the approximation. We find for $\chi^2$, corresponding to the correct critical $P = 0{\cdot}05$ or $0{\cdot}01$ value of $r_{12}$, the values :

| | $P = 0{\cdot}05$ | $P = 0{\cdot}01$ |
|---|---|---|
| $n = 10$ | 3·84 | 6·60 |
| $n = 20$ | 3·84 | 6·63 |
| $n = \infty$ | 3·841 | 6·635 |

*Effect of Eliminating the Larger Roots.* The approximate effect of eliminating the larger roots in the factor analysis has next to be considered. To some extent it is possible to give an argument analogous to that used for an external analysis, when the roots eliminated are well determined. In this case the orthogonal directions corresponding to these roots in the geometrical representation of the variables in the $n$-dimensional sample space will be well determined, and we may consider the analysis of the remaining $p - k$ roots in a space of $n - k$ dimensions (if $k$ roots have been eliminated).

The adjustment of the multiplying factor gives

$$(n-k) - \tfrac{1}{6}\{2(p-k)+5\} = n - \tfrac{1}{6}(2p+5) - \tfrac{2}{3}k ,$$

as given in (3). The expression for $R_{p-k}$ given in (3) is merely a convenient method of obtaining the expression in the remaining roots analogous to the determinant $|R| = \lambda_1 \lambda_2 \ldots \lambda_p$, when we are given $|R|$ and $\lambda_1 \lambda_2 \ldots \lambda_k$. It should be remembered that significance in an internal analysis can only be on the basis of the relative values of the roots, which must after each elimination be in effect re-scaled to unit mean variance. This explains the factor $[(p-k)/(\lambda_{k+1} + \ldots \lambda_p)]^{p-k}$ in $R_{p-k}$, which may thus be written

$$\frac{|R|}{\lambda_1 \lambda_2 \ldots \lambda_k}\left\{\frac{p-k}{p - \lambda_1 - \lambda_2 \ldots - \lambda_k}\right\}^{p-k}.$$

[1] An independent and prior mention of this approximation is included in the recent comprehensive paper on such approximations by Box (11).

It must be admitted that the above justification of the proposed test for the residual roots is not as complete as could be wished. The problem would be slightly simpler if we were alternatively considering the analysis of test scores known to have true equal variances, but standardized to unity only for the *mean* variance. In this case the above form of the test for the residual roots can be given further justification, but of course the total $\chi^2$ now has $p - 1$ more degrees of freedom, these representing the test of homogeneity of the $p$ test variances about their unit mean variance. Correspondingly, the total $\chi^2$ for the last $p - k$ factors would have (approximately) $\frac{1}{2}(p - k - 1)(p - k + 2)$ degrees of freedom instead of $\frac{1}{2}(p - k)(p - k - 1)$. One point that remains in some doubt pending a more detailed examination is whether the reduction in degrees of freedom that ensues from the individual standardization of the tests is automatically felt in the residual factor components (an assumption implicit in the proposed test), or is mainly absorbed by the larger roots.

*The Alternative Forms of Lawley's $\chi^2$ Test.* In his 1940 paper (2), Lawley gave two alternative forms for the large-sample $\chi^2$ test he developed for use with his maximum likelihood method (Method I). One was an arithmetically convenient sum of squares, but it is worth pointing out that the equivalent determinant formula arises more exactly than he perhaps emphasized, and should therefore be the more accurate one for moderate-sized samples.

To demonstrate this, it is convenient to define a 'homogeneous' likelihood function from the variance-covariance sampling distribution by its logarithm $L$, where

$$-2L + \text{constant} = n \log | C | - n \log | A | + n \operatorname{trace} (C^{-1}A) . \quad (9)$$

In this expression, where $C$ denotes the true variance-covariance matrix and $A$ the corresponding sample matrix, $-2L$ tends to $\chi^2$ with $\frac{1}{2}p(p + 1)$ degrees of freedom as $n$ increases. To determine the constant in (9) we substitute the estimates $A$ for $C$, giving

$$\text{constant} = n \operatorname{trace} (A^{-1}A) = np . \quad (10)$$

If we alternatively substitute estimates of $C$ from (efficient) factor estimates, with $q$ independent unknown coefficients, we obtain

$$-2\hat{L} + \text{constant} = n \log | \hat{C} | - n \log | A | + n \operatorname{trace} (\hat{C}^{-1}A) , \quad (11)$$

with $\frac{1}{2} p(p + 1) - q$ degrees of freedom.

From (10) and (11),

$$\begin{aligned} -2\hat{L} &= -n \log | A | / | \hat{C} | + n \operatorname{trace} (\hat{C}^{-1}A) - np \\ &= -n \log | A | / | \hat{C} | + n \operatorname{trace} (\hat{C}^{-1}A - \hat{C}^{-1}\hat{C}) \\ &= -n \log | A | / | \hat{C} | , \end{aligned} \quad (12)$$

the last term vanishing when $\hat{C}$ corresponds to the maximum likelihood estimates (see Lawley (2) ).

If no general factors are fitted, the estimates $\hat{C}$ consist simply of the specific factor variances coinciding with the test variances, and (12) reduces to $-n \log | R |$. It is then, as we should expect, identical with the total $\chi^2$ proposed in the principal components analysis, and the modified multiplying coefficient $n - \frac{1}{6}(2p + 5)$ is consequently more accurate. When $k$ general factors have also been fitted, (12) may be expressed in terms of the ratio of the observed and estimated *correlation* determinants. The more accurate multiplying coefficient $n - \frac{1}{6}(2p + 5)$ should still be more appropriate than the crude coefficient $n$, and the further modified coefficient $n - \frac{1}{6}(2p + 5) - \frac{2}{3}k$ would have some justification here also, though this last refinement is of course not very firmly established. The number of degrees of freedom comes out at $\frac{1}{2}(p - k)(p - k - 1) - k$, the number lost per component being one more than in the principal components analysis.

Finally, it might be noticed that an alternative analysis of the total $\chi^2$ along the lines given in the first section of Part III would sometimes be of practical interest, and since, as was there pointed out, it would then rest on the known $\chi^2$ approximation for *external* analyses, such an analysis would have a somewhat more precise justification than it has yet been found possible to give either for the principal components analysis or for the maximum likelihood factor analysis.

## REFERENCES


1. Bartlett, M. S. (1948). 'Internal and external factor analysis.' *Brit. J. Psych., Stat. Sect.*, I, 73.
2. Lawley, D. N. (1940). 'The estimation of factor loadings by the method of maximum likelihood.' *Proc. Roy. Soc. Edin.*, IX, 64.
3. Lawley, D. N. (1942). 'Further investigations in factor estimation.' *Proc. Roy. Soc. Edin.*, LXI, 176.
4. Lawley, D. N. (1949). 'Problems in factor analysis.' *Proc. Roy. Soc. Edin.*, LXII, 1.
5. Hotelling, H. (1933). 'Analysis of a complex of statistical variables into principal components.' *J. Educ. Psych.*, XXIV, 417 and 498.
6. Burt, C. (1947). 'Factor analysis and analysis of variance.' *Brit. J. Psych., Stat. Sect.*, I, 3.
7. Kendall, M. G. (1950). 'Factor analysis as a statistical technique.' Symposium on factor analysis to be published in *J. Roy. Statist. Soc.* (Series B).
8. Bartlett, M. S. (1949). 'Fitting a straight line when both variables are subject to error.' *Biometrics*, 5, 207.
9. Bartlett, M. S. (1938). 'Further aspects of the theory of multiple regression.' *Proc. Camb. Phil. Soc.*, XXXIV, 33.
10. Wilks, S. S. (1932). 'Certain generalizations in analysis of variance.' *Biometrika*, XXIV, 471.
11. Box, G. E. P. (1949). 'A general distribution theory for a class of likelihood criteria.' *Biometrika*, XXXVI, 317.