DOCUMENT RESUME

ED 078 053 TM 002 868

AUTHOR	Jennrich, Robert I.; Thayer, Dorothy T.
TITLE	A Note on Lawley's Formulas for Standard Errors in Maximum Likelihood Factor Analysis.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RB-73-31
PUB DATE	Apr 73
NOTE	19p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Error Patterns; *Factor Analysis; *Statistical Analysis; Tables (Data); Technical Reports
IDENTIFIERS	*Lawleys Formulas

ABSTRACT

Evidence is given to indicate that Lawley's formulas for the standard errors of maximum likelihood loading estimates do not produce exact asymptotic results. A small modification is derived which appears to eliminate this difficulty. (Author)

RESEARCH BULLETIN

RB-73-31

A NOTE ON LAWLEY'S FORMULAS FOR STANDARD ERRORS IN MAXIMUM LIKELIHOOD FACTOR ANALYSIS

Robert I. Jennrich

and

Dorothy T. Thayer

Educational Testing Service

Princeton, New Jersey

April 1973

A NOTE ON LAWLEY'S FORMULAS FOR STANDARD ERRORS IN MAXIMUM LIKELIHOOD FACTOR ANALYSIS

Abstract

Evidence is given to indicate that Lawley's formulas for the standard errors of maximum likelihood loading estimates do not produce exact asymptotic results. A small modification is derived which appears to eliminate this difficulty.

A NOTE ON LAWLEY'S FORMULAS FOR STANDARD ERRORS IN
MAXIMUM LIKELIHOOD FACTOR ANALYSIS

1. Introduction

In two important papers Lawley [1953, 1967] derived formulas for the standard errors of factor loading estimates produced in maximum likelihood factor analysis. Recent work [Archer & Jennrich, 1973; Jennrich, 1973a] has made it possible to use Lawley's results to obtain standard errors for analytically rotated loadings as well. There is some evidence, however, that Lawley's formulas do not produce exact asymptotic results. We shall attempt to demonstrate that this is in fact the case and propose a modification which will make Lawley's results asymptotically exact. This modification amounts to the insertion of a single already defined symbol into his formulas. While this change may produce little practical effect in standard error estimates computed from real data, the modification is necessary to eliminate discrepancies which arise when Lawley's formulas are used in conjunction with asymptotic results from other sources.

The evidence that there is a problem is quite simple. Table 1 contains standard errors for maximum likelihood loading estimates arising in an example given by Lawley and Maxwell [1971, p. 63] and obtained through the use of Lawley's formulas. The authors have recomputed these values and obtained results which agree with those of Lawley and Maxwell to within one digit in the last decimal place presented. This makes us confident that both we and they have implemented the required formulas correctly.

Insert Table 1 about here

Table 2 contains asymptotic standard errors for the same example which led to Table 1, but which were obtained by inverting the appropriate augmented information matrix [Jennrich, 1973b]. The differences in these two tables, while not large, demonstrate clearly that they cannot both represent true asymptotic standard errors.

Insert Table 2 about here

Finally, Table 3 contains standard errors for the same example using Lawley's formulas with the modification proposed in Section 3. The nearly perfect agreement between Tables 2 and 3 is strong evidence that with this modification, exact standard errors are obtained through the use of Lawley's formulas.

Insert Table 3 about here

Because Lawley's formulas are more efficient than those of Jennrich [1973b] and avoid the inversion of a large matrix, the authors feel this modification is important whenever exact asymptotic standard errors are required.

2. Locating the Difficulty

Lawley originally [1953] derived his standard error formulas under the assumption that the unique variances ψ in a factor analytic decomposition

$$(1) \qquad \Sigma = \Lambda\Lambda' + \psi$$

of a population covariance matrix Σ were known. The formulas were later [1967] generalized to the case of unknown unique variances. The difficulty lies in this latter article. The results there are derived in greater detail in Lawley's text [1971] co-authored with Maxwell. Because of this and because the text contains extensions of the original results to cover the effect of standardization, we shall base our comments on the 1971 text ratner than on the 1967 article. Unless otherwise indicated we shall use the notation and definitions of the text.

On page 53 Lawley and Maxwell define functions $g_{ir}(\psi,S)$ to be the right-hand side of (5.15) in the text divided by $(\theta_r - 1)$ regarded as a function of ψ and a sample covariance matrix S. They assert that (5.15) becomes (5.21) which is

$$(2) \qquad \hat{\lambda}_{iro} - \lambda_{ir} = g_{ir}(\psi,S) \quad .$$

This like (5.15), however, is only an approximate equality. Here $\hat{\lambda}_{iro}$ denotes the maximum likelihood estimate of the population loading λ_{ir} computed under the assumption that ψ is known. To computed standard errors in the unrestricted case Lawley and Maxwell introduce coefficients

$$(3) \qquad b_{j,ir} = \operatorname*{prob\,lim}_{n \to \infty} \frac{\partial g_{ir}}{\partial \psi_j} \quad ,$$

and express them in terms of population parameters by means of equation (5.27) of the text which is

$$(4) \qquad b_{j,ir} = -\lambda_{jr}(\theta_r-1)^{-1}\psi_j^{-2}[\delta_{ij}\psi_j - \tfrac{1}{2}\lambda_{ir}\lambda_{ij}/(\theta_r - 1)$$

$$+ \Sigma_m'\lambda_{im}\lambda_{jm}/(\theta_r - \theta_m)] \quad .$$

Here θ_r denotes the r -th largest eigenvalue of $\psi^{-\frac{1}{2}}\Sigma\psi^{-\frac{1}{2}}$, δ_{ij} denotes the Kronecker delta, and the symbol Σ_m' denotes summation on m from 1 to the number of factors k skipping the value r . Because of the approximation in (2) we cannot agree that (4), when used in (5.29) of the text, gives asymptotic covariances for the maximum likelihood estimates $\hat{\lambda}_{ir}$ of the population loadings λ_{ir} . In order to isolate the problem let $\tilde{g}_{ir}(\psi,S)$ be the function which maps ψ and S into $\hat{\lambda}_{iro}$ minus λ_{ir} . Then replacing g_{ir} by $\tilde{g}_{ir}$ makes the approximate equality in (2) an exact equality. Similarly let $\tilde{b}_{j,ir}$ be the value obtained when g_{ir} is replaced by $\tilde{g}_{ir}$ in the definition of $b_{j,ir}$. As will be seen in the next section these modifications lead to a formula for $\tilde{b}_{j,ir}$ which differs slightly from that for $b_{j,ir}$. When, however, $\tilde{b}_{j,ir}$ replaces $b_{j,ir}$ in formula (5.29) of the text the modified standard error formulas appear to give exact standard errors as observed in the previous section.

3. Deriving the Required Modification

Proceeding here with exact equalities only,

(5) $$\hat{\lambda}_{iro} - \lambda_{ir} = \tilde{g}_{ir}(\psi, S)$$

and

(6) $$\tilde{b}_{j,ir} = \underset{n \to \infty}{\text{prob lim}} \frac{\partial \tilde{g}_{ir}}{\partial \psi_j} \quad .$$

Since Σ is the probability limit of S ,

(7) $$\frac{\partial \tilde{g}_{ir}}{\partial \psi_j} \to \left. \frac{\partial \tilde{g}_{ir}}{\partial \psi_j} \right|_{(\psi,\Sigma)}$$

in probability as $n \to \infty$. Consequently (6) becomes

(8) $$\tilde{b}_{j,ir} = \left. \frac{\partial \tilde{g}_{ir}}{\partial \psi_j} \right|_{(\psi,\Sigma)} \quad .$$

What is required are the values of the partial derivatives $\partial \tilde{g}_{ir} / \partial \psi_j$ at $(\psi, S) = (\psi, \Sigma)$. These are a little difficult to obtain, but implicit differentiation seems to work nicely.

Relations (5.21) and (4.9) from the text and the fact that $\Lambda' \psi^{-1} \Lambda$ is by assumption (p. 27) diagonal may be expressed in the form:

(9) $$\hat{\Lambda}_0 - \Lambda = \tilde{g}(\psi, S)$$

(10) $$(S - \hat{\Lambda}_0 \hat{\Lambda}_0' - \psi)\psi^{-1}\hat{\Lambda}_0 = 0$$

(11) $$\text{non-diag } \hat{\Lambda}_0' \psi^{-1} \hat{\Lambda}_0 = 0 \quad .$$

In order to compute the partial derivatives $\partial\tilde{g}_{ir}/\partial\psi_j$ at (ψ,Σ) by implicit differentiation, we differentiate (10) and (11) holding S fixed at Σ and evaluate the result at $(\psi,S,\hat{\Lambda}_0) = (\psi,\Sigma,\Lambda)$. This leads to the differential relations:

$$(12) \qquad (\Lambda d\Lambda' + d\Lambda\Lambda' + d\psi)\psi^{-1}\Lambda = 0$$

$$(13) \qquad \text{non-diag}(d\Lambda'\psi^{-1}\Lambda + \Lambda\psi^{-1}d\Lambda - \Lambda'\psi^{-1}d\psi\psi^{-1}\Lambda) = 0 \quad .$$

Contributions from the second and third factors in (10) disappear since the first term is zero when $(\psi,S,\hat{\Lambda}_0) = (\psi,\Sigma,\Lambda)$. The required partial derivatives are found by solving (12) and (13) for $d\Lambda$ in terms of $d\psi$. This may be accomplished by algebraic manipulations similar to those which led to (5.15) in the text.

Letting

$$(14) \qquad \Delta = \Lambda'\psi^{-1}\Lambda$$

which is a diagonal matrix and multiplying (12) on the left by $\Lambda'\psi^{-1}$ gives

$$(15) \qquad \Delta d\Lambda'\psi^{-1}\Lambda + \Lambda'\psi^{-1}d\Lambda\Delta + \Lambda'\psi^{-1}d\psi\psi^{-1}\Lambda = 0 \quad .$$

Multiplying (13) on the right by Δ and subtracting the result from the non-diagonal part of (15) gives

$$(16) \qquad \text{non-diag}(\Delta d\Lambda'\psi^{-1}\Lambda - d\Lambda'\psi^{-1}\Lambda\Delta + \Lambda'\psi^{-1}d\psi\psi^{-1}\Lambda(I + \Delta)) = 0 \quad .$$

Solving (15) and (16) for the diagonal and nondiagonal components of $d\Lambda'\psi^{-1}\Lambda$ respectively gives

$$(17) \qquad (d\Lambda'\psi^{-1}\Lambda)_{rr} = -(2\Delta_r)^{-1}(\Lambda'\psi^{-1}d\psi\psi^{-1}\Lambda)_{rr}$$

and

$$(18) \qquad (d\Lambda'\psi^{-1}\Lambda)_{rs} = -(\Delta_r - \Delta_s)^{-1}(\Lambda'\psi^{-1}d\psi\psi^{-1}\Lambda)_{rs}(1 + \Delta_s)$$

for all $r \neq s$. Here Δ_r denotes the r -th diagonal element of Δ and $(\ \cdot\ )_{rs}$ denotes the element in row r and column s of the expression inside the parentheses. Using (14), equation (12) can be written in the form

$$(19) \qquad d\Lambda\Delta = -d\psi\psi^{-1}\Lambda - \Lambda d\Lambda'\psi^{-1}\Lambda \quad .$$

Writing this in coordinate form and using (17) and (18) gives

$$(20) \qquad \begin{aligned} d\lambda_{ir}\Delta_r &= -(d\psi\psi^{-1}\Lambda)_{ir} - \sum_{m=1}^{k} \lambda_{im}(d\Lambda'\psi^{-1}\Lambda)_{mr} \\ &= -(d\psi\psi^{-1}\Lambda)_{ir} + \lambda_{ir}(2\Delta_r)^{-1}(\Lambda'\psi^{-1}d\psi\psi^{-1}\Lambda)_{rr} \\ &\quad +(1 + \Delta_r) \sum_{m\neq r} \lambda_{im}(\Delta_m - \Delta_r)^{-1}(\Lambda'\psi^{-1}d\psi\psi^{-1}\Lambda)_{mr} \quad . \end{aligned}$$

The partial derivative $\partial\tilde{g}_{ir}/\partial\psi_j$ evaluated at (ψ,Σ) is the value of $d\lambda_{ir}$ corresponding to $d\psi = J(j,j)$ where $J(j,j)$ is the p by p matrix which has a one in its j -th diagonal position and is zero elsewhere. Since $\tilde{b}_{j,ir}$ is the value of this partial derivative, replacing $d\lambda_{ir}$ by $\tilde{b}_{j,ir}$ and $d\psi$ by $J(j,j)$ in (20) gives

(21) $$\tilde{b}_{j,ir}\Delta_r = -\delta_{ij}\psi_j^{-1}\lambda_{jr} + \lambda_{ir}(2\Delta_r)^{-1}\psi_j^{-2}\lambda_{jr}^2 + (1 + \Delta_r)\sum_{m\neq r}\lambda_{im}(\Delta_m - \Delta_r)^{-1}\psi_j^{-2}\lambda_{jm}\lambda_{jr} \quad .$$

We will put this in a form which looks as similar as possible to that of (5.27) of the text. To that end we note that because of (4.12) of the text,

(22) $$\Delta_r = \Theta_r - 1 \quad .$$

Making this replacement in (21) and rearranging a bit gives

(23) $$\tilde{b}_{j,jr} = -\lambda_{jr}(\Theta_r - 1)^{-1}\psi_j^{-2}[\delta_{ij}\psi_j - \tfrac{1}{2}\lambda_{ir}\lambda_{jr}/(\Theta_r - 1) + \Theta_r \sum_m{}' \lambda_{im}\lambda_{jm}/(\Theta_r - \Theta_m)] \quad .$$

The only difference between this and (5.27) which is reproduced in (4) is the appearance of Θ_r immediately preceding the summation symbol.

4. Modification for Standardization

The results of the previous sections dealt with standard errors for maximum likelihood estimates of (natural) factor loadings. It is common in practice to estimate standardized loadings

(24) $$\lambda_{ir}^* = \lambda_{ir}/\sigma_i \quad .$$

Here σ_i denotes the population standard deviation of the i-th score. Maximum likelihood estimates of standardized loadings are computed from a sample correlation matrix in precisely the same way that maximum likelihood estimates for natural loadings are computed from a sample covariance matrix.

Lawley and Maxwell [1971, p. 61] give standard erro standardized loading estimates in addition to those alr the unstandardized case. Since these formulas also invo derivatives $b_{j,ir}$ we expect that they will not give ex results and that they could be made exact if the $b_{j,ir}$ $\tilde{b}_{j,ir}$. This in fact seems to be the case.

Lawley and Maxwell's formulas for standardized loadi two alternative forms. The first, which is based on (5.4 $b_{j,ir}$ explicitly and requires no modification except for of $b_{j,ir}$ by $\tilde{b}_{j,ir}$. The second is based on (5.47) and quired modification involves replacing the -1 in the firs the summation symbol by $-\theta_r$. The modified formulas in the text are

$$(25) \qquad n\, \mathrm{cov}(\hat{\lambda}_{ir}, s_{jj}) = 2 \sum_m (\lambda_{jm} a_{ir,jm}) + 2\psi_j^2 \tilde{b}_{j,ir}$$

and

$$(26) \qquad n\, \mathrm{cov}(\hat{\lambda}_{ir}, s_{jj}) = (\theta_r - 1)^{-1} \lambda_{jr} [2\theta_r \sigma_{ij} - 2\delta_{ij} \psi_i$$

$$+ 2 \sum_s{}' (\theta_r \theta_s - \theta_r)(\theta_r - \theta_s)^{-1}$$

To check these modifications the authors have impleme sions of Lawley and Maxwell's formulas and applied them to discussed in Section 1. Both versions produced results wh within one digit in the last decimal place with those give Maxwell [1971, p. 64]. The latter are reproduced in Table ment suggests again that all parties correctly implemented the text. Table 5 contains the corresponding asymptotic standard errors obtained by inverting the appropriate augmented information matrix [Jennrich, 1973b]. It displays the anticipated lack of agreement with the results in Table 4. Finally, the results in Table 6 were obtained using the first form of the Lawley and Maxwell formulas with the modification discussed in this section. The agreement between Tables 5 and 6, which again is to

----

Insert Tables 4, 5, and 6 about here

----

within one digit in the last decimal place presented, suggests that the proposed modification does produce exact asymptotic standard errors. Using the second version of the Lawley and Maxwell formulas with the proposed modification produced no surprises. The results were identical to those in Table 6.

5. Comments

We have avoided the assertion that Lawley's standard error results are incorrect. To our knowledge it was never claimed that they were exact asymptotic results. What is important here we believe is that we have furnished an almost completely independent verification of a slightly modified form of his formulas. For this the modification played a necessary but not the central role. At present, the whole area of standard errors for factor loading estimates is fairly new and fairly complicated. Since there is little that is as fallible as a mathematical derivation, independent verifications like those presented here are not only comforting but essential.

References

Archer, C. O. & Jennrich, R. I. Standard errors for rotated factor loadings. Research Bulletin 73-17. Princeton, N.J.: Educational Testing Service, 1973.

Jennrich, R. I. Standard errors for obliquely rotated factor loadings. Research Bulletin 73-28. Princeton, N.J.: Educational Testing Service, 1973.

Jennrich, R. I. Simplified formulas for standard errors in maximum likelihood factor analysis. Research Bulletin 73-00. Princeton, N.J.: Educational Testing Service, 1973, in press. (b)

Lawley, D. N. A modified method of estimation in factor analysis and some large sample results. Uppsala Symposium on Psychological Factor Analysis. Nordisk Psykologis' Monograph Series No. 3. Stockholm: Almqvist and Wiksell, 1953. Pp. 35-42.

Lawley, D. N. Some new results in maximum likelihood factor analysis. Proceedings of the Royal Society of Edinburgh, 1967, A67, 256-264.

Lawley, D. N. & Maxwell, A. E. Factor analysis as a statistical method. New York: American Elsevier, 1971.

TABLE 1

Standard Errors Using Lawley's Formulas

	Factor		
Variate	I	II	III
1	.066	.058	.076
2	.064	.061	.068
3	.070	.071	.083
4	.060	.046	.045
5	.065	.057	.072
6	.064	.046	.037
7	.068	.057	.066
8	.073	.093	.142
9	.066	.047	.055

TABLE 2

Standard Errors Using an Augmented Information Matrix

	Factor		
Variate	I	II	III
1	.068	.066	.091
2	.066	.076	.069
3	.072	.081	.084
4	.061	.068	.054
5	.067	.073	.081
6	.069	.049	.036
7	.071	.068	.080
8	.076	.103	.124
9	.070	.058	.078

TABLE 3

Standard Errors Using the Modified Lawley Formulas

	Factor		
Variate	I	II	III
1	.068	.066	.091
2	.066	.076	.069
3	.072	.081	.084
4	.061	.067	.054
5	.067	.073	.081
6	.069	.049	.036
7	.071	.068	.080
8	.076	.103	.124
9	.070	.058	.078

TABLE 4

Standard Errors for Standardized Loadings

Using Lawley and Maxwell's Formulas

	Factor		
Variate	I	II	III
1	.044	.057	.076
2	.041	.060	.068
3	.058	.069	.082
4	.027	.048	.045
5	.040	.057	.073
6	.032	.050	.038
7	.046	.056	.066
8	.062	.090	.140
9	.037	.050	.055

TABLE 5

Standard Errors for Standardized Loadings Using an

Augmented Information Matrix

	Factor		
Variate	I	II	III
1	.047	.065	.091
2	.043	.076	.069
3	.060	.078	.083
4	.030	.069	.054
5	.043	.073	.081
6	.042	.052	.037
7	.050	.067	.080
8	.066	.101	.121
9	.044	.060	.078

TABLE 6

Standard Errors for Standardized Loadings Using the

Modified Lawley and Maxwell Formulas

	Factor		
Variate	I	II	III
1	.047	.065	.091
2	.043	.076	.069
3	.060	.078	.083
4	.030	.069	.054
5	.043	.073	.081
6	.042	.052	.037
7	.050	.067	.080
8	.066	.101	.122
9	.044	.060	.078