*Multivariate Behavioral Research*, 1980, 15, 335-352

# TESTING PATTERN HYPOTHESES ON CORRELATION MATRICES: ALTERNATIVE STATISTICS AND SOME EMPIRICAL RESULTS

JAMES H. STEIGER
University of British Columbia

ABSTRACT

The goodness-of-fit of correlational pattern hypotheses has traditionally been assessed either with a likelihood ratio statistic (in conjunction with maximum likelihood estimation) or with a quadratic form statistic (in conjunction with generalized least squares estimates). In the present paper, several alternative statistics, based on the use of the Fisher r-to-z transform, are proposed, and their performance (as well as that of the traditional statistics) is assessed in a Monte Carlo experiment. The new statistics are shown to have Type I error rate performance at smaller sample sizes which is notably superior to their more traditional counterparts.

A correlational "pattern hypothesis" specifies that certain groups of elements in a correlation matrix are equal to each other, and/or to a specified numerical value. Such hypotheses have wide application in the social sciences. For example, they can be used to test whether correlations among variables measured on the same subjects have remained stable over time, or to test significance in "cross-lagged panel correlation" analysis. For other applications, see such references as McDonald (1975), or Jöreskog (1978).

The statistical testing of pattern hypotheses is not straightforward, because correlation coefficients measured on the same individuals are, in general, dependent random variables. The possibility of obtaining maximum likelihood estimates for the parameters of a general analysis of covariance structures model (Jöreskog, 1970) made a broad range of correlational tests possible and was a considerable breakthrough in this area. Details of such testing procedures are given elsewhere (*e.g.*, Jöreskog, 1978) and need not be repeated here. Essentially, in the analysis of covariance structures approach, the hypothesis is that the population covariance matrix C, of order $m \times m$, is of the form

$$\text{[1]} \qquad \mathbf{C} = \mathbf{DRD}$$

Research support by The National Research Council of Canada (NRC Grant # A4640) and the University of British Columbia (U.B.C. HSSE Grant #26-9914) is gratefully acknowledged.

where **D** is a diagonal matrix of scale factors to be estimated, and **R** is a patterned correlation matrix. If $\mathbf{C}^*$ is the matrix of maximum likelihood estimates of **C** under the null hypothesis, and **A** is the sample covariance matrix, then for

$$[2] \qquad \varphi_1 = \log|\mathbf{C}^*| - \log|\mathbf{A}| + \mathrm{Tr}(\mathbf{A}\mathbf{C}^{*-1}) - m$$

the statistic $U_1 = N\varphi_1$ is distributed asymptotically as a chi-square variate with degrees of freedom given by McDonald (1974, Eq. 5), among others.

A drawback of the likelihood ratio test approach is that it requires computer iteration of maximum likelihood estimates of **D** and **R**. McDonald's (1974) ingenious derivation of the matrix of first and second derivatives of the likelihood function allowed the use of the classical Newton method for iterating these estimates and reduced computation time significantly, without changing the essential nature of the statistical procedure.

An alternative approach to estimation and significance testing for correlational pattern hypotheses uses the generalized least squares principle. Browne (1977) gave details of this method, together with a quadratic form test statistic.

Both the likelihood ratio statistic and the generalized least squares quadratic form are asymptotic chi-square statistics, and their convergence properties and small sample usefulness are largely unknown. Although expansion formulae are available for obtaining close approximations to the distribution of $U_1$ for some hypotheses, such formulas are cumbersome to derive and implement for the general case. Hence, although the user of the traditional statistics is informed that they are "large sample," there is seldom any practical advice about how "large" a "large" $N$ should be.

The original primary purpose of the present research was to shed some light on this question through the use of Monte Carlo simulation methods. However, preliminary results indicated that Type I error rate control for the traditional asymptotic statistics was simply not adquate when $N$ was less than $10p$. This, in turn, led to attempts to develop test statistics with improved small-sample performance qualities.

The approach used in the present paper for accomplishing this objective relies on the normalizing and variance-stabilizing properties of the Fisher $r$-to-$z$ transform. Neill & Dunn (1975) reported that, for testing the extremely simple "pattern hypothesis" that two elements of a correlation matrix are equal, the use of the Fisher transform in what can be seen to be a special case of the quadratic form statistic led to excellent small sample performance. This, in turn, suggested the more general use of the transform in modified quadratic form statistics.

After a review of some required preliminary algebra, the modified quadratic form statistics are derived. The performance of the new and old statistics is then assessed and compared in a series of Monte Carlo experiments. The encouraging results suggest very strongly that, with the use of the newer test statistics, adequate Type I error rate control can be maintained with sample sizes as small as $5p$, thus yielding, effectively, "small sample" inference for correlational pattern hypotheses with either the maximum-likelihood or generalized least squares approaches.

## Some Preliminary Results

Let $\mathbf{p}$ be a vector of the $k = (m^2 - m)/2$ unique off-diagonal elements of $\mathbf{P}$, the $m \times m$ population correlation matrix of a MVN random vector $\mathbf{x}$. Let $\mathbf{r}$ be a corresponding vector of sample correlation coefficients, based on a sample of size $N$ on $\mathbf{x}$.

A *pattern hypothesis* on $\mathbf{p}$ states that some elements of $\mathbf{p}$ are equal to each other, or to specified values. Such hypotheses may be written in the form

$$\mathrm{H}_0: \quad \mathbf{p} = \Delta\gamma + \mathbf{p}^* = \mathbf{p}(\gamma, \mathbf{p}^*) \ ,$$

where $\Delta$ is a $k \times q$ matrix of zeroes and ones with elements given by $\delta_{ij} = \partial \mathrm{p}_i / \partial \gamma_j$, $\gamma$ is a $q \times 1$ vector of common (unspecified) correlations, and $\mathbf{p}^*$ is a $k \times 1$ vector containing specified values for elements of $\mathbf{p}$, and zeroes in all other positions.

For example, let $\mathbf{P}$ be $4 \times 4$. Let $\mathrm{H}_0$ be that $p_{31} = p_{21} = \gamma_1$, that $p_{32} = .6$, and that $p_{41} = p_{42} = p_{43} = \gamma_2$. Then $\mathrm{H}_0$ may be written

$$\mathrm{H}_0: \begin{bmatrix} p_{21} \\ p_{31} \\ p_{41} \\ p_{32} \\ p_{42} \\ p_{43} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ .6 \\ 0 \\ 0 \end{bmatrix}$$

In practice, since $\mathbf{p}^*$ contains specified values, the estimation of $p$ reduces immediately to the problem of estimating $\gamma$.

In this paper, discussion is restricted to three types of estimates of $\gamma$, $\hat{\gamma}_{ML}$ ("maximum likelihood estimates"), which minimize the loss function $\varphi_1$, in [2], $\hat{\gamma}_{OLS}$ ("ordinary least squares" estimates), which minimize $\varphi_2 = (\mathbf{r} - \hat{\mathbf{p}})' (\mathbf{r} - \hat{\mathbf{p}})$, and $\hat{\gamma}_{GLS}$ ("generalized least squares estimates") which minimize the loss function $\varphi_3 = (\mathbf{r} - \hat{\mathbf{p}})' \hat{\Sigma}^{-1} (\mathbf{r} - \hat{\mathbf{p}})$, where $\hat{\Sigma}$ is a consistent estimator of $\Sigma$, the asymptotic variance-covariance matrix of $N^{1/2}\mathbf{r}$. $\Sigma$ has typical element $\sigma_{jk,lm}$ given by (see Pearson & Filon, 1898)

$$[3] \qquad \sigma_{jk,lm} = \tfrac{1}{2}[(p_{jl} - p_{jk}p_{kl})(p_{km} - p_{kl}p_{lm}) + (p_{jm} - p_{jl}p_{lm})(p_{kl} - p_{kj}p_{jl}) + (p_{jl} - p_{jm}p_{ml})(p_{km} - p_{kj}p_{jm}) + (p_{jm} - p_{jk}p_{km})(p_{kl} - p_{km}p_{ml})]$$

It is well known that $\mathbf{r}$ has an asymptotic distribution which is multivariate normal, with mean $\mathbf{p}$ and variance-covariance matrix $\Sigma/N$. $\hat{\gamma}_{ML}$, in general, is not expressible in closed form, and must be obtained by iteration. (See, for example, McDonald, 1975.) $\hat{\gamma}_{LS}$ is given by

$$[4] \qquad \hat{\gamma}_{LS} = (\Delta'\Delta)^{-1}\Delta'\mathbf{r}$$

$\hat{\gamma}_{GLS}$ may be written as

$$[5] \qquad \hat{\gamma}_{GLS} = (\Delta'\hat{\Sigma}_{OLS}{}^{-1}\Delta)^{-1}\Delta'\hat{\Sigma}_{OLS}{}^{-1}(\mathbf{r} - \mathbf{p}^*)$$

where $\hat{\Sigma}_{OLS}$ is the estimate of $\Sigma$ obtained by substituting elements of $\hat{\mathbf{p}}_{OLS}$ for $p_{ij}$ in [3].

Let $\mathbf{z}_r$ be a vector of Fisher transforms of the elements of $\mathbf{r}$, *i.e.*,

$$[6] \qquad z(r_{ij}) = \tfrac{1}{2} \ln [(1 + r_{ij})/(1 - r_{ij})] \; .$$

Straightforward application of the multivariate "delta" theorem (Olkin and Siotani, 1964, Note 1; Szatrowski, 1979) yields the result that $\mathbf{z}_r$ has an asymptotic distribution which is multivariate normal, with mean $\mathbf{z}_p$, and variance-covariance matrix $\Sigma^*/(N - 3)$, where $\Sigma^*$ has typical element given by

$$[7] \qquad \sigma^*_{jk,lm} = \sigma_{jk,lm}/(1 - \rho_{jk}^2)\;(1 - \rho_{lm}^2)$$

(This result is given in several sources. For example, see Dunn and Clark, 1969.)

## Quadratic Form Statistics

Some important equivalencies between maximum likelihood and GLS estimators have already been established. (See Browne, 1974, 1977.) Perhaps the most important result is that $\hat{\gamma}_{ML}$ and $\hat{\gamma}_{GLS}$ are asymptotically equivalent, which may be verified as follows. The asymptotic distribution of $\mathbf{r}$ is $MVN(\mathbf{p},\Sigma/N)$. Hence, the asymptotic log likelihood is maximized under a null pattern hypothesis by minimizing the quadratic form, $N(\mathbf{r} - \Delta\hat{\gamma} - \mathbf{p}^*)'\Sigma^{-1}(\mathbf{r} - \Delta\hat{\gamma} - \mathbf{p}^*)$, under choice of $\hat{\gamma}$. This is accomplished by setting $\hat{\gamma}$ equal to the asymptotic MLE, $\hat{\gamma}^*_{ML} = (\Delta'\Sigma^{-1}\Delta)^{-1}\Delta'\Sigma^{-1}(\mathbf{r} - \mathbf{p}^*)$. $\hat{\gamma}^*_{ML}$ and $\hat{\gamma}_{GLS}$ are asymptotically equivalent, since $\text{plim}_{N\to\infty} N^{1/2} (\hat{\gamma}_{GLS} - \hat{\gamma}^*_{ML}) = 0$. This last fact may be proven quite directly. Specifically,

$$\begin{aligned} N^{1/2}(\hat{\gamma}_{GLS} - \hat{\gamma}^*_{ML}) &= [(\Delta'\hat{\Sigma}_{OLS}{}^{-1}\Delta)^{-1}\Delta'\hat{\Sigma}_{OLS}{}^{-1} - \\ &\qquad (\Delta'\Sigma^{-1}\Delta)^{-1}\Delta'\Sigma^{-1}]N^{1/2}(\mathbf{r} - \mathbf{p}^* - \Delta\gamma) \\ &= BN^{1/2}(\mathbf{r} - \mathbf{p}^* - \Delta\gamma) \end{aligned}$$

However, $\text{plim}_{N\to\infty} B = 0$, and, on the other hand, $N^{1/2}(\mathbf{r} - \mathbf{p}^* - \Delta\gamma)$ has a limiting distribution, thus establishing the result.

Next, we give an abbreviated proof that the quadratic form

$$[8] \qquad X = N(\mathbf{r} - \hat{\mathbf{p}}_{GLS})\hat{\Sigma}_{OLS}{}^{-1}(\mathbf{r} - \hat{\mathbf{p}}_{GLS})$$

is asymptotically $\chi^2_{k-q}$. First, we recall the following (Timm, 1975, p. 132)

> Lemma. Let **y** be MVN($\mu$,**C**). Then the quadratic form $X$ = **y'Ay** has a *chi square* distribution with $v$ degrees of freedom and non-centrality parameter $\lambda = \mu'\mathbf{A}\mu$ if and only if **AC** is idempotent and of rank $v$.

Asymptotically, the result will be applicable if $\operatorname{plim}_{N\to\infty}\mathbf{AC}$ satisfies the Lemma. In the present case, $X$ may be written

[9] $$X = N(\mathbf{r} - \Delta\hat{\gamma}_{GLS} - \mathbf{p}^*)'\hat{\Sigma}_{OLS}^{-1}(\mathbf{r} - \Delta\hat{\gamma}_{GLS} - \mathbf{p}^*)$$

$$= N(\mathbf{r}^{*\prime}\hat{\Sigma}_{OLS}^{-1}\mathbf{r}^*)$$

with

[10] $$\mathbf{r}^* = (\mathbf{I} - \Delta\mathbf{Q})\ (\mathbf{r} - \mathbf{p}^*), \text{ and } \mathbf{Q} = (\Delta'\hat{\Sigma}_{OLS}^{-1}\Delta)^{-1}\Delta'\hat{\Sigma}_{OLS}^{-1}$$

Since, under a true null hypothesis, $\operatorname{plim}_{N\to\infty}\hat{\Sigma}_{OLS}^{-1} = \Sigma^{-1}$, it follows that $\mathbf{Q}^* = \operatorname{plim}_{N\to\infty}\mathbf{Q} = (\Delta'\Sigma^{-1}\Delta)^{-1}\Delta'\Sigma^{-1}$, and that, under true $H_0$, $\mathbf{r}^*$ is asymptotically MVN($\emptyset$, $(\mathbf{I} - \Delta\mathbf{Q}^*)\ (\Sigma/N)\ (\mathbf{I} - \Delta\mathbf{Q}^*)'$). To prove that $X$ is asymptotically $\chi^2_{k-q}$, we must show that $\Sigma^{-1}(\mathbf{I} - \Delta\mathbf{Q}^*)\ \Sigma\ (\mathbf{I} - \Delta\mathbf{Q}^*)'$ is idempotent and of rank $k-q$. Idempotency is easily established by substitution and recombination. Specifically, we find that $\Sigma^{-1}(\mathbf{I} - \Delta\mathbf{Q}^*)\ \Sigma\ (\mathbf{I} - \Delta\mathbf{Q}^*)' = \mathbf{I} - \Sigma^{-1}\Delta(\Delta'\Sigma^{-1}\Delta)^{-1}\Delta'$. In this latter form, we clearly recognize that the rank of the expression depends on the rank of $\Delta$, which is $k-q$, and idempotency is also immediately evident.

Since GLS and ML estimators are asymptotically equivalent, it follows directly that the quadratic form

[11] $$U_2 = N(\mathbf{r} - \hat{\mathbf{p}}_{ML})'\ \hat{\Sigma}_{ML}^{-1}\ (\mathbf{r} - \hat{\mathbf{p}}_{ML})$$

is also asymptotically $\chi^2_{k-q}$. (It perhaps should be noted that any consistent estimator of $\Sigma$ can be substituted in [8] and [11] without affecting the asymptotic result.) The preceding theory generalizes readily to monotonic, differentiable functions $\mathbf{f}(\mathbf{p})$ of $\mathbf{p}$. Specifically, any pattern hypothesis on $\mathbf{p}$ can be reexpressed as an equivalent pattern hypothesis on $\mathbf{f}(\mathbf{p})$, *i.e.*, $\mathbf{f}(\mathbf{p}) = \Delta\mathbf{f}(\gamma) + \mathbf{f}(\mathbf{p}^*)$. Via the multivariate delta theorem, we find that $\mathbf{f}(\mathbf{r})$ is asymptotically $\text{MVN}(\mathbf{f}(\mathbf{p}), \Sigma^+/N)$. $\Sigma^+$ is calculated from $\Sigma$ as $\Sigma^+ = \mathbf{D}'\Sigma\mathbf{D}$, with $\mathbf{D}$ having elements $\delta_{ij} = \partial f(r_i)/\partial r_j|\mathbf{r}=\mathbf{p}$. With arguments analogous to those preceding, it may be verified that the maximum likelihood and generalized least squares estimators of $\mathbf{f}(\gamma)$ are equivalent, and that chi-square quadratic forms analogous to [8] and [11], but with $\hat{\mathbf{f}}(\mathbf{p})$, $\mathbf{f}(\mathbf{r})$, and $\hat{\Sigma}^+$ substituted for $\mathbf{p}$, $\mathbf{r}$, and $\hat{\Sigma}$, respectively, may be constructed. The GLS estimator $\hat{\mathbf{f}}(\mathbf{p})$ is given by $\hat{\mathbf{f}}(\mathbf{p})_{\text{GLS}} = \hat{\mathbf{f}}(\gamma)_{\text{GLS}} + \hat{\mathbf{f}}(\mathbf{p}^*)$, with $\hat{\mathbf{f}}(\gamma)_{\text{GLS}} = (\Delta'\hat{\Sigma}^+{}_{\text{OLS}}{}^{-1}\Delta)^{-1}$ $\Delta'\hat{\Sigma}^+{}_{\text{OLS}}{}^{-1}(\mathbf{f}(\mathbf{r}) - \mathbf{f}(\mathbf{p}^*))$.

Alternatively, since $\hat{\mathbf{f}}(\mathbf{p})_{\text{ML}} = \mathbf{f}(\hat{\mathbf{p}}_{\text{ML}})$, and since, asymptotically, because of the equivalence of GLS and ML estimators, $\text{plim}_{N\to\infty}$ $N^{1/2}(\mathbf{f}(\hat{\mathbf{p}}_{\text{ML}}) - \mathbf{f}(\hat{\mathbf{p}}_{\text{GLS}})) = 0$, we have $\chi^2$ forms with $\mathbf{f}(\hat{\mathbf{p}})$, $\mathbf{f}(\mathbf{r})$, and $\hat{\Sigma}^{+-1}$ substituted for $\hat{\mathbf{p}}$, $\mathbf{r}$, and $\hat{\Sigma}^{-1}$, respectively. Our present interest is in normalizing and variance stabilizing transforms, such as Fisher's *r*-to-*z* [6] or Hotelling's (1953) extensions, which have the potential for improving the small-sample performance of the chi-square statistic. In the case of the Fisher transform, we have the quadratic forms

[12] $$U_2^* = (N - 3)\ (\mathbf{z}_r - \hat{\mathbf{z}}_{p_{\text{ML}}})'\ \hat{\Sigma}^*{}_{\text{ML}}{}^{-1}\ (\mathbf{z}_r - \hat{\mathbf{z}}_{p_{\text{ML}}}), \text{ and}$$

[13] $$X^* = (N - 3)\ (\mathbf{z}_r - \hat{\mathbf{z}}_{p_{\text{GLS}}})'\ \hat{\Sigma}^*{}_{\text{OLS}}{}^{-1}\ (\mathbf{z}_r - \hat{\mathbf{z}}_{p_{\text{GLS}}})\ .$$

The multiplier $N-3$, used in place of $N$, is a practical concession to the fact that the variances of $\mathbf{z}_r$ are known to be, for small samples, approximately $1/(N-3)$. The multiplier has no effect on the asymptotic result. Similar quadratic forms may be constructed for the Hotelling transforms.

$X$ and $X^*$ involve less computational effort than $U_1$, $U_2$, or $U_2^*$, because the latter statistics require (sometimes extensive) iteration to obtain maximum likelihood estimates. It should also be noted that the $k \times k$ matrix $\hat{\Sigma}^*$ need not be inverted, as the elements of $\hat{\Sigma}^{*-1}$ can be obtained more directly with adaptations of a formula given in Jennrich (1970). Jennrich's method requires only $p \times p$ and $q \times q$ matrix inversions and can yield considerable savings in computation time when $k > 10$.

As a "quick approximation" to $X_1$ and $X_2$, one might substitute OLS estimators and obtain

$$[14] \qquad X_0 = (N - 3)\ (\mathbf{z}_r - \hat{\mathbf{z}}_{p_{\text{OLS}}})'\ \hat{\Sigma}^*{}_{\text{OLS}}{}^{-1}\ (\mathbf{z}_r - \hat{\mathbf{z}}_{p_{\text{OLS}}})\ .$$

$X_0$ is somewhat easier to compute than $X_2$, is usually very close to it in numerical value, and is, for many pattern hypotheses, formally equivalent. (Browne, 1977, discusses a number of situations where OLS and GLS estimators are formally identical.) $X_0$ is not always a $\chi^2_{k-q}$ statistic although (as shown in the Monte Carlo experiments to follow) it appears to be a rather good approximation.

## Empirical Performance of the Statistics—Some Monte Carlo Results

All of the test statistics discussed are based on asymptotic distribution theory. However, there is virtually no information available about how such statistics (even the more traditional $U_1$ and $X$) perform at small to medium sample sizes. Neill and Dunn (1975) showed that a statistic similar to $X_0$ decisively outperformed $U$ at small sample sizes in Type I error control. One might conjecture that this small sample superiority is due to the distributional stability of $\mathbf{z}_r$, and would generalize to more complicated hypotheses, but further empirical evidence is necessary to establish this.

In the present paper we investigate, through Monte Carlo simulation experiments, the performance of various statistics in testing several pattern hypotheses. The investigation is limited by a number of practical considerations. First, any of the statistics may be applied to a virtually unlimited variety of hypotheses and the present study investigated only a small cross-section of these. Second, the computation time required by the maximum likelihood approach makes Monte Carlo investigation, even for small and moderate size matrices, extremely expensive, and so very large correlation matrices could not be investigated. Despite its limitations, the study provides valuable information on the relative performance of the various methods. (Indeed, there seems to be little information available on the performance of the likelihood ratio test, despite its wide popularization.)

The results reported here concentrate on Type I error rate control, rather than power, for several reasons. First, the author feels that, in comparing statistics, Type I error rate control is of primary importance. If statistics control Type I error equally well, then relative power becomes important, but, in our opinion, power is not of great importance if a statistic does not control Type I error. Second, for situations where violation of the null hypothesis is moderate, the asymptotic chi-square statistics are known to have a non-central chi-square distribution with non-centrality parameter which is very close to the value which would be obtained if the statistic were computed on the population correlation matrix. (See, for example, Kendall and Stuart, 1961, p. 231, Szatrowski, 1979.) The non-central chi-square can be closely approximated by an adjusted central chi-square, and so power estimates for the various statistics may be obtained rather easily. In situations where the statistics have converged to their asymptotic distributions, we would expect them to have approximately the same power. Third, empirical assessment of power is extremely expensive, because for any given type of pattern, a whole range of violations of the null hypothesis must be examined.[1]

The general method used for simulating the sampling of a correlation matrix from a multivariate normal population was a standard approach for studies of this type. Vectors of independent random normal deviates were constructed using a standard (Marsaglia rectangle wedge-tail method, given in Knuth (1968)) random number generator on an Amdahl V/6-II computer. These vectors were then linearly recombined using a Gram-factor of the desired correlation matrix to yield the simulated sample data matrix. Sample correlation matrices were then computed and analyzed by the various computing methods. (Unless otherwise noted, 2500 replications were obtained in each condition.) Maximum likelihood

1. The author has examined empirical power performance for the statistics for a limited number of situations where $N$ was sufficiently large to minimize differences in Type I error rates. In these cases, power for the various statistics was virtually identical, and empirical power was very close to that predicted by the non-central chi square approximation.

estimates were obtained by TESPAR (McDonald, 1975) which is by all accounts the most efficient program of its type available. Convergence criterion was set, as in McDonald (1975), at .0005. (Preliminary investigation seemed to indicate that lowering the convergence criterion to .0001 had virtually no effect on Type I error rates.)

Type I error rate control was assessed empirically for six different types of pattern hypotheses, which, for convenience, we label (a) Identity, (b) Equicorrelation, (c) Partial Equicorrelation, (d) Matrix Equality, (e) Circumplex, and (f) Toeplitz.

The Identity hypothesis is simply that all off-diagonal elements of the correlation matrix are zero. Two levels of matrix size ($p$ = 4, 8) were crossed factorially with six levels of sample size ($N$ = 25, 50, 75, 100, 150, 200) in this investigation. Results are presented in Tables 1 and 2, which contain empirical Type I error rates when nominal $\alpha$ was .05. (Results at other levels of nominal

**Table 1**
**Empirical Type I Error Rates—Identity Hypothesis**

| | | | m = 4 | | | |
|---|---|---|---|---|---|---|
| | $X_0$ | $U_1$ | $U_c$ | $U_2^*$ | $X$ | $X^*$ |
| N = 25 | .0508 | .0888 | .0500 | .0508 | .0476 | .0508 |
| 50 | .0408 | .0584 | .0416 | .0408 | .0400 | .0408 |
| 75 | .0532 | .0604 | .0496 | .0532 | .0524 | .0532 |
| 100 | .0428 | .0552 | .0452 | .0428 | .0424 | .0428 |
| 150 | .0468 | .0492 | .0452 | .0468 | .0424 | .0468 |
| 200 | .0536 | .0532 | .0520 | .0536 | .0532 | .0536 |
| | | | m = 8 | | | |
| | $X_0$ | $U_1$ | $U_c$ | $U_2^*$ | $X$ | $X^*$ |
| N = 25 | .0636 | .2144 | .0596 | .0636 | .0676 | .0636 |
| 50 | .0468 | .1008 | .0500 | .0468 | .0496 | .0468 |
| 75 | .0568 | .0908 | .0568 | .0568 | .0568 | .0568 |
| 100 | .0480 | .0792 | .0524 | .0480 | .0516 | .0480 |
| 150 | .0540 | .0692 | .0552 | .0540 | .0548 | .0540 |
| 200 | .0504 | .0572 | .0492 | .0504 | .0508 | .0504 |

$\alpha$ (.02 and .01) essentially parallel those for the .05 significance level. The .05 level results have the lowest inherent variability and so we present them alone in the interest of brevity.) The left hand column gives results for $X_0$, the quadratic form "approximate" chi-square statistic using OLS estimates. The next three columns give statistics based on maximum likelihood estimates. $U_1$ is the ("uncorrected") likelihood ratio statistic. $U_c$ is a "corrected" likelihood ratio statistic, obtained by substituting the constant

Table 2
Empirical Type I Error Rates—Equicorrelation Hypothesis

| | $X_0$ | $U_1$ | $U_c$ | $U_2$* | $X$ | $X$* |
|---|---|---|---|---|---|---|
| m = 4  γ = .90 | | | | | | |
| N = 25 | .0644 | .0940 | .0572 | .0644 | .0620 | .0644 |
| 50 | .0500 | .0600 | .0460 | .0500 | .0500 | .0500 |
| 75 | .0564 | .0632 | .0536 | .0564 | .0552 | .0564 |
| 100 | .0512 | .0572 | .0532 | .0512 | .0520 | .0512 |
| 150 | .0516 | .0536 | .0488 | .0516 | .0492 | .0516 |
| 200 | .0608 | .0652 | .0624 | .0608 | .0604 | .0608 |
| m = 4  γ = .50 | | | | | | |
| | $X_0$ | $U_1$ | $U_c$ | $U_2$* | $X$ | $X$* |
| N = 25 | .0452 | .0884 | .0528 | .0452 | .0520 | .0452 |
| 50 | .0468 | .0616 | .0432 | .0468 | .0464 | .0468 |
| 75 | .0492 | .0624 | .0536 | .0496 | .0512 | .0492 |
| 100 | .0512 | .0592 | .0504 | .0516 | .0512 | .0512 |
| 150 | .0564 | .0624 | .0548 | .0564 | .0556 | .0564 |
| 200 | .0548 | .0616 | .0576 | .0548 | .0564 | .0548 |
| m = 4  γ = .10 | | | | | | |
| | $X_0$ | $U_1$ | $U_c$ | $U_2$* | $X$ | $X$* |
| N = 25 | .0452 | .0836 | .0484 | .0448 | .0492 | .0452 |
| 50 | .0464 | .0640 | .0492 | .0468 | .0464 | .0464 |
| 75 | .0468 | .0572 | .0484 | .0468 | .0492 | .0468 |
| 100 | .0584 | .0692 | .0592 | .0584 | .0612 | .0584 |
| 150 | .0420 | .0496 | .0432 | .0420 | .0436 | .0420 |
| 200 | .0512 | .0532 | .0500 | .0512 | .0516 | .0512 |

$N - [(2m + 11)/6]$, as suggested by Box (1949), for $N$ in computing $U_1$. ($U_c$ was derived specifically for testing the Identity hypothesis at small sample sizes. However, preliminary indications from an earlier Monte Carlo investigation were that the use of the correction constant might also produce considerable improvement in the performance of the likelihood ratio statistic in testing other hypotheses. Consequently, we present results for $U_c$ in testing these other hypotheses.) $U_2^*$ is the quadratic form statistic given in [9], which uses Fisher-transformed MLE's. The rightmost columns give results for statistics based on GLS estimators. $X$ is the original "raw correlation" quadratic form statistic given by Browne (1977). $X^*$ [13] is a revised quadratic form, using Fisher-transformed correlations. It should be noted that, for the Identity hypothesis, $X$, $X^*$, and $U_2^*$ are all identical, and reduce to the simplified formula

$$[15] \qquad X_0 = X^* = U_2^* = (N - 3) \sum_{j<i} z_{ij}^2 \; .$$

$X$ has the very simple reduced formula

[16] $$X = N \sum_{j<i} r_{ij}^2 \ .$$

The equicorrelation hypothesis states that all off-diagonal elements of the correlation matrix are equal to a common, but unspecified value. Three different population matrices of order 4x4 were tested, with $\gamma$ values of .90, .50, and .10. The partial equicorrelation hypothesis (results are in Table 3) specifies that two

Table 3
Empirical Type I Error Rates—Partial Equicorrelation Hypothesis

| $\gamma = .50$, m = 4 | | | | | | |
|---|---|---|---|---|---|---|
| | $X_0$ | $U_1$ | $U_c$ | $U_2{}^*$ | $X$ | $X^*$ |
| N = 25 | .0516 | .0554 | .0424 | .0445 | .0580 | .0516 |
| 50 | .0456 | .0520 | .0444 | .0460 | .0508 | .0456 |
| 75 | .0476 | .0512 | .0468 | .0480 | .0500 | .0476 |
| 100 | .0460 | .0484 | .0452 | .0460 | .0480 | .0460 |
| 150 | .0468 | .0492 | .0468 | .0468 | .0476 | .0468 |
| 200 | .0472 | .0492 | .0472 | .0476 | .0488 | .0472 |

elements of the correlation matrix are equal to each other and to a common unspecified value. In this investigation the population

Table 4
Empirical Type I Error Rates—Matrix Equality

| | $X_0$ | $X$ | $X^*$ |
|---|---|---|---|
| m = 4 $\gamma$ = .90 | | | |
| N = 25 | .0556 | .1472 | .0556 |
| 50 | .0524 | .1100 | .0524 |
| 75 | .0460 | .0852 | .0460 |
| 100 | .0488 | .0832 | .0488 |
| 150 | .0524 | .0624 | .0524 |
| 200 | .0488 | .0676 | .0488 |
| m = 4 $\gamma$ = .50 | $X_0$ | $X$ | $X^*$ |
| N = 25 | .0536 | .0920 | .0536 |
| 50 | .0548 | .0724 | .0548 |
| 75 | .0580 | .0692 | .0580 |
| 100 | .0488 | .0596 | .0488 |
| 150 | .0536 | .0620 | .0536 |
| 200 | .0476 | .0600 | .0476 |
| m = 4 $\gamma$ = .10 | $X_0$ | $X$ | $X^*$ |
| N = 25 | .0496 | .0528 | .0496 |
| 50 | .0460 | .0464 | .0460 |
| 75 | .0456 | .0456 | .0456 |
| 100 | .0492 | .0500 | .0492 |
| 150 | .0508 | .0524 | .0508 |
| 200 | .0524 | .0508 | .0524 |

matrices were identical to those used in testing the equicorrelation condition. The hypothesis tested was that $p_{12} = p_{13}$.

The matrix equality hypothesis states that $\mathbf{p} = \mathbf{p}^*$, *i.e.* that the population correlations are exactly equal to specified values. In this case, two sizes of population matrix (4x4, 8x8) were investigated. The 4x4 population matrices were the same as in the equicorrelation and partial equicorrelation conditions. The 8x8 matrix had all off-diagonal elements equal to .50. Results for this hypothesis are presented in Tables 4 and 5. (Since McDonald's (1975) program cannot test this hypothesis except for $\mathbf{P} = \mathbf{I}$, statistics based on ML estimation were not calculated in this condition.)

**Table 5**
**Empirical Type I Error Rates—Matrix Equality 8 x 8**
**(1000 replications)**

| | $\gamma = .50$ | | |
|---|---|---|---|
| | $X_0$ | $X$ | $X^*$ |
| N = 25 | .0580 | .1760 | .0580 |
| 50 | .0650 | .1200 | .0650 |
| 75 | .0610 | .0990 | .0610 |
| 100 | .0510 | .0950 | .0510 |
| 150 | .0480 | .0740 | .0480 |
| 200 | .0370 | .0600 | .0370 |

Results for the circumplex and Toeplitz patterns are presented in Table 6. The circumplex hypothesis, for a 6x6 correlation

**Table 6**
**Empirical Type I Error Rates—Circumplex Hypothesis**

| | | | m = 6 | | | |
|---|---|---|---|---|---|---|
| | $X_0$ | $U_1$ | $U_c$ | $U_2^*$ | $X$ | $X^*$ |
| N = 25 | .0644 | .1468 | .0724 | .0672 | .0760 | .0644 |
| 50 | .0460 | .0780 | .0516 | .0480 | .0564 | .0460 |
| 75 | .0492 | .0672 | .0536 | .0500 | .0528 | .0492 |
| 100 | .0460 | .0588 | .0472 | .0456 | .0480 | .0460 |
| 150 | .0460 | .0608 | .0492 | .0464 | .0488 | .0460 |
| 200 | .0448 | .0504 | .0440 | .0456 | .0456 | .0448 |

**Empirical Type I Error Rates—Toeplitz pattern**

| | | | m = 6 | | | |
|---|---|---|---|---|---|---|
| | $X_0$ | $U_1$ | $U_c$ | $U_2^*$ | $X$ | $X^*$ |
| N = 25 | .0648 | .1318 | .0510 | .0490 | .0576 | .0484 |
| 50 | .0608 | .0848 | .0528 | .0500 | .0536 | .0496 |
| 75 | .0552 | .0640 | .0492 | .0480 | .0544 | .0484 |
| 100 | .0520 | .0604 | .0480 | .0468 | .0500 | .0444 |

matrix, specifies **P** to be of the form presented in Table 7. In this

Table 7
The 6 x 6 Circumplex

| | | | | | |
|---|---|---|---|---|---|
| 1 | | | | | |
| $\gamma_1$ | 1 | | | | |
| $\gamma_2$ | $\gamma_1$ | 1 | | | |
| $\gamma_3$ | $\gamma_2$ | $\gamma_1$ | 1 | | |
| $\gamma_2$ | $\gamma_3$ | $\gamma_2$ | $\gamma_1$ | 1 | |
| $\gamma_1$ | $\gamma_2$ | $\gamma_3$ | $\gamma_2$ | $\gamma_1$ | 1 |

case, we used as population values, $\gamma_1 = .3952$ $\gamma_2 = .2950$ $\gamma_3 = .2497$. These are the OLS estimates obtained from analysis of Guttman's (1954) circumplex data. (See, for example, Browne, 1977; Jöreskog, 1978.) The Toeplitz pattern hypothesis specifies **R** is of the form given in Table 8. The same population matrix was

Table 8
Toeplitz Pattern

| | | | | | |
|---|---|---|---|---|---|
| 1 | | | | | |
| $\gamma_1$ | 1 | | | | |
| $\gamma_2$ | $\gamma_1$ | 1 | | | |
| $\gamma_3$ | $\gamma_2$ | $\gamma_1$ | 1 | | |
| $\gamma_4$ | $\gamma_3$ | $\gamma_2$ | $\gamma_1$ | 1 | |
| $\gamma_5$ | $\gamma_4$ | $\gamma_3$ | $\gamma_2$ | $\gamma_1$ | 1 |

used for this hypothesis as for testing the circumplex hypothesis, *i.e.*, $\gamma_1 = .3952$, $\gamma_2 = .2950$, $\gamma_3 = .2497$, $\gamma_4 = .2950$, $\gamma_5 = .3952$.

Some major trends are evident in Tables 1 through 6. First, $U_1$, the uncorrected likelihood ratio statistic generally advocated in treatments of ML estimation, performs rather poorly in all conditions. $X$, the "raw correlation" quadratic form, performs reasonably well in most conditions, but very poorly in the "matrix equality" condition when correlations are high. This problem is in a sense, to be expected, because $X$ assumes raw correlations are normally distributed, and departures from normality are known to be quite severe when $\rho_{ij}$ are high in absolute value and $N$ is not large. $X_0$, the "approximate statistic," performs well in a number of conditions. Indeed, for many of the hypotheses, it appears to be identical to $X^*$. However, for the Toeplitz pattern $X_0$ is notably inferior to $X^*$ though superior to $U_1$.

Clearly $U_c$, $U_2^*$, and $X^*$ appear to be the best performing statistics. They all perform quite well, even at small sample sizes, and they perform consistently for all hypotheses tested.

To help summarize the performance of the test statistics, two chi-square goodness of fit values, based on a normal approximation to the binomial, were computed for each statistic in each condition. The statistics were

$$\chi^2 = R\Sigma(\hat{\alpha} - .05)^2/[\hat{\alpha}(1 - \hat{\alpha})]$$

where $R$ is the number of Monte Carlo repetitions, $\hat{\alpha}$ the empirical estimate of Type I error. The first chi-square was computed for the three values of $N$ from 25 to 75, the second (except for the Toeplitz pattern) for $N = 100\text{-}200$. Small-sample statistics are presented in Table 9 while large-sample statistics are given in Table 10.

Table 9
Summary Chi-Square Statistics
Small Sample Values (N = 25, 50, 75)

| | Statistic | | | | | |
|---|---|---|---|---|---|---|
| *Hypothesis* | $X_0$ | $U_1$ | $U_c$ | $U_2^*$ | $X$ | $X^*$ |
| Identity (m = 4) | 5.95 | 54.49 | 4.43 | 5.95 | 7.12 | 5.95 |
| Identity (m = 8) | 10.50 | 522.75 | 6.27 | 10.50 | 14.45 | 10.50 |
| Equicorrelation ($\gamma$ = .9) | 10.53 | 68.62 | 3.95 | 10.53 | 7.49 | 10.53 |
| Equicorrelation ($\gamma$ = .1) | 2.64 | 47.42 | 0.31 | 2.73 | 0.80 | 2.64 |
| Equicorrelation ($\gamma$ = .5) | 1.94 | 58.14 | 3.83 | 1.92 | 1.01 | 1.94 |
| Circumplex | 9.55 | 226.08 | 19.45 | 12.02 | 26.38 | 9.55 |
| Toeplitz | 15.44 | 193.08 | 0.48 | 0.27 | 4.24 | 0.29 |
| Partial Equicorrelation | 1.56 | 1.67 | 5.98 | 2.91 | 2.96 | 1.56 |
| Overall $\chi^2_{24}$ | 58.11 | 1172.55 | 44.70 | 46.83 | 64.45 | 42.96 |
| Matrix Equality (m = 4, $\gamma$ = .9) | 2.69 | | | | 319.83 | 2.69 |
| Matrix Equality (m = 4, $\gamma$ = .5) | 4.68 | | | | 85.78 | 4.68 |
| Matrix Equality (m = 4, $\gamma$ = .1) | 2.03 | | | | 2.24 | 2.03 |
| Matrix Equality (m = 8, $\gamma$ = .5) | 6.99 | | | | 182.79 | 6.99 |

Table 10
Summary Chi-Square Statistics
Large Sample Values (N = 100, 150, 200)

| | Statistic | | | | | |
|---|---|---|---|---|---|---|
| *Hypothesis* | $X_0$ | $U_1$ | $U_c$ | $U_2$* | $X$ | $X$* |
| Identity (m = 4) | 4.38 | 1.84 | 2.87 | 4.38 | 7.62 | 4.38 |
| Identity (m = 8) | 1.01 | 45.94 | 1.62 | 1.01 | 1.28 | 1.01 |
| Equicorrelation (γ = .9) | 5.31 | 12.52 | 7.16 | 5.31 | 5.00 | 5.31 |
| Equicorrelation (γ = .1) | 7.26 | 14.82 | 6.60 | 7.26 | 8.04 | 7.26 |
| Equicorrelation (γ = .5) | 3.11 | 16.19 | 3.78 | 3.17 | 3.49 | 3.17 |
| Circumplex | 3.40 | 8.61 | 2.61 | 2.96 | 1.41 | 3.40 |
| Toeplitz (N = 100 only) | 0.20 | 4.76 | 0.22 | 0.57 | 0.00 | 1.85 |
| Partial Equicorrelation | 1.92 | 0.21 | 2.34 | 1.80 | 0.61 | 1.80 |
| Overall $\chi^2_{22}$ | 26.59 | 104.89 | 27.20 | 26.46 | 27.45 | 28.18 |
| Matrix Equality (m = 4, γ = .9) | 0.45 | | | | 54.98 | 0.45 |
| Matrix Equality (m = 4, γ = .5) | 1.03 | | | | 14.73 | 1.03 |
| Matrix Equality (m = 4, γ = .1) | 0.36 | | | | 0.32 | 0.36 |
| Matrix Equality (m = 8, γ = .5) | 4.85 | | | | 33.73 | 4.85 |

As a summary measure of performance, an overall $\chi^2$ statistic was obtained by summing values of $\chi^2$ over all conditions. These values confirm that $X^*$, $U_c$, and $U_2^*$ perform very well, $X$ and $X_0$ are adequate (except for $X$ in the matrix equality condition) but that $U_1$ is notably inferior to the other statistics. The chi-square summary statistics for the smaller sample sizes fall between the 60th and 70th percentiles for a $\chi^2_{24}$ variate, and those for large samples are around the mean for $\chi^2_{22}$, indicating that, overall, Type I error rate performance of $U_c$, $X_1^*$ and $U_2^*$ is very close to the nominal value of .05.

## Summary

The preceding results are rather limited. They investigated only a few hypotheses, and for smaller size matrices. To a considerable extent, this limitation was unavoidable, due to the extreme expense of simulating many replications of an iterative procedure. However, the major trends manifested in this data were very consistent, and it seems fair to say that they probably generalize to most if not all pattern hypotheses tests.

One important finding is that $U_1$ is, indeed, a "large sample" statistic. For example, with $N = 75$, $m = 8$, $U_1$ has an empirical

$\alpha$ of .09 at the nominal .05 level. Indications are that this trend toward excessiveness would be much more pronounced when larger matrices are tested. $U_2{}^*$ performs very well in all conditions. Since the time required to compute $U_2{}^*$ is a very small fraction of the time required to obtain MLE's, the computation of this statistic, especially when $m$ is large and $N$ small, will generally yield much more accurate inferences than the use of $U_1$, and with very little increase in cost.

$U_c$ is easily obtained from $U_1$, is easier to compute than $U_2{}^*$, and performed just as well as $U_2{}^*$ in the conditions we tested. However, it should be used with some caution as a general statistic, because the correction factor was derived for a situation where all correlations are specified to be zero. In conditions (such as the partial equicorrelation hypothesis) where degrees of freedom are small relative to $\binom{m}{2}$, the correction factor may be excessively small, and the test statistic too conservative. Nevertheless, $U_c$ seems likely to be a major improvement over $U_1$ in most situations.

Statistic $X$ (Browne, 1977) tends to be highly excessive in testing matrix equality, and this drawback would seem to outweigh any advantages in its use. Fortunately, this does not prove to be a serious limitation, because the modified quadratic form $X$ statistic, presented here appears to allow accurate small sample inference in all situations. The excellent small sample performance of $X^*$ suggests that, indeed, there is little if any necessary loss in accuracy when GLS estimates are used (with the proper test statistic) in place of MLE's. This, in turn, raises the prospect of considerable saving in computational effort in testing pattern hypotheses, especially for large correlation matrices.

## Reference Note

1. Olkin, I., & Siotani, M. Asymptotic distribution functions of a correlation matrix. Technical Report No. 6. Laboratory for quantitative research in education, Stanford University, 1964.

## References

Box, G. E. P. A general distribution theory for a class of likelihood criteria. *Biometrika*, 1949, *36*, 317-346.

Browne, M. W. Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal*, 1974, *8*, 1-24.

Browne, M. W. The analysis of patterned correlation matrices by generalized least squares. *British Journal of Mathematical and Statistical Psychology*, 1977, *30*, 113-124.

Dunn, O. J., & Clark, V. A. Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, 1969, *64*, 366-377.

Guttman, L. A new approach to factor analysis: The radex. In P. F. Lazarsfeld (ed.), *Mathematical thinking in the social sciences*. New York: Columbia University Press, 1954.

Hotelling, H. New light on the correlation coefficient and its transforms. *Journal of the Royal Statistical Society, Series B*, 1953, *15*, 193.

Jennrich, R. I. An asymptotic $\chi^2$ test for the equality of two correlation matrices. *Journal of the American Statistical Association*, 1970, *65*, 904-912.

Jöreskog, K. G. A general method for the analysis of covariance structures. *Biometrika*, 1970, *57*, 239-251.

Jöreskog, K. G. Structural analysis of covariance and correlation matrices. *Psychometrika*, 1978, *43*, 443-477.

Kendall, M. G., & Stuart, A. *The advanced theory of statistics*. Vol. 2. (3rd Ed.) London: Griffin, 1961.

Knuth, D. *The art of computer programming*. Volume 2. Semi-Numerical Algorithms. Reading, Mass. Addison-Wesley, 1968.

McDonald, R. P. Testing pattern hypotheses for covariance matrices. *Psychometrika*, 1974, *39*, 189-203.

McDonald, R. P. Testing pattern hypotheses for correlation matrices. *Psychometrika*, 1975, *40*, 253-255.

Neill, J. J., & Dunn, O. J. Equality of dependent correlation coefficients. *Biometrics*, 1975, *31*, 531-543.

Pearson, K. & Filon, L. N. G. Mathematical contributions to the theory of evolution. IV. On the probable errors of frequency constants and on the influence of random selection on variation and correlation. *Philosophical Transactions of the Royal Society of London, Series A*, 1898, *191*, 229-311.

Szatrowski, T. H. Asymptotic nonnull distributions for likelihood ratio statistics in the multivariate normal patterned mean and covariance matrix testing problem. *The Annals of Statistics*, 1979, *7*, 823-837.

Timm, N. H. *Multivariate analysis with applications in education and psychology*. Monterey, California: Brooks/Cole, 1975.