

## Research Article

# Exact Inference for the Dispersion Matrix

Alan D. Hutson, Gregory E. Wilding, Jihnhee Yu, and Albert Vexler

Department of Biostatistics, University at Buffalo, 706 Kimball Tower, 3435 Main Street, Buffalo, NY 14214-3000, USA

Correspondence should be addressed to Alan D. Hutson; [ahutson@buffalo.edu](mailto:ahutson@buffalo.edu)

Received 10 April 2014; Accepted 4 September 2014; Published 14 September 2014

Academic Editor: Lynn Kuo

Copyright © 2014 Alan D. Hutson et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We develop a new and novel exact permutation test for prespecified correlation structures such as compound symmetry or spherical structures under standard assumptions. The key feature of the work contained in this note is the distribution free aspect of our procedures that frees us from the standard and sometimes unrealistic multivariate normality constraint commonly needed for other methods.

## 1. Introduction

Let  $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$  be an iid  $p$ -dimensional multivariate sample from an absolutely continuous distribution  $F$  with  $p \times p$  dispersion matrix of  $\mathbf{X}$  as

$$\Sigma = \begin{pmatrix} \sigma_{x_1}^2 & \rho_{12}\sigma_{x_1}\sigma_{x_2} & \cdots & \rho_{1p}\sigma_{x_1}\sigma_{x_p} \\ \rho_{21}\sigma_{x_2}\sigma_{x_1} & \sigma_{x_2}^2 & \cdots & \rho_{2p}\sigma_{x_2}\sigma_{x_p} \\ \vdots & & \ddots & \vdots \\ \rho_{p1}\sigma_{x_p}\sigma_{x_1} & \cdots & \cdots & \sigma_{x_p}^2 \end{pmatrix}. \quad (1)$$

Inference about the dispersion matrix  $\Sigma$  takes the general form

$$\begin{aligned} H_0 : \Sigma &= \Sigma_0, \\ H_1 : \Sigma &\neq \Sigma_0, \end{aligned} \quad (2)$$

where we assume that  $\Sigma_0$  is specified in a particular manner, for example, a block diagonal matrix or a spherical type structure or simply an unstructured form.

In general, research and testing methods of this form assume an underlying multivariate normal distribution with associated exact and approximate tests; for example, for a thorough overview and history of this testing problem, see Seber [1] and the references therewithin. In practice one can safely say that it would be rare that the multivariate normality

assumption holds. Hence, we were motivated to develop an exact permutation method approach to this problem. To the best of our knowledge no so-called exact permutation tests have been developed or explored with the exception of the very special case of  $p = 2$  dimensions and testing  $H_0 : \rho_{12} = 0$ ; for example, see Good [2]. Martin [3] provides a bootstrap algorithm for testing  $H_0 : \rho_{12} = \rho_{12,0}$ , which asymptotically can be shown to have the appropriate type I error rate. Unfortunately, the bootstrap methods given by Martin [3] relative to first standardizing the variables and rotating the data so as to transform the problem to the setting of testing  $H_0 : \rho_{12} = 0$  do not work in the permutation setting. The permutation test for the case  $H_0 : \rho_{12} = 0$  follows by permuting the second column of the  $n \times 2$  data matrix  $(\mathbf{X}_1, \mathbf{X}_2)$  and calculating the test statistic  $\hat{\rho}_{12}$ , where  $\hat{\rho}_{12}$  refers to the standard sample Pearson correlation coefficient, over all permutations. This can be done directly via a computationally expensive algorithm or via the more widely used Monte Carlo techniques. With respect to the Monte Carlo methods we generate  $B$  random permutations of the data and denote the permuted value of the test statistic by  $\hat{\rho}_{12,i}^*$ . Then the one-sided  $P$  value for the alternative  $H_1 : \rho_{12} > 0$  is given as  $\sum_{i=1}^B I(\hat{\rho}_{12,i}^* \geq \hat{\rho}_{12})/B$ , where the index  $i$  corresponds to a given permutation and  $I$  denotes the indicator function. Alternative approaches found in software packages such as SAS PROC FREQ (SAS version 9.3, Cary, NC) utilize hypergeometric probabilities similar to

how Fisher's exact test is carried out via treating the fixed data as discrete.

In general permutation testing is most often used for comparing two groups in the context of location differences or other features of distributions such as scale measures. Most of the theoretical work has been done in this setting such as type I error control. For a technical treatment of permutation testing see Romano [4] with respect to a theoretical examination for the behavior of the type I error control for permutation tests under exchangeability versus nonexchangeability conditions. In order to ensure true bounded type I error control in the permutation testing setting either the null hypothesis has to be specified in such a way that exchangeability holds by definition under  $H_0$  or some design feature such as randomization or matching needs to be employed. Commenges [5] studies the more general transformation approach used to preserve exchangeability. Also, see Zhang [6], Huang et al. [7], and Janssen and Pauls [8] with respect to the inflation of type I error rates when comparing means in the two-sample setting along with other types of comparisons. In terms of permutation testing related to correlation structure based hypotheses very little has been accomplished. This paper represents one of the only investigations of this type to date.

In Section 2 we develop the general  $p$ -dimensional exact tests given a prespecified covariance structure. Special cases include testing for sphericity, compound symmetry, and block diagonality, to name a few. This presentation is followed by a simulation study in Section 3. We then apply our method in Section 4 to an example involving repeated measures mice weight data.

## 2. Exact Tests for Covariance Structures in $p$ Dimensions

The focus of the work in this setting is with respect to two-sided alternatives. In certain instances a subset of these tests with one-sided alternative structure may be constructed. Those tests will not be included as part of this discussion due to the specificity of their applications.

*2.1. Unequal Variance Setting.* Now let  $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$  be an iid  $p$ -dimensional multivariate sample from an absolutely continuous distribution  $F$  with the first two finite central moments corresponding to each component of  $\mathbf{X}$  given as  $E(X_i) = \mu_{x_i}$  and  $\text{Var}(X_i) = \sigma_{x_i}^2$ ,  $i = 1, 2, \dots, p$ . Let  $\text{Corr}(X_i, X_j) = \rho_{ij}$ ,  $i = 1, 2, \dots, p$ ,  $j = 1, 2, \dots, p$ , ( $i \neq j$ ). Furthermore, denote the  $p \times p$  dispersion matrix of  $\mathbf{X}$  by

$$\Sigma = \begin{pmatrix} \sigma_{x_1}^2 & \rho_{12}\sigma_{x_1}\sigma_{x_2} & \cdots & \rho_{1p}\sigma_{x_1}\sigma_{x_p} \\ \rho_{21}\sigma_{x_2}\sigma_{x_1} & \sigma_{x_2}^2 & \cdots & \rho_{2p}\sigma_{x_2}\sigma_{x_p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1}\sigma_{x_p}\sigma_{x_1} & \cdots & \cdots & \sigma_{x_p}^2 \end{pmatrix}, \quad (3)$$

where  $\Sigma$  is defined to be a  $p \times p$  positive definite matrix. We represent the Cholesky decomposition of the  $p \times p$  matrix  $\Sigma$  as

$$\Sigma = \mathbf{A}'\mathbf{A}, \quad (4)$$

such that  $\mathbf{A}'^{-1}$  is defined. The Cholesky decomposition is a key component of the permutation test we propose; however it is not unique to the problem; that is, other decomposition methods may yield similar results and alternative solutions. From a practical standpoint the Cholesky decomposition is built in to several statistical softwarepackages, thus making our methodology more feasible for a larger group of practitioners.

Now our more general hypothesis of interest takes the form

$$\begin{aligned} H_0: \Sigma &= \Sigma_0, \\ H_1: \Sigma &\neq \Sigma_0, \end{aligned} \quad (5)$$

where  $\Sigma_0$  are the hypothesized value of  $\Sigma$  at (3).

*Test Statistic.* Let the  $p \times p$  matrix  $\mathbf{A}_0^{-1}$  denote the transpose of  $\mathbf{A}'^{-1}$  with the hypothesized values as given elements from (5). Let the  $n \times p$  matrix  $\mathbf{Z} = \mathbf{A}_0^{-1}\mathbf{X}$  denote the data matrix following transformation. Then the dispersion matrix corresponding to the  $n \times p$  matrix  $\mathbf{Z}$  will be a diagonal matrix such that  $\text{Corr}(Z_i, Z_j) = 0 \forall i < j, j < i$  if and only if  $H_0$  at (5) holds true. Under these conditions testing  $H_0$  at (5) is equivalent to the test:

$$\begin{aligned} H_0: \Sigma_z &= \Sigma_{z_0}, \\ H_1: \Sigma_z &\neq \Sigma_{z_0}, \end{aligned} \quad (6)$$

where the off-diagonal elements of  $\Sigma_{z_0}$  are equal to 0 under  $H_0$  at (6).

An exact  $\alpha$ -level permutation test of  $H_0$  can be defined for (6) by considering the permutation of each column of  $\mathbf{Z}$  and employing the Pearson correlation coefficient for each combination of columns. Towards this end let us denote the  $\text{Corr}(Z_i, Z_j) = \rho_{z_{ij}}$  with the corresponding Pearson estimator by  $\hat{\rho}_{z_{ij}}$ . The test statistic of interest in the two-sided case with respect to detecting departures from  $H_0$  at (6) is defined as

$$T(\mathbf{Z}) = \sum_{i=1}^{p-1} \sum_{j=i+1}^p \left| \hat{\rho}_{z_{ij}} \right|. \quad (7)$$

The exactness of the test in terms of the type I error control follows from a straightforward generalization of the form of the dispersion matrix for the  $2 \times 2$  case, where

$$\begin{aligned} H_0: \rho &= \rho_0, & \sigma_{x_1} &= \sigma_{x_{10}}, & \sigma_{x_2} &= \sigma_{x_{20}}, \\ H_1: \rho &\neq \rho_0, & \sigma_{x_1} &\neq \sigma_{x_{10}}, & \sigma_{x_2} &\neq \sigma_{x_{20}}. \end{aligned} \quad (8)$$

In the  $2 \times 2$  case the off-diagonal elements of the dispersion matrix are given as

$$\Sigma_{\mathbf{Z}|H_1} = \begin{pmatrix} \frac{\sigma_{x_1}^2}{\sigma_{x_{10}}^2} & \frac{\sigma_{x_1}(\rho\sigma_{x_{10}}\sigma_{x_2} - \rho_0\sigma_{x_1}\sigma_{x_{20}})}{\sqrt{1 - \rho_0\sigma_{x_{10}}^2\sigma_{x_{20}}^2}} \\ \frac{\sigma_{x_1}(\rho\sigma_{x_{10}}\sigma_{x_2} - \rho_0\sigma_{x_1}\sigma_{x_{20}})}{\sqrt{1 - \rho_0\sigma_{x_{10}}^2\sigma_{x_{20}}^2}} & \frac{\sigma_{x_{10}}^2\sigma_{x_2}^2 - 2\rho\rho_0\sigma_{x_1}\sigma_{x_2}\sigma_{x_{10}}\sigma_{x_{20}} + \rho_0^2\sigma_{x_1}^2\sigma_{x_{20}}^2}{(1 - \rho_0)\sigma_{x_{10}}^2\sigma_{x_{20}}^2} \end{pmatrix}. \tag{9}$$

An examination of the covariance term corresponding to  $\Sigma_{\mathbf{Z}|H_1}$  at (9) clearly indicates that it has the value of 0 if and only if  $H_0$  at (8) is true. When testing  $H_0$  at (8) the Pearson correlation estimate between the transformed variates through  $\mathbf{A}_0^{-1}$ ,  $Z_1$ , and  $Z_2$  serves to appropriately detect departures from  $H_0$ . Within the permutation testing framework provides an exact  $\alpha$ -level test; that is, the covariance of  $Z_1$  and  $Z_2$  is 0 if and only if the correlation of  $Z_1$  and  $Z_2$  is 0.

We resort to a Monte Carlo approximation in order to obtain the  $P$  value for testing  $H_0$  defined at (5). The steps for performing the Monte Carlo approximation with respect to estimating the  $P$  value are as follows.

- (1) Define  $H_0$  at (5).
- (2) Obtain the new random variates  $\mathbf{Z}$  by applying the transformation  $\mathbf{Z} = \mathbf{A}_0^{-1}\mathbf{X}$  to the observed data  $\mathbf{X}$ .
- (3) Calculate  $T(\mathbf{Z})$  at (7).
- (4) Permute each column of  $\mathbf{Z}$  independently such that we have the permuted  $n \times p$  matrix denoted by  $\mathbf{Z}^*$ .
- (5) Calculate  $T(\mathbf{Z}^*)$  applying the resampled values to  $T(\mathbf{Z})$  at (7).
- (6) Repeat steps (4) and (5)  $B$  times.
- (7) Calculate the Monte Carlo estimated permutation  $P$  value as  $p = (\sum_{l=1}^B I_{(T(\mathbf{Z}^*)_l > T(\mathbf{Z}))})/B$ , where  $I_{(\cdot)}$  denotes the indicator function.

**2.2. Nontransformation Special Cases.** In certain special cases we can test specific forms of hypothesis (5) using our permutation approach without specifying a specific subset of the  $\sigma_{x_i}^2$ 's or  $\rho_{ij}$ 's. One obvious special case relative to testing hypothesis (5) is the test given for a diagonal dispersion matrix versus nondiagonal dispersion matrix such that under  $H_0$  all  $\rho_{ij} = 0$ . Historical tests of this form have relied on assuming  $p$ -variate multivariate normality; for example, see Mudholkar et al. [9] for a description of a likelihood ratio approximation to this test. In this instance we have

$$\Sigma_0 = \begin{pmatrix} \sigma_{x_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{x_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_{x_p}^2 \end{pmatrix}, \tag{10}$$

with unspecified  $\sigma_{x_i}^2$ 's under  $H_0$ .

In this case there is no transformation of the data required. An exact  $\alpha$ -level permutation test of  $H_0$  can be

defined simply by considering the permutation of each column of  $\mathbf{X}$  and employing the Pearson correlation coefficient for each combination of columns. Towards this end denote the  $\text{Corr}(X_i, X_j) = \rho_{x_{ij}}$  with the corresponding Pearson estimator by  $\hat{\rho}_{x_{ij}}$ . The test statistic of interest with respect to detecting departures from the diagonal structure is defined as

$$T(\mathbf{X}) = \sum_{i=1}^{p-1} \sum_{j=i+1}^p |\hat{\rho}_{x_{ij}}|. \tag{11}$$

The Monte Carlo estimated permutation  $P$  value is calculated similarly as before, where  $p = (\sum_{l=1}^B I_{(T(\mathbf{X}^*)_l > T(\mathbf{X}))})/B$ , where  $I_{(\cdot)}$  denotes the indicator function.

Another special case where we can have a set of unspecified  $\sigma_{x_i}^2$ 's or  $\rho_{ij}$ 's is when we may be interested in testing for a block diagonal dispersion matrix structure such that under  $H_0$  at (5) we now have

$$\Sigma_0 = \begin{pmatrix} \Sigma_{11} & 0 & \cdots & 0 \\ 0 & \Sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \Sigma_{bb} \end{pmatrix}, \tag{12}$$

where the partitioned  $q_j \times q_j$  matrices are given as

$$\Sigma_{jj} = \begin{pmatrix} \sigma_{x_1}^2 & \rho_{12}\sigma_{x_1}\sigma_{x_2} & \cdots & \rho_{1q}\sigma_{x_1}\sigma_{x_q} \\ \rho_{21}\sigma_{x_2}\sigma_{x_1} & \sigma_{x_2}^2 & \cdots & \rho_{2q}\sigma_{x_2}\sigma_{x_q} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{q1}\sigma_{x_q}\sigma_{x_1} & \cdots & \cdots & \sigma_{x_q}^2 \end{pmatrix}. \tag{13}$$

The dispersion matrix  $\Sigma_{jj}$  may have different dimensions  $q_j < p$  ( $\sum q_j = p$ ),  $j = 1, 2, \dots, b$ , with unspecified  $\sigma_{x_i}^2$ 's and  $\rho_{ij}$ 's under  $H_0$ .

As in the test for a diagonal dispersion matrix above there is no transformation of the data required. An exact  $\alpha$ -level permutation test of  $H_0$  can again be defined simply by considering the permutation of each column of  $\mathbf{X}$  and employing the Pearson correlation coefficient for each combination of columns. The test statistic of interest with respect to detecting departures from the diagonal structure is a slight modification of the test statistic at (11) defined as

$$T(\mathbf{X}) = \sum_{i=1}^{p-1} \sum_{j=i+1}^p |\hat{\rho}_{x_{ij}}| I_{(\rho_{0,ij}=0)}, \tag{14}$$

where the "off-block" correlation elements at (12),  $\rho_{0,ij} = 0$ , under  $H_0$  and  $I_{(\cdot)}$  denote the indicator function. The Monte

Carlo estimated permutation  $P$  value is calculated similarly as before, where  $p = (\sum_{l=1}^B I_{(T(\mathbf{X}^*)_{(l)} > T(\mathbf{X}))})/B$  and  $I_{(\cdot)}$  denotes the indicator function.

2.3. *Equal Variance Setting.* For the equal variance  $p$ -dimensional case we have

$$\begin{aligned} H_0 : \Sigma &= \Sigma_0, \\ H_1 : \Sigma &\neq \Sigma_0, \end{aligned} \tag{15}$$

where we now define the  $p \times p$  dispersion matrix under  $H_0$  as

$$\Sigma_0 = \sigma \Gamma_0, \tag{16}$$

where  $\Sigma_0$  is defined to be a  $p \times p$  positive definite matrix,  $\text{Var}(\mathbf{X}_i) = \sigma^2$ ,  $i = 1, 2, \dots, p$ , under  $H_0$  and the  $p \times p$  correlation matrix  $\Gamma_0$  is given by

$$\Gamma_0 = \begin{pmatrix} 1 & \rho_{12,0} & \cdots & \rho_{1p,0} \\ \rho_{21,0} & 1 & \cdots & \rho_{2p,0} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{p1,0} & \cdots & \cdots & 1 \end{pmatrix}. \tag{17}$$

The Cholesky decomposition of matrix  $\Gamma_0$  is as

$$\Gamma_0 = \mathbf{B}'_0 \mathbf{B}_0. \tag{18}$$

In order to test the hypothesis at (15) we utilize the transformation  $\mathbf{Z} = \mathbf{B}'_0^{-1} \mathbf{X}$ . Note that  $H_0$  at (15) will be sensitive to departures from  $\Gamma_0$  and unequal marginal variances. Furthermore, explicit values for  $\sigma^2$  do not need to be specified within this hypothesis testing framework.

The steps for performing the Monte Carlo approximation with respect to estimating the  $P$  value are as follows.

- (1) Define  $H_0$  at (15).
- (2) Obtain the new random variates  $\mathbf{Z}$  by applying the transformation  $\mathbf{Z} = \mathbf{B}'_0^{-1} \mathbf{X}$  to the observed data  $\mathbf{X}$ .
- (3) Calculate  $T(\mathbf{Z})$  at (7).
- (4) Permute each column of  $\mathbf{Z}$  independently such that we have the permuted  $n \times p$  matrix denoted by  $\mathbf{Z}^*$ .
- (5) Calculate  $T(\mathbf{Z}^*)$ .
- (6) Repeat steps (4) and (5)  $B$  times.
- (7) Calculate the Monte Carlo estimated permutation  $P$  value as  $p = (\sum_{l=1}^B I_{(T(\mathbf{Z}^*)_{(l)} > T(\mathbf{Z}))})/B$ , where  $I_{(\cdot)}$  denotes the indicator function.

A special case relative to testing hypothesis (15) is the test that the dispersion matrix is diagonal and all  $\sigma_{x_i} = \sigma$ ,  $i = 1, 2, \dots, p$ .

Other special cases of the test at (15) may be of interest and written in the form

$$\begin{aligned} H_0 : \Sigma &= \sigma \Gamma_0(\rho_0), \\ H_1 : \Sigma &\neq \sigma \Gamma_0(\rho_0). \end{aligned} \tag{19}$$

Examples of specific dispersion structures of importance corresponding to the test at (19) include

(1) sphericity:

$$\Gamma_0(\rho_0) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}; \tag{20}$$

(2) compound symmetry:

$$\Gamma_0(\rho_0) = \begin{pmatrix} 1 & \rho_0 & \cdots & \rho_0 & \rho_0 \\ \rho_0 & 1 & \cdots & \rho_0 & \rho_0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_0 & \rho_0 & \cdots & 1 & \rho_0 \\ \rho_0 & \rho_0 & \cdots & \rho_0 & 1 \end{pmatrix}; \tag{21}$$

(3) first-order autoregressive:

$$\Gamma_0(\rho_0) = \begin{pmatrix} 1 & \rho_0 & \cdots & \rho_0^{p-1} & \rho_0^p \\ \rho_0 & 1 & \cdots & \rho_0^{p-2} & \rho_0^{p-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_0^{p-1} & \rho_0^{p-2} & \cdots & 1 & \rho_0 \\ \rho_0^p & \rho_0^{p-1} & \cdots & \rho_0 & 1 \end{pmatrix}; \tag{22}$$

(4) spatial power:

$$\Gamma_0(\rho_0) = \begin{pmatrix} 1 & \rho_0^{d_{1,2}} & \cdots & \rho_0^{d_{1,p-1}} & \rho_0^{d_{1,p}} \\ \rho_0^{d_{2,1}} & 1 & \cdots & \rho_0^{d_{2,p-2}} & \rho_0^{d_{2,p-1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_0^{d_{p-1,1}} & \rho_0^{d_{p-1,2}} & \cdots & 1 & \rho_0^{d_{p-1,p}} \\ \rho_0^{d_{p,1}} & \rho_0^{d_{p,2}} & \cdots & \rho_0^{d_{p,p-1}} & 1 \end{pmatrix}. \tag{23}$$

Several other well-known spatial dispersion matrices similar to the spatial power matrix presented above fit within this same framework and will not be presented here.

### 3. Simulation Study

In this section we examine the test at (19), where we specify  $\rho_0$  and the form of the correlation structure  $\Gamma$  at (17), for example, compound symmetry. Our simulation study for the  $p \times p$  case will utilize a  $p$ -variate standardized multivariate normal distribution with  $p = 5$  and a special case mixing the marginal distributions across normal, exponential, and uniform forms. Again, differing location and scale do not vary the general conclusions. In terms of our simulation study we set the null value of  $\rho_0 = 0, 0.5, 0.9$  under a compound symmetry assumption and a first-order autoregressive assumption, where  $\Gamma$  will take the forms:

(1) compound symmetry:

$$\Gamma_0(\rho_0) = \begin{pmatrix} 1 & \rho_0 & \cdots & \rho_0 & \rho_0 \\ \rho_0 & 1 & \cdots & \rho_0 & \rho_0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_0 & \rho_0 & \cdots & 1 & \rho_0 \\ \rho_0 & \rho_0 & \cdots & \rho_0 & 1 \end{pmatrix}; \tag{24}$$

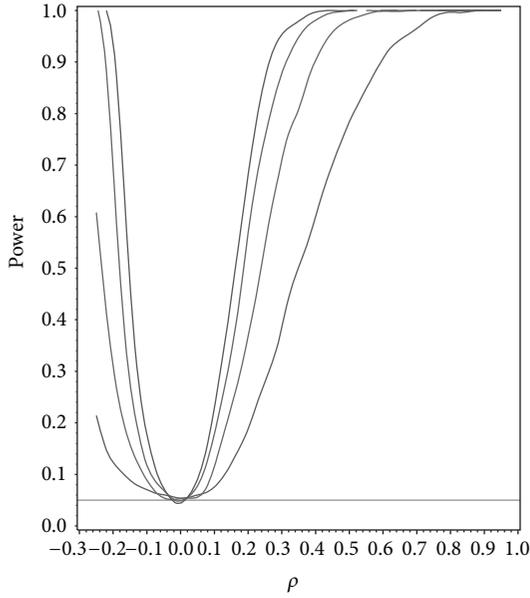


FIGURE 1: Power for testing for compound symmetry covariance as a function of  $\rho$  given  $\rho_0 = 0$ .

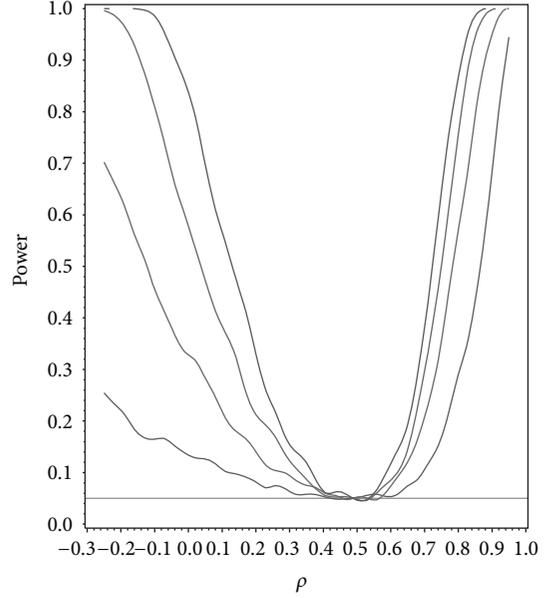


FIGURE 2: Power for testing for compound symmetry covariance as a function of  $\rho$  given  $\rho_0 = 0.5$ .

(2) first-order autoregressive:

$$\Gamma_0(\rho_0) = \begin{pmatrix} 1 & \rho_0 & \cdots & \rho_0^{p-1} & \rho_0^p \\ \rho_0 & 1 & \cdots & \rho_0^{p-2} & \rho_0^{p-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_0^{p-1} & \rho_0^{p-2} & \cdots & 1 & \rho_0 \\ \rho_0^p & \rho_0^{p-1} & \cdots & \rho_0 & 1 \end{pmatrix}. \quad (25)$$

Note that the special case  $\rho_0 = 0$  is the same for both covariance structures and is only presented once. It should also be noted that under the assumption of multivariate normality testing  $H_0 : \Sigma = \sigma\Gamma_0(0)$  under the compound symmetry or first-order autoregressive structure is a special case (equal variance assumption) of the well-known test for “complete independence”; for example, see Mudholkar et al. [9]. Under nonnormality we are essentially testing the “complete uncorrelated” case. In this special case the methods presented here are the first exact methods developed for tackling this particular hypothesis. In terms of large sample theory around similar results see Jiang [10] and Xiao and Wu [11].

For our simulation study we used 1000 replicates for our study at  $n = 10, 20, 30, 40$  and set  $\alpha = 0.05$ . The covariance structure was the same under  $H_0$  and  $H_1$  for this set of simulations. The results are contained in Figures 1, 2, 3, 4, and 5. As anticipated we see the expected results of appropriate type I error control and monotone power functions increasing in either direction about the null value for  $\rho$ . The range of  $\rho$  under the alternative was dictated by the constraint that  $\Gamma_0(\rho_0)$  is defined to be positive definite.

For the sake of example we modified our simulation and took  $\rho_0 = 0.5$  with marginals given by  $X_1, X_2 \sim N(0, 1)$ ,  $X_3, X_4 \sim U(-1/\sqrt{12}, 1/\sqrt{12})$ , and  $X_5 \exp(1, -1)$  with  $\Gamma_0(\rho_0)$

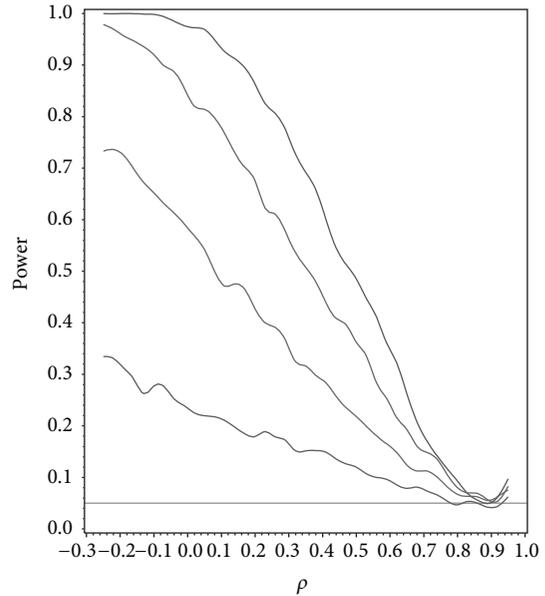


FIGURE 3: Power for testing for compound symmetry covariance as a function of  $\rho$  given  $\rho_0 = 0.9$ .

assumed to be compound symmetric under  $H_0$  and under  $H_1$ . The results are shown in Figure 6 and as we can see they do not differ dramatically from Figure 2 assuming multivariate normality, thus illustrating the flexibility and nonparametric nature of our methodology.

As an additional result we studied the power under the correctly specified  $\rho$  under  $H_0$  with  $\Gamma_0(\rho_0)$  differing in structure. For this study we set  $\rho_0 = 0.5, 0.9$  with  $\Gamma_0(\rho_0)$  set to compound symmetry under  $H_0$  and  $\Gamma_0(\rho_0)$  set to the first-order autoregressive structure under  $H_1$ . In other words

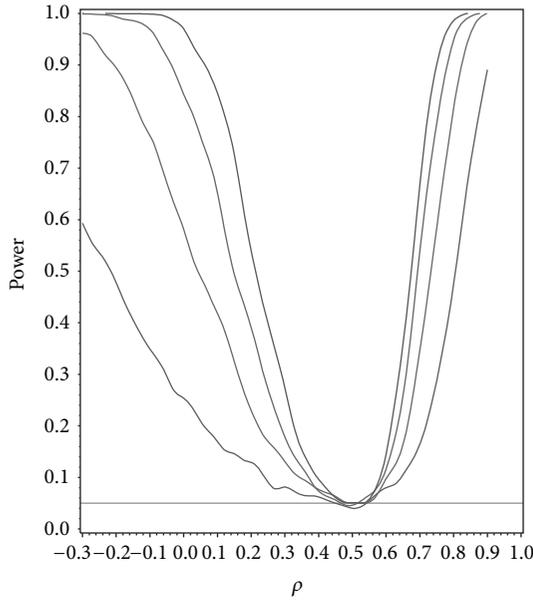


FIGURE 4: Power for testing for first-order autoregressive covariance as a function of  $\rho$  given  $\rho_0 = 0.5$ .

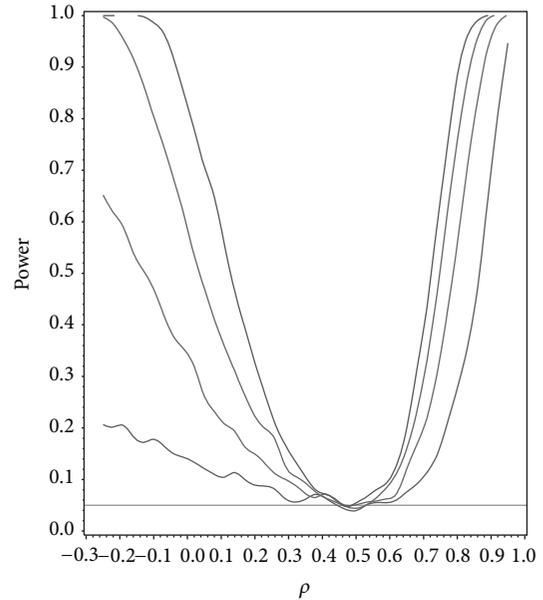


FIGURE 6: Power for testing for compound symmetry covariance as a function of  $\rho$  given  $\rho_0 = 0.5$  and mixed marginal distributions.

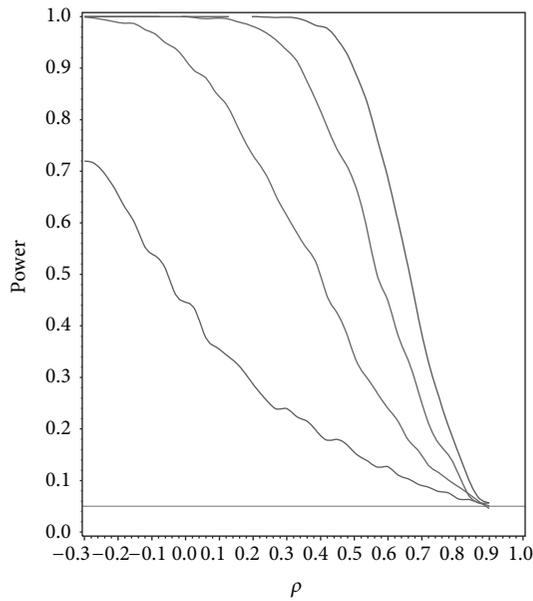


FIGURE 5: Power for testing for first-order autoregressive covariance as a function of  $\rho$  given  $\rho_0 = 0.9$ .

what is the power to detect a correlation structure different from the null structure given that  $\rho_0$  is the true correlation. At  $\alpha = 0.05$  the power to detect a different correlation structure under the alternative for  $\rho_0 = 0.5$  and  $0.9$  at  $n = 10, 20, 30, 40$  and  $\alpha = 0.05$  was  $0.245, 0.539, 0.761, \text{ and } 0.894$  and  $0.520, 0.858, 0.951, \text{ and } 0.998$ , respectively.

#### 4. Example

As an illustration of our method we will use phenotypic weight data from  $n = 16$  mice as contained in Table 1 from

a recent unpublished study conducted within Roswell Park Cancer Institute. The estimated correlation matrix is provided in Table 2. The respective sample variance estimates were  $\hat{\sigma}_{x_1}^2 = 3.2, \hat{\sigma}_{x_2}^2 = 3.3, \hat{\sigma}_{x_3}^2 = 2.6, \hat{\sigma}_{x_4}^2 = 3.4, \text{ and } \hat{\sigma}_{x_5}^2 = 3.1$ . For example, suppose we were interested in testing

$$\begin{aligned} H_0 : \Sigma &= \sigma \Gamma_0(.95), \\ H_1 : \Sigma &\neq \sigma \Gamma_0(.95), \end{aligned} \tag{26}$$

for both  $\Gamma(\cdot)$  having the compound symmetry structure or the first-order autoregressive structure as defined in (23). In this instance the test corresponding to the above hypothesis under a compound symmetry correlation structure yielded a Monte Carlo estimated  $P$  value  $< 0.0001$  ( $B = 10,000$ ). While the test corresponding to the above hypothesis under the first-order autoregressive correlation structure yielded a Monte Carlo estimated  $P$  value  $= 0.002$  ( $B = 10,000$ ). For this example this provides some measure of evidence that the correlation structure does not fit the compound symmetry structure and that the first-order autoregressive structure assuming  $\rho = 0.95$  may be more appropriate. Similarly, the test for diagonality under the equal variances assumption (sphericity), which does not assume a value for  $\rho_0$ , yielded a Monte Carlo estimated  $P$  value  $< 0.0001$  ( $B = 10,000$ ). Note that we may be rejecting  $H_0$  at some specified level  $\alpha$  under at least one of 3 scenarios: unequal marginal variances,  $\rho \neq \rho_0$  or  $\Gamma \neq \Gamma_0$ .

Given our overall  $P$  value from above which was  $0.002$ , we may wish to examine in further detail what is driving us to reject  $H_0$ . In this case we can examine specific submatrices of the dispersion matrix of  $\mathbf{Z} = \mathbf{B}'^{-1}\mathbf{X}$ . For this example we could test  $H_0 : \Sigma = \sigma \Gamma_0(.95)$  using days 0, 1, and 2, only or any other combinations of days such that the appropriate correlation substructure is extracted from the

TABLE 1: Mice weights (grams).

Mouse	Day 0	Day 1	Day 2	Day 3	Day 4
1	21.4	21.0	21.0	21.3	21.5
2	18.4	18.4	18.0	18.1	17.8
3	17.8	17.6	17.4	17.2	17.1
4	18.2	18.5	17.6	18.0	17.5
5	20.0	19.4	19.1	19.6	18.9
6	20.2	19.3	19.3	19.3	18.6
7	16.5	16.3	16.6	16.4	16.3
8	17.6	17.4	17.5	17.9	17.6
9	19.8	20.4	20.1	20.6	20.3
10	22.0	22.3	21.3	21.9	20.7
11	17.5	17.4	17.5	17.5	17.3
12	20.3	20.2	19.8	20.4	19.7
13	16.3	16.2	16.2	15.9	15.7
14	18.0	17.2	17.2	17.0	16.4
15	20.3	19.9	19.0	18.9	18.1
16	21.4	21.2	20.7	21.1	20.7

TABLE 2: Estimated correlation matrix for mouse weight example data.

	Day 0	Day 1	Day 2	Day 3	Day 4
Day 0	1.000	0.977	0.974	0.959	0.921
Day 1	0.977	1.000	0.982	0.979	0.945
Day 2	0.974	0.982	1.000	0.993	0.979
Day 3	0.959	0.979	0.993	1.000	0.983
Day 4	0.921	0.945	0.979	0.983	1.000

original hypothesized values for  $\Gamma$ . For our example subtest we get  $P = 0.32$  indicating no strong evidence against a first-order autoregressive “substructure” with equal variances and  $\rho_0 = 0.95$ . If we add day 3, our  $P$  value = 0.04, indicating either the correlation structure may be misspecified at this point or the variance is different. Note that further work relative to the multiple comparison problem of subtests and their relative correlation is needed. This is simply an exploratory approach to this issue relative to the example at hand.

### 5. Concluding Remarks

In this note we provided a method for exact testing around specific covariance structures. We employed the Cholesky decomposition for this purpose. It was noted by a reviewer that other decomposition methodologies may lead to extensions of this methodology, which we will consider in terms of future work.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgments

This research is supported by the NIH Grant IR03DE020851-01A1, the National Institute of Dental and Craniofacial Research. The authors wish to thank the reviewers for their time and effort.

### References

- [1] G. A. F. Seber, *Multivariate Observations*, John Wiley & Sons, New York, NY, USA, 1984.
- [2] P. Good, “Robustness of Pearson correlation,” *Interstat*, vol. 15, no. 5, pp. 1–6, 2009.
- [3] M. A. Martin, “Bootstrap hypothesis testing for some common statistical problems: a critical evaluation of size and power properties,” *Computational Statistics & Data Analysis*, vol. 51, no. 12, pp. 6321–6342, 2007.
- [4] J. P. Romano, “On the behavior of randomization tests without a group invariance assumption,” *Journal of the American Statistical Association*, vol. 85, no. 411, pp. 686–692, 1990.
- [5] D. Commenges, “Transformations which preserve exchangeability and application to permutation tests,” *Journal of Non-parametric Statistics*, vol. 15, no. 2, pp. 171–185, 2003.
- [6] S. P. Zhang, “The split sample permutation  $t$ -tests,” *Journal of Statistical Planning and Inference*, vol. 139, no. 10, pp. 3512–3524, 2009.
- [7] Y. Huang, H. Xu, V. Calian, and J. C. Hsu, “To permute or not to permute,” *Bioinformatics*, vol. 22, no. 18, pp. 2244–2248, 2006.
- [8] A. Janssen and T. Pauls, “A Monte Carlo comparison of studentized bootstrap and permutation tests for heteroscedastic two-sample problems,” *Computational Statistics*, vol. 20, no. 3, pp. 369–383, 2005.
- [9] G. S. Mudholkar, M. C. Trivedi, and C. T. Lin, “An approximation to the distribution of the likelihood ratio statistics for testing complete independence,” *Technometrics*, vol. 24, no. 2, pp. 139–143, 1982.
- [10] T. Jiang, “The asymptotic distributions of the largest entries of sample correlation matrices,” *The Annals of Applied Probability*, vol. 14, no. 2, pp. 865–880, 2004.
- [11] H. Xiao and W. B. Wu, “Asymptotic theory for maximum deviations of sample covariance matrix estimates,” *Stochastic Processes and their Applications*, vol. 123, no. 7, pp. 2899–2920, 2013.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

