The Korean Communications
in Statistics Vol. 5, No. 3, 1998
pp. 591-597

On the Distribution of the Scaled Residuals under Multivariate Normal Distributions

Cheolyong Park[1)]

Abstract

We prove (at least empirically) that some forms of the scaled residuals calculated from i.i.d. multivariate normal random vectors are ancillary. We further show that, if the scaled residuals are ancillary, then they have the same distribution whatever form of rotation is used to remove sample correlations.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous p-variate distribution with mean vector μ and nonsingular covariance matrix Σ and let $\overline{X}$ and S be the sample mean vector and sample covariance matrix of $X_1, X_2, \ldots, X_n$, respectively. Here we assume $n > p$ so that S is nonsingular. To remove the sample means and sample correlations, we use the scaled residuals

$$Z_i(R) = R(S)(X_i - \overline{X}), \quad i = 1, 2, \ldots, n, \tag{1}$$

where a $p \times p$ matrix $R(S)$ is chosen such that the sample covariance matrix of $Z_1(R), Z_2(R), \ldots, Z_n(R)$ becomes the identity matrix, i.e. $R(S)\, S R^t(S) = I$ with "t" being the notation for transpose. There are many ways of choosing such $R(S)$ but we will mainly focus on two special choices: the square root matrix $R(S) = S^{-1/2}$ or the lower triangular matrix $R(S) = L^{-1} \in C$, where C is the class of all lower triangular matrices with positive diagonal elements and L is the unique element in C such that $LL^t = S$. For the convenience of notations we will denote the scaled residuals based on the square root and lower triangular method by Z_i^{sr} and Z_i^{tr}, respectively. Note that the notation $R(S)$ is used to clearly state that R is the function of $X_1, X_2, \ldots, X_n$ only through S.

In this paper, we are interested in the distribution of the scaled residuals $Z_i(R)$ when

1) Assistant Professor, Department of Statistics, Keimyung University, Taegu 704-701

$X_1, X_2, \ldots, X_n$ are from some multivariate normal distribution. First, we start with Z_i^{tr}, a special form of the scaled residuals. We show that Z_i^{tr} are ancillary and then express the distribution of Z_i^{tr} as a conditional distribution of X_i given the sufficient statistics $\overline{X}$ and S. We next focus on the distribution of the scaled residuals $Z_i(R)$ in the case where they are ancillary. In this case, we can show that the distribution of $Z_i(R)$ is the same as that of Z_i^{tr}.

Many tests for the multivariate normality of $X_1, X_2, \ldots, X_n$ are based on Z_i^{sr}. For example, Moore and Stubblebine's(1981) chi-squared test and Mardia's(1970) skewness and kurtosis tests are based on the radii $d_i = \|Z_i^{sr}\|$, where $\|\cdot\|$ is the Euclidean norm on R^p. Also the Rayleigh statistic(Koziol, 1983) is based on the unit residuals $U_i = Z_i^{sr}/d_i$, and Quiroz and Dudley's(1991) tests are based on both the radii and unit residuals. However, there has been little study on the distribution of Z_i^{sr} under normality. If Z_i^{sr} are ancillary, then they will have the same distribution as Z_i^{tr}. However, as noted in Quiroz and Dudley(1991, p.544), it is not easy to prove that Z_i^{sr} are ancillary. Therefore, we provide a simulation study to empirically show that they are ancillary.

In Section 2, we present main results with proofs. We show that Z_i^{tr} are ancillary and express the distribution of Z_i^{sr} as a conditional distribution. We then show that, if the scaled residuals $Z_i(R)$ are ancillary, then $Z_i(R)$ have the same distribution as Z_i^{tr}. In Section 3, we will provide a simulation study to support the conjecture that Z_i^{sr} are ancillary.

2. Main Results

In this section, we assume that $X_1, X_2, \ldots, X_n$ are a random sample from the p-variate normal distribution $N_p(\mu, \Sigma)$ where Σ is nonsingular and $n > p$. First of all, we provide an easy and self-contained proof of the ancillarity of Z_i^{tr} and express the distribution of Z_i^{tr} as a conditional distribution of X_i given the sufficient statistics $\overline{X}$ and S. We, then, show that, if the scaled residuals $Z_i(R)$ are ancillary, then they have the same distribution as Z_i^{tr}.

We first prove that Z_i^{tr} are ancillary. The main idea of proof is to show that Z_i^{tr} are invariant on the transformation based on lower triangular matrices with positive diagonal

elements. In other words, we will show that $Z_i^{tr}(X)=Z_i^{tr}(LX)$ for all $L\in C$, where $Z_i^{tr}(X)$ is a notation used to clearly state that it is a function of $X_1, X_2, \ldots, X_n$.

Theorem 1. Z_i^{tr} are ancillary.

Proof: Since Z_i^{tr} depend only on $X_1-\overline{X}, \ldots, X_n-\overline{X}$, it is clear that the distribution of Z_i^{tr} is free of μ. Thus, we may assume from now on that $X_1, \ldots, X_n$ are from $N_p(0, \Sigma)$.

We will now prove that the distribution of Z_i^{tr} is free of Σ. Let $L^*\in C$ be arbitrary but fixed. Let $L\in C$ be such that $LL^t=S$ and let $Y_i=L^*X_i$ for $i=1,2,\ldots,n$. Then, the sample mean vector $\overline{Y}$ and sample covariance matrix S_y of $Y_1, Y_2, \ldots, Y_n$ are given by

$$\overline{Y}=L^*\overline{X}, \quad S_y=L^*SL^{*t}=(L^*L)(L^*L)^t.$$

It is clear that L^*L belongs to C and so we have

$$Z_i^{tr}(Y)=(L^*L)^{-1}(Y_i-\overline{Y})=L^{-1}(X_i-\overline{X})=Z_i^{tr}(X).$$

This shows that $Z_i^{tr}(X)=Z_i^{tr}(L^*X)$ for all $L^*\in C$.

Choose $L^*\in C$ to be such that $L^*\Sigma(L^*)^t=I$, then $Y_i=L^*X_i$ have the distribution $N_p(0, I)$, which is free of Σ. Thus the distribution of $Z_i^{tr}(X)=Z_i^{tr}(L^*X)$ is free of Σ and this completes the proof. □

Next, we express the distribution of Z_i^{tr} as a conditional distribution of X_i given the sufficient statistics $\overline{X}$ and S. Since Z_i^{tr} are ancillary, we may assume that the sampling model is $N_p(0, I)$, which will be denoted by $\mathcal{L}_0$. By Basu's Theorem, the sufficient statistics $\overline{X}$ and S are independent of Z_i^{tr}. Thus we have

$$\mathcal{L}_0(Z_i^{tr})=\mathcal{L}_0(Z_i^{tr}\mid \overline{X}=0, S=I)=\mathcal{L}_0(X_i\mid \overline{X}=0, S=I), \tag{2}$$

where the last equality follows since Z_i^{tr} becomes X_i given that $\overline{X}=0$, $S=I$.

Now we focus on the distribution of the scaled residuals $Z_i(R)$ given in (1). Next theorem shows that the ancillarity of the scaled residuals is a sufficient condition for that they have the same distribution as (2).

Theorem 2. If the scaled residuals $Z_i(R)$ in (1) are ancillary, then the distribution of $Z_i(R)$ is given by (2).

Proof: By the same argument given just prior to (2), we have

$$\mathcal{L}_0(Z_i(R)) = \mathcal{L}_0(R(I)X_i \mid \overline{X}=0, S=I).$$

Since $R(S)$ is chosen such that $R(S)\,S\,R^t(S)=I$, a $p\times p$ matrix $R(I)$ satisfies $R(I)\,R^t(I)=I$. Therefore $R(I)$ is an orthogonal matrix of constants and this implies that both $Y_i=R(I)\,X_i$ and X_i have the same distribution $N_p(0,I)$. Also, it is easy to show that $\overline{Y}=0$ and $S_y=I$, given that $\overline{X}=0$ and $S=I$. This shows that the distribution of $Z_i(R)$ is the same as (2). This completes the proof. □

3. A Simulation Study on the Distribution of Z_i^{sr}

In this section, we provide a simulation study on the distribution of Z_i^{sr}. As noted in Theorem 2, the ancillarity of Z_i^{sr} is a crucial part in the study of the distribution. If Z_i^{sr} are ancillary, then Z_i^{sr} and Z_i^{tr} have the same distribution (2). However, as noted in Section 1, it is not easy to prove that the distribution of Z_i^{sr} is free of Σ. (The distribution of Z_i^{sr} is clearly free of μ since Z_i^{sr} depend only on $X_1-\overline{X},\ldots,X_n-\overline{X}$.)

Thus we provide a simulation study to support the conjecture that the distribution of Z_i^{sr} is free of Σ. Here are some restrictions we impose on this simulation study:

1. We consider small samples such as $n=10,20$ since Z_i^{sr} are asymptotically $N_p(0,I)$ for large n.
2. Only Z_1^{sr} is considered since Z_i^{sr} have the same distribution for all i.
3. We restrict ourselves to the bivariate case, i.e. $p=2$.
4. The sampling models we consider are $N_2(\mu,\Sigma)$ with $\mu=0$ and four covariance matrices:

$$\Sigma_0=\begin{pmatrix}1&0\\0&1\end{pmatrix},\ \Sigma_1=\begin{pmatrix}1&0\\0&4\end{pmatrix},\ \Sigma_2=\begin{pmatrix}1&.25\\.25&1\end{pmatrix},\ \Sigma_3=\begin{pmatrix}1&.75\\.75&1\end{pmatrix}.$$

These four models will be denoted by Models 0-3, respectively.

5. We compare only the marginal distributions of Z_1^{sr} since the sample correlations among the components of Z_i^{sr} are all 0 and so the components are asymptotically independent.

Next, we will explain how our simulation study is conducted. For each sample size

$n=10,20$, we generate 500 independent vectors Z_1^{sr} from each sampling model and then compare Z_1^{sr}'s from Models 1-3 with those from Model 0. Comparison of Z_1^{sr}'s from two Models are made by applying univariate techniques to each component of Z_1^{sr}'s: Two methods we use are informal quantile-quantile plot(Q-Q plot) and formal Kolmogorov-Smirnov two-sample test(K-S test). S-Plus is used to generate random numbers and then to make Q-Q plots and K-S tests. The Q-Q plots and P-values of K-S test are given in Figure 1 for $n=10$ and in Figure 2 for $n=20$.

In each figure, there are six plots arranged in 3×2 matrix format. The i-th row compares the components of Z_1^{sr}'s from Model i with those from Model 0. The j-th column compares the j-th component of Z_1^{sr}'s from Models 1-3 with that from Model 0. Thus the (i,j) plot compares the j-the component of Z_1^{sr}'s from Model i (plotted as y-axis) with that from Model 0 (plotted as x-axis). We also provide a reference line with intercept 0 and slope 1. The P-value of K-S test is given on top of each plot. From Figures 1 and 2, we find that all points fall fairly close to the reference line and that the points in most Q-Q plots of $n=20$ are closer to the reference line than those of $n=10$. These findings are well confirmed by the P-values of K-S test. Thus, even though this simulation is limited to bivariate case, we can infer that the distribution of Z_1^{sr} does not depend on Σ.

References

[1] Koziol, J.A. (1983). On Assessing Multivariate Normality. *Journal of the Royal Statistical Society-Series B*, Vol. 45, 358-361.

[2] Mardia, K.V. (1970). Measures of Multivariate Skewness and Kurtosis with Applications. *Biometrika*, Vol. 57, 519-530.

[3] Moore, D.D, and Stubblebine, J.B. (1981). Chi-square Test for Multivariate Normality with Application to Common Stock Prices. *Communications in Statistics-Theory and Methods*, Vol. 10, 713-738.

[4] Quiroz, A.J., and Dudley, R.M. (1991). Some New Tests for Multivariate Normality. *Probability Theory and Related Fields*, Vol. 87, 521-546.

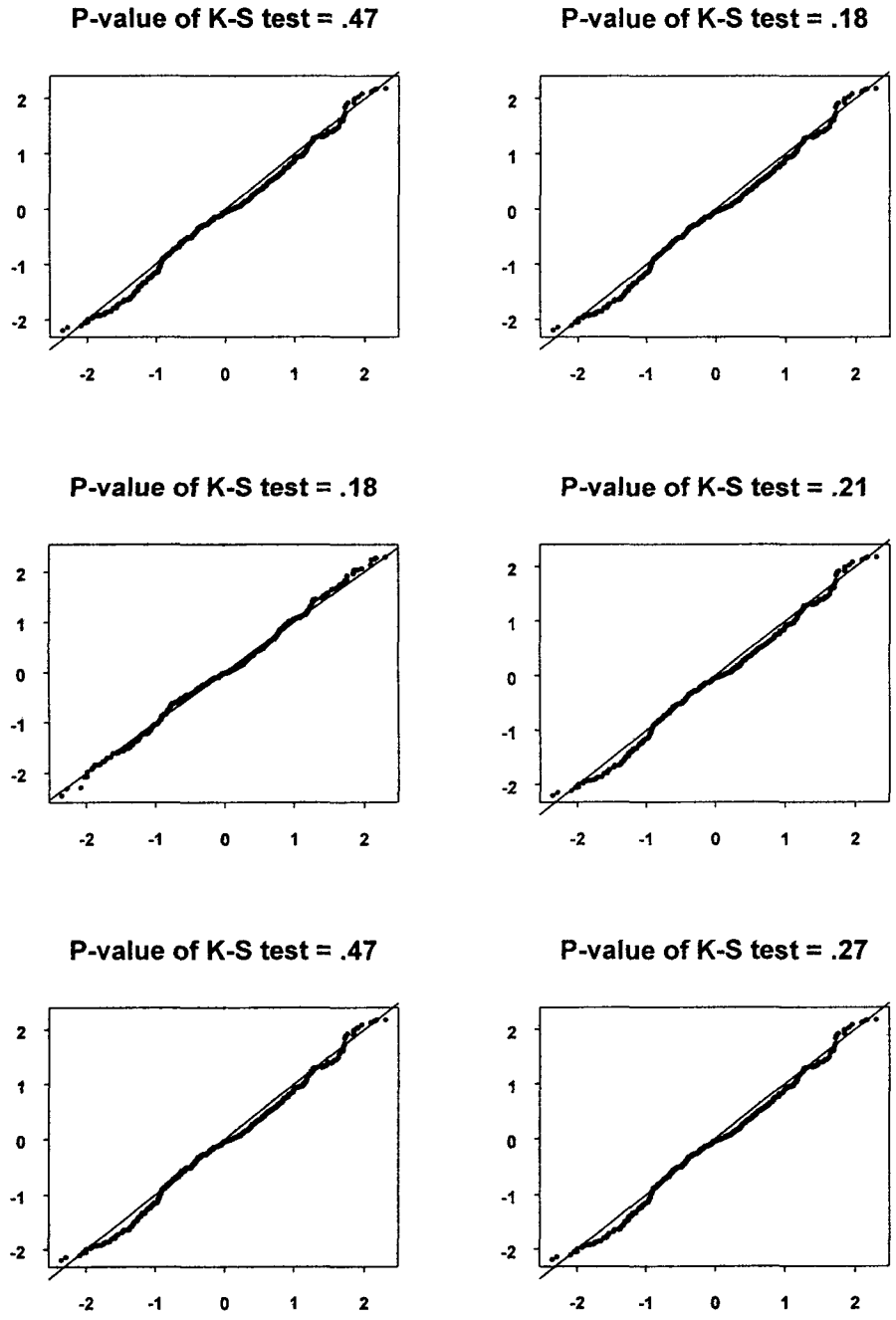

Figure 1. Q-Q plots and P-values of K-S tests for $n = 10$

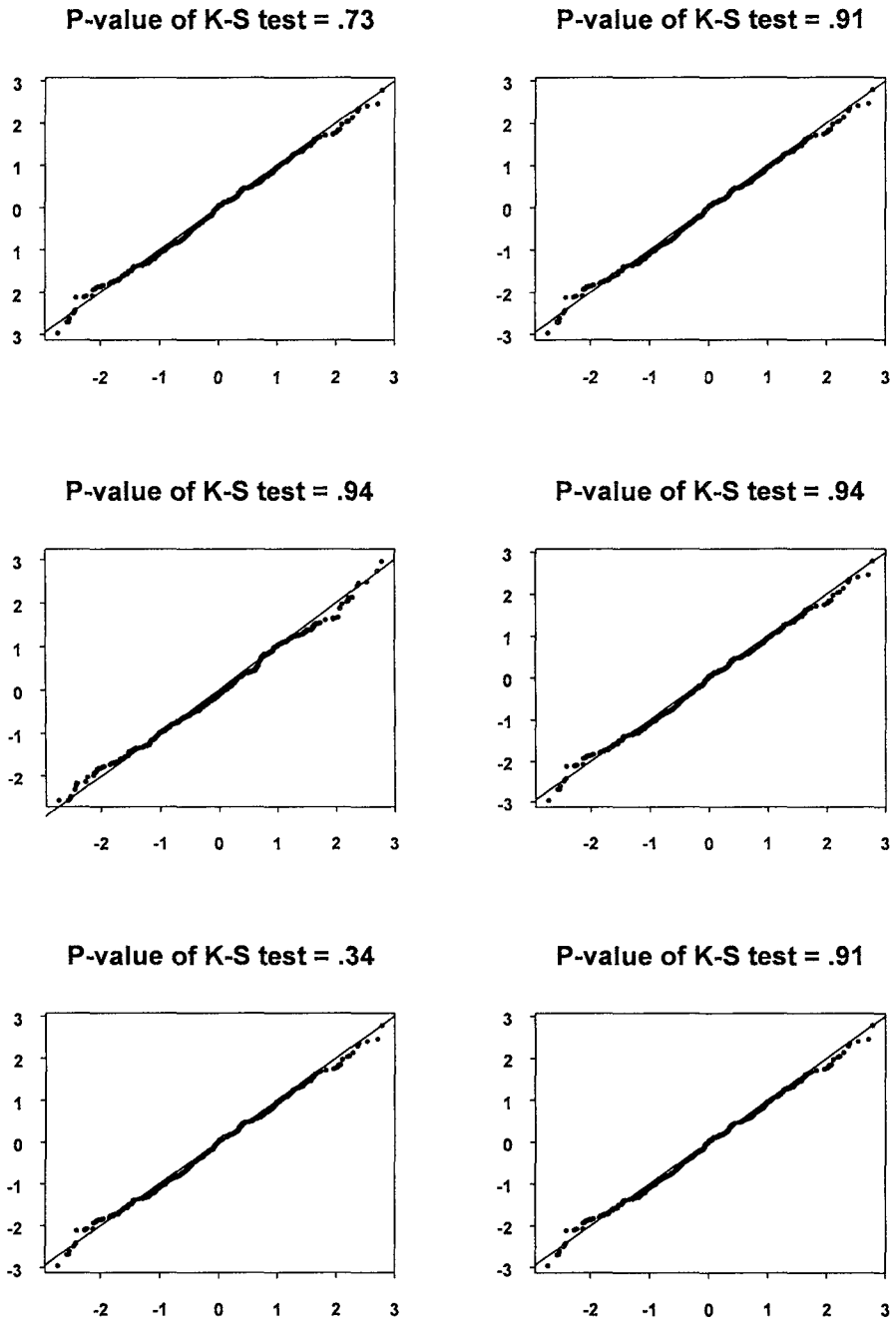

Figure 2. Q-Q plots and P-values of K-S tests for $n = 20$