
Kalimuthu Krishnamoorthy & Yanping Xia


To link to this article: https://doi.org/10.1080/00031305.2020.1829048
LETTER TO THE EDITOR


Let \( X_1, \ldots, X_n \) be a sample from a bivariate normal distribution with mean vector \( \mu \) and the covariance matrix \( \Sigma, N_2(\mu, \Sigma) \). Define

\[
\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{and} \quad S = \sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})'.
\]

Hu, Jung, and Qin (2020) have considered the problem of interval estimating the correlation coefficient \( \rho = \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}} \), where \( \sigma_{ij} \) is the \((i,j)\)th element of \( \Sigma \). In Section 3.1 of their article, the authors have developed the generalized pivotal quantity (GPQ) using the generalized variable approach. To describe their GPQ, let \( S_{ij} \) denote the \((i,j)\)th element of \( S \), and let \( s_{ij} \) be an observed value of \( S_{ij} \). Let \( s_{21} = s_{22} - s_{12}/s_{11} \). Furthermore, let \( Z \) be a standard normal random variable and \( U_i \) be a \( \chi^2_{n-1} \) random variable, \( i = 1, 2 \). Assume that these random variables \( Z, U_1 \), and \( U_2 \) are independent. In terms of these random variables and observed values \( s_{ij} \), the GPQ given in (3.9) of Hu, Jung, and Qin (2020) can be expressed as

\[
G_{\rho} = \frac{(s_{21}/s_{11} \sqrt{s_{21}/(s_{11} U_2)}) \sqrt{s_{11}/U_1}}{\left[(s_{21}/s_{11} - Z\sqrt{s_{21}/(s_{11} U_2)})^2 s_{11}/U_1 + s_{21}/U_2 \right]^{1/2}}.
\]

(1)

For a given \((s_{11}, s_{21}, s_{22})\), the 100\(\alpha/2\) percentile and the 100\((1-\alpha/2)\) percentile of \(G_{\rho}\) form a 100\((1-\alpha)\) confidence interval (CI) for \( \rho \). This CI is referred to as the generalized CI.

The above GPQ is not new and it has been already given in our article (Krishnamoorthy and Xia 2007). Indeed, we have obtained the GPQ for the correlation coefficient \( \rho \) as a special case from the GPQs of the elements of a \( p \times p \) normal covariance matrix \( \Sigma \). The GPQs for the elements of \( \Sigma \) were obtained from the GPQs for the elements \( \theta_{ij} \) of \( \theta \), where \( \theta \) is the Cholesky factor of \( \Sigma \). Using our GPQs, one can find CI for the difference between two overlapping correlation coefficients \( \rho_{ij} - \rho_{ik} \) and for the difference between two non-overlapping correlation coefficients \( \rho_{ij} - \rho_{kl}, i \neq j \neq k \neq l \). To write our GPQ for the simple correlation coefficient \( \rho \), let \( r = s_{21}/\sqrt{s_{11}s_{22}} \), the sample correlation coefficient, and let \( r^* = r/\sqrt{1-r^2} \). Then our GPQ (see Krishnamoorthy and Xia 2007, eq. (16)) can be expressed as

\[
Q_{\rho} = \frac{r^* \sqrt{U_2} - Z}{\sqrt{(r^* \sqrt{U_2} - Z)^2 + U_1}}.
\]

(2)

For a given \((s_{11}, s_{21}, s_{22})\), the 100\(\alpha/2\) and 100\((1-\alpha/2)\) percentiles of \(Q_{\rho}\) form a \((1-\alpha)\) confidence interval for \( \rho \).

We can show that the GPQs \( G_{\rho} \) and \( Q_{\rho} \) are the same as follows. We can write the numerator of (1) as

\[
\left(\frac{s_{21}}{s_{11}} - Z\sqrt{s_{21}/(s_{11} U_2)}\right) \sqrt{s_{11}/U_1} = \left(r^* \sqrt{U_2} - Z\right) \sqrt{s_{21}/U_1 U_2}.
\]

Substituting the above expression in the numerator and in the denominator of (1), it can be readily verified that \( G_{\rho} = Q_{\rho} \).

The simulation study by Krishnamoorthy and Xia (2007) indicated that the generalized CI for the correlation coefficient \( \rho \) is practically exact. Later, Krishnamoorthy (2013) has shown that the generalized CI is exact, and extended the results to the case of missing data.

References


Kalimuthu Krishnamoorthy
University of Louisiana at Lafayette
krisha@louisiana.edu

Yanping Xia
University of Nevada at Las Vegas
yanping.xia@unlv.edu