

Multivariate Approach to Partial Correlation Analysis

Onyeneke Casmir Chidiebere

Mathematics and Statistics Department, University of Calabar, Calabar, Nigeria

Email address:

casdyx@yahoo.com

To cite this article:

Onyeneke Casmir Chidiebere. Multivariate Approach to Partial Correlation Analysis. *Science Journal of Applied Mathematics and Statistics*. Vol. 3, No. 3, 2015, pp. 165-170. doi: 10.11648/j.sjams.20150303.20

Abstract: Multivariate approach to generate variance covariance and partial correlation coefficients of one or more independent variables has been the concern of advanced statisticians and users of statistical tools. This work tackled the problem by keeping one or some variables constant and partitioned the variance covariance matrices to find multivariate partial correlations. Due to the challenges that faced the analysis and computation of complex variables, this research used matrix to ascertain the level of relationship that exist among these variables and obtained correlation coefficients from variance covariance matrices. It was proved that partial correlation coefficients are diagonal matrices that are normally distributed. (Work count = 101).

Keywords: Multivariate, Correlation, Partial, Normality, Coefficients, Variables, Matrices

1. Introduction

Multivariate statistics is the science of collecting and analyzing multiple data. It involves observing, describing and presenting complex events and distributions that occur simultaneously. Multivariate partial correlation is the analysis made on many variables to ascertain the degrees of relationship existing when one or more of such variables are held constant (Francois, 2010). Here the main point is to examine how these multiple variance, covariance and correlation are partly related. This measures variance and correlation among the many variables. It examines the relationship between several variables of interest where the objective is to ascertain the interdependence effect existing between them. It checks how one variable relates with another. Neter, et al., (1996) suggested that cases with more than one independent variable are appropriate to compute with matrices. Therefore, we mostly adopted matrix principles to carry out the analysis in this work.

2. Multivariate Correlation

The statistical methodology used in this research is multivariate application to generate both variance covariance and correlation matrices. The correlation matrix was used to established and prove the theories of first and second order partial correlation. Since multivariate correlation employs matrix method to generate all the possible correlation confidents in the variables of consideration, we examined the cases of variance covariance matrix to enable us get the multivariate relationship between one index and another. It was used to measure the strength of association among variables. The impact of independent variables was studied simultaneously. For instance, the correlation coefficients among the standard of living indices in rural communities in a state was given as income (X_1), feeding (X_2), accommodation (X_3) and education (X_4), the multivariate correlation coefficients were computed as:

$$X = \begin{pmatrix} x_1^2 & x_1x_2 & x_1x_3 & x_1x_4 \\ x_2x_1 & x_2^2 & x_2x_3 & x_2x_4 \\ x_3x_1 & x_3x_2 & x_3^2 & x_3x_4 \\ x_4x_1 & x_4x_2 & x_4x_3 & x_4^2 \end{pmatrix}, \bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \end{bmatrix}$$

$$X = \begin{pmatrix} x_1^2 & x_1x_2 & x_1x_3 & x_1x_4 \\ x_2x_1 & x_2^2 & x_2x_3 & x_2x_4 \\ x_3x_1 & x_3x_2 & x_3^2 & x_3x_4 \\ x_4x_1 & x_4x_2 & x_4x_3 & x_4^2 \end{pmatrix} = \begin{pmatrix} 75251.14 & 14965.38 & 7979.07 & 9968.05 \\ 14965.38 & 3804.55 & 2065.95 & 2543.68 \\ 7979.07 & 2065.95 & 1146.70 & 1377.43 \\ 9968.05 & 2543.68 & 1377.43 & 1718.22 \end{pmatrix}$$

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \end{bmatrix} = \begin{bmatrix} 48.136 \\ 12.316 \\ 6.712 \\ 8.256 \end{bmatrix}$$

When N = 25

$$\begin{aligned} X - \bar{x}N\bar{x}' &= \begin{bmatrix} 75251.14 & 14965.38 & 7979.07 & 9968.05 \\ 14965.38 & 3804.55 & 2065.95 & 2543.68 \\ 7979.07 & 2065.95 & 1146.70 & 1377.43 \\ 9968.05 & 2543.68 & 1377.43 & 1718.22 \end{bmatrix} - \begin{bmatrix} 48.136 \\ 12.316 \\ 6.712 \\ 8.256 \end{bmatrix} 25(48.136 \quad 12.316 \quad 6.712 \quad 8.256) \\ &= \begin{bmatrix} 75251.14 & 14965.38 & 7979.07 & 9968.05 \\ 14965.38 & 3804.55 & 2065.95 & 2543.68 \\ 7979.07 & 2065.95 & 1146.70 & 1377.43 \\ 9968.05 & 2543.68 & 1377.43 & 1718.22 \end{bmatrix} - \begin{bmatrix} 57926.862 & 14821.07 & 8077.228 & 9935.2704 \\ 14821.07 & 3792.096 & 2066.625 & 2542.0224 \\ 8077.2208 & 2066.625 & 1126.2736 & 1385.3568 \\ 9935.27 & 2542.0224 & 1385.3568 & 1704.038 \end{bmatrix} \\ &= \begin{bmatrix} 17324.278 & 144.3056 & -98.1508 & 32.7796 \\ 144.3056 & 12.4536 & -6.748 & 1.6576 \\ -98.1508 & -67.48 & 20.4264 & -7.9268 \\ 32.7796 & 165.76 & -7.9268 & 14.1816 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} [(X - \bar{x}N\bar{x}')_{ii} (X - \bar{x}N\bar{x}')_{jj}]^{1/2} &= \begin{bmatrix} 17324.14 & 464.4886 & 594.8719 & 495.6672 \\ 464.4886 & 12.4536 & 15.9494 & 13.2895 \\ 594.8719 & 15.9494 & 20.4264 & 17.0210 \\ 495.6672 & 13.2895 & 17.0200 & 14.1816 \end{bmatrix} \frac{X - \bar{x}N\bar{x}'}{[(X - \bar{x}N\bar{x}')_{ii} (X - \bar{x}N\bar{x}')_{jj}]^{1/2}} \\ &= \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & \rho_{22} & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & \rho_{33} & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & \rho_{44} \end{bmatrix} = \begin{bmatrix} 1 & 0.3107 & -0.1650 & 0.0661 \\ 0.3107 & 1 & -0.0423 & 0.1247 \\ -0.1650 & -0.0423 & 1 & -0.4657 \\ 0.0661 & 0.1247 & -0.4657 & 1 \end{bmatrix} \end{aligned}$$

These coefficients take values between -1 and +1, where $0 \leq |\rho_{ij}| \leq 1$ (Rao, 2008)

Anderson (2003), stated the first partial correlation as

$$\rho_{ij.k} = \frac{\rho_{ij} - \rho_{ik}\rho_{jk}}{[(1 - \rho_{ik}^2)(1 - \rho_{jk}^2)]^{1/2}}$$

3. First Partial Correlation

According to Morison (2007), to calculate the partial relationship between the indices, we examined the cases where only one variable is kept constant and others were varied. In situations where the system has multiple variables and factors influencing them, multivariate partial correlation analysis becomes very relevant. This is seen in physical, agricultural, business and experimental sciences where the aim is how to study separately the effect and control of these variables. The approach is very significant during design or analysis of experiment and principle component analysis.

$i, j, k=1, \dots, 4$

Therefore, the first partial correlation matrix can be calculated by first computing the numerator as follows

$$\rho_{ij} - \rho_{ik}\rho_{jk} = \begin{bmatrix} 1 & 0.0090 & 0.1042 & -0.1519 \\ 0.0274 & 1 & -0.4548 & 0.3037 \\ -0.0107 & 0.3024 & 1 & -0.1342 \\ -0.4605 & 0.1050 & 0.0158 & 1 \end{bmatrix}$$

$$[(1 - \rho_{ik}^2)(1 - \rho_{jk}^2)]^{1/2} = \begin{bmatrix} 1 & 0.9375 & 0.9484 & 0.9497 \\ 0.9431 & 1 & 0.9841 & 0.9854 \\ 0.8728 & 0.9900 & 1 & 0.8830 \\ 0.9911 & 0.8841 & 0.8780 & 1 \end{bmatrix}$$

$$\frac{\rho_{ij} - \rho_{ik}\rho_{jk}}{[(1 - \rho_{ik}^2)(1 - \rho_{jk}^2)]^{1/2}} = \rho_{ij.k} = \begin{bmatrix} 1 & \rho_{23.1} & \rho_{24.1} & \rho_{13.2} \\ \rho_{14.2} & 1 & \rho_{34.1} & \rho_{12.3} \\ \rho_{14.3} & \rho_{12.4} & 1 & \rho_{13.4} \\ \rho_{34.2} & \rho_{24.3} & \rho_{23.4} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.095 & 0.1098 & -0.1599 \\ 0.0290 & 1 & -0.4622 & 0.3082 \\ -0.0122 & 0.3055 & 1 & -0.1520 \\ -0.4645 & 0.1188 & 0.0180 & 1 \end{bmatrix}$$

Theorem one: Given that $X = x_1, x_2, \dots, x_n$ be variables with variance covariance matrix $V = v_{ij}$, X, V and \bar{x} can be partitioned to derive multivariate partial correlation coefficient matrices. We are required to prove that

Let

$$\chi = (X_2 X_3 / X_4) \text{ and } \bar{\chi} = (\bar{x}_2 \quad \bar{x}_3 / \bar{x}_4), V = \begin{pmatrix} V_{22} & V_{23} \\ V_{32} & V_{33} \end{pmatrix}$$

$$\rho_{23.4} = \frac{\rho_{23} - \rho_{24}\rho_{34}}{[(1 - \rho_{24}^2)(1 - \rho_{34}^2)]^{1/2}}$$

$$V = \begin{pmatrix} V_{22} & V_{23} \\ V_{32} & V_{33} \end{pmatrix} = \begin{bmatrix} \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix} = \begin{bmatrix} \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix}$$

$$V_{22.3} = V_{22} - V_{23}V_{33}^{-1}V_{32} = \begin{bmatrix} \sigma_2^2 & \sigma_{23} \\ \sigma_{32} & \sigma_3^2 \end{bmatrix} - 1/\sigma_4^2 \begin{bmatrix} \sigma_{24} \\ \sigma_{34} \end{bmatrix} \begin{bmatrix} \sigma_{42} & \sigma_{43} \end{bmatrix} = \frac{1}{\sigma_4^2} \left\{ \left[\begin{pmatrix} \sigma_2^2 \sigma_4^2 & \sigma_{23} \sigma_4^2 \\ \sigma_{32} \sigma_4^2 & \sigma_3^2 \sigma_4^2 \end{pmatrix} \right] - \left[\begin{pmatrix} \sigma_{24}^2 & \sigma_{24} \sigma_{43} \\ \sigma_{34} \sigma_{42} & \sigma_{34}^2 \end{pmatrix} \right] \right\}$$

$$= \frac{1}{\sigma_4^2} \begin{bmatrix} \sigma_2^2 \sigma_4^2 - \sigma_{24}^2 & \sigma_{23} \sigma_4^2 - \sigma_{24} \sigma_{43} \\ \sigma_{32} \sigma_4^2 - \sigma_{34} \sigma_{42} & \sigma_3^2 \sigma_4^2 - \sigma_{34}^2 \end{bmatrix}$$

$$\rho_{23.4} = \frac{[\sigma_{23} \sigma_4^2 - \sigma_{24} \sigma_{43}] / \sigma_4^2}{\{[(\sigma_2^2 \sigma_4^2 - \sigma_{24}^2) / \sigma_4^2][(\sigma_3^2 \sigma_4^2 - \sigma_{34}^2) / \sigma_4^2]\}^{1/2}}$$

Divide through by $\sigma_{22}\sigma_{33}$

$$\rho_{23.4} = \frac{\frac{\sigma_{23}}{\sigma_{22}\sigma_{33}} - \left(\frac{\sigma_{24}}{\sigma_{22}\sigma_{44}}\right)\left(\frac{\sigma_{43}}{\sigma_{33}\sigma_{44}}\right)}{\left[\left(\frac{\sigma_2^2}{(\sigma_{22}\sigma_{33})} - \frac{\sigma_{24}^2}{\sigma_{22}\sigma_{33}\sigma_4^2}\right)\left(\frac{\sigma_3^2}{\sigma_{22}\sigma_{33}} - \frac{\sigma_{34}^2}{\sigma_{22}\sigma_{33}\sigma_4^2}\right)\right]^{1/2}} = \frac{\rho_{23} - \rho_{24}\rho_{34}}{\left[\left(\frac{\sigma_{22}}{\sigma_{33}} - \rho_{24}\frac{\sigma_{24}}{\sigma_{33}\sigma_{44}}\right)\left(\frac{\sigma_{33}}{\sigma_{22}} - \rho_{34}\frac{\sigma_{34}}{\sigma_{33}\sigma_{44}}\right)\right]^{1/2}}$$

$$= \frac{\rho_{23} - \rho_{24}\rho_{34}}{(1 - \rho_{24}^2 - \rho_{34}^2 + \rho_{24}^2\rho_{34}^2)^{1/2}} = \frac{\rho_{23} - \rho_{24}\rho_{34}}{[(1 - \rho_{24}^2)(1 - \rho_{34}^2)]^{1/2}}$$

This has proved that

$$\rho_{23.4} = \frac{\rho_{23} - \rho_{24}\rho_{34}}{[(1 - \rho_{24}^2)(1 - \rho_{34}^2)]^{1/2}}$$

Let us now apply the theorem in the numbers to get the partial relationship.
Since

$$V = \begin{pmatrix} V_{22} & V_{23} \\ V_{32} & V_{33} \end{pmatrix} = \begin{bmatrix} \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix} = \begin{bmatrix} 1245.36 & & -67.48 & & 165.76 \\ -67.48 & v_{22} & 2042.64 & & v_{23} & -792.68 \\ 165.76 & & v_{32} - 792.68 & & v_{33} & 1418.16 \end{bmatrix}$$

$$V_{22.3} = V_{22} - V_{23}V_{33}^{-1}V_{32} = \begin{bmatrix} 1245.36 & -67.48 \\ -67.48 & 2042.64 \end{bmatrix} - 1/1418.16 \begin{bmatrix} 165.76 \\ -792.68 \end{bmatrix} \begin{bmatrix} 165.76 & -792.68 \end{bmatrix}$$

$$= \frac{1}{1418.16} \left\{ \left[\begin{pmatrix} 1245.36(1418.16) & -67.48(1418.16) \\ -67.48(1418.16) & 2042.64(1418.16) \end{pmatrix} \right] - \left[\begin{pmatrix} 165.76^2 & 165.76(-792.68) \\ 165.76(-792.68) & -792.68^2 \end{pmatrix} \right] \right\}$$

$$= \frac{1}{1418.16} \begin{bmatrix} 1245.36(1418.16) - 165.76^2 & -67.48(1418.16) - 165.76(-792.68) \\ -67.48(1418.16) - 165.76(-792.68) & 2042.64(1418.16) - (-792.68^2) \end{bmatrix}$$

So

$$\rho_{23.4} = \frac{[-67.48(1418.16) - 165.76(-792.68)] / 1418.16}{\{[(1245.36(1418.16) - 165.76^2) / 1418.16][2042.64(1418.16) - (-792.68^2) / 1418.16]\}^{1/2}}$$

Divide through by $\sigma_{22}\sigma_{33} = (35.2897)(45.1956)$

$$\rho_{23.4} = \frac{\frac{-67.48}{(35.29)(45.20)} - \left(\frac{165.76}{(35.29)(37.66)}\right)\left(\frac{-792.68}{(45.20)(37.66)}\right)}{\left\{ \left[\frac{(35.29)^2}{(35.29)(45.20)} - \frac{(165.76)^2}{(35.29)(45.20)(37.66)^2} \right] \left[\frac{(45.20)^2}{(35.29)(45.20)} - \frac{(-792.68)^2}{(35.29)(45.20)(37.66)^2} \right] \right\}^{1/2}}$$

$$= \frac{-0.0423 - (0.1247)(-0.4657)}{\left[\left(\frac{35.29}{45.20} - (0.1247)\frac{165.79}{(45.20)(37.66)}\right)\left(\frac{45.20}{35.29} - (-0.4657)\frac{-792.68}{(45.20)(37.66)}\right) \right]^{1/2}}$$

$$= \frac{-0.0423 - (0.1247)(-0.4657)}{[1 - (0.1247)^2 - (-0.4657)^2 + (0.1247)^2(-0.4657)^2]^{1/2}} = \frac{-0.0423 - (0.1247)(-0.4657)}{\{[1 - (-0.1247)^2][1 - (-0.4667)^2]\}^{1/2}}$$

4. Second Partial Correlation

This is to compute the partial relationship between two variables when two other indices are held constant. Theorem two: At this time, we prove that;

$$\rho_{ij.km} = \frac{\rho_{ij.k} - \rho_{im.k}\rho_{jm.k}}{[(1 - \rho_{im.k}^2)(1 - \rho_{jm.k}^2)]^{1/2}} = \frac{\rho_{ij.m} - \rho_{ik.m}\rho_{jk.m}}{[(1 - \rho_{ik.m}^2)(1 - \rho_{jk.m}^2)]^{1/2}} = \frac{A^n}{B^n}$$

The above expression means $A^n = \rho_{ij.k} - \rho_{im.k}\rho_{jm.k} = -A^n + \rho_{ij.k} = \rho_{im.k}\rho_{jm.k}$. So given X_1, X_2, \dots, X_n with the corresponding $\rho_{i1}, \rho_{i2}, \dots, \rho_{in}$, the variance covariance of

$$(-A^n + \rho_{ij.k})^2 = (\rho_{im.k})^2(\rho_{jm.k})^2$$

$$(\rho_{im.k})^2(\rho_{jm.k})^2 = \left\{ \left[\begin{pmatrix} \rho_{im} \\ \rho_{ik} \\ \rho_{mk} \end{pmatrix} (\rho_{im} \ \rho_{ik} \ \rho_{mk}) \right] \left[\begin{pmatrix} \rho_{jm} \\ \rho_{jk} \\ \rho_{mk} \end{pmatrix} (\rho_{jm} \ \rho_{jk} \ \rho_{mk}) \right] \right\}$$

$$= \begin{pmatrix} 1 & \rho_{im}\rho_{ik} & \rho_{im}\rho_{mk} \\ \rho_{im}\rho_{ik} & 1 & \rho_{ik}\rho_{mk} \\ \rho_{im}\rho_{mk} & \rho_{im}\rho_{mk} & 1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{jm}\rho_{jk} & \rho_{jm}\rho_{mk} \\ \rho_{im}\rho_{jk} & 1 & \rho_{jk}\rho_{mk} \\ \rho_{jm}\rho_{mk} & \rho_{jk}\rho_{mk} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \rho_{mk.i} & \rho_{ik.m} \\ \rho_{mk.i} & 1 & \rho_{im.k} \\ \rho_{ik.m} & \rho_{im.k} & 1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{mk.j} & \rho_{jk.m} \\ \rho_{mk.j} & 1 & \rho_{jm.k} \\ \rho_{jk.m} & \rho_{jm.k} & 1 \end{pmatrix}$$

Multiplying by each value with the corresponding terms

$$= \begin{pmatrix} 1 & \rho_{mk.i}\rho_{mk.j} & \rho_{ik.m}\rho_{jk.m} \\ \rho_{mk.i}\rho_{mk.j} & 1 & \rho_{im.k}\rho_{jm.k} \\ \rho_{ik.m}\rho_{jk.m} & \rho_{im.k}\rho_{jm.k} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \rho_{mk.ij} & \rho_{ij.km} \\ \rho_{mk.ij} & 1 & \rho_{ij.km} \\ \rho_{ij.m} & \rho_{ji.mk} & 1 \end{pmatrix}$$

This implies

$$\rho_{ik.m}\rho_{jk.m} = \rho_{im.k}\rho_{jm.k} = \rho_{ij.km}$$

$$\rho_{ij.km} = \rho_{ij.k} - \rho_{im.k}\rho_{jm.k} = \rho_{ij.m} - \rho_{ik.m}\rho_{jk.m}$$

$$\rho_{ij.km} = \frac{\rho_{ij.k} - \rho_{im.k}\rho_{jm.k}}{[(1 - \rho_{im.k}^2)(1 - \rho_{jm.k}^2)]^{1/2}} = \frac{\rho_{ij.m} - \rho_{ik.m}\rho_{jk.m}}{[(1 - \rho_{ik.m}^2)(1 - \rho_{jk.m}^2)]^{1/2}}$$

Also applying the above prove to the standard of living indices we could achieve the following result. For instance,

$$A^1 = \rho_{12.4} - \rho_{13.4}\rho_{23.4} = -0.4632$$

$$A^2 = \rho_{12.4} - \rho_{13.4}\rho_{23.4} = -0.4632$$

$$B^1 = [(1 - \rho_{13.4}^2)(1 - \rho_{23.4}^2)]^{1/2} = 0.9939$$

$$B^2 = [(1 - \rho_{14.3}^2)(1 - \rho_{24.3}^2)]^{1/2} = 0.9867$$

$$\frac{A^1}{B^1} = \frac{\rho_{12.4} - \rho_{13.4}\rho_{23.4}}{[(1 - \rho_{13.4}^2)(1 - \rho_{23.4}^2)]^{1/2}} = \frac{-0.4632}{0.9939} = -0.4661$$

$$\frac{A^2}{B^2} = \frac{\rho_{12.4} - \rho_{13.4}\rho_{23.4}}{[(1 - \rho_{14.3}^2)(1 - \rho_{24.3}^2)]^{1/2}} = \frac{-0.4599}{0.9867} = -0.4661$$

$$\rho_{ij.km} = \frac{\rho_{ij.k} - \rho_{im.k}\rho_{jm.k}}{[(1 - \rho_{im.k}^2)(1 - \rho_{jm.k}^2)]^{1/2}} = \frac{\rho_{ij.m} - \rho_{ik.m}\rho_{jk.m}}{[(1 - \rho_{ik.m}^2)(1 - \rho_{jk.m}^2)]^{1/2}} = \frac{A^n}{B^n}$$

Other results are seen in the second order partial correlation matrix of the correlation matrices as presented below.

Notice that n in this case is from 1 to 16

$$A^n = \rho_{ij.k} - \rho_{im.k}\rho_{jm.k} = A = \begin{bmatrix} 1 & -0.4632 & 0.1143 & 0.0603 \\ -0.4599 & 1 & -0.0452 & -0.1464 \\ 0.1226 & -0.0489 & 1 & 0.3096 \\ 0.0644 & -0.1575 & 0.3082 & 1 \end{bmatrix}$$

$$B^n = [(1 - \rho_{im.k}^2)(1 - \rho_{jm.k}^2)]^{1/2} = B = \begin{bmatrix} 1 & 0.9939 & 0.8868 & 0.8814 \\ 0.9867 & 1 & 0.8742 & 0.8852 \\ 0.9513 & 0.9446 & 1 & 0.9928 \\ 0.9411 & 0.9520 & 0.9882 & 1 \end{bmatrix}$$

$$\frac{A^n}{B^n} = \rho_{ij.km} = \begin{bmatrix} 1 & \rho_{34.12} & \rho_{24.13} & \rho_{23.14} \\ \rho_{34.12} & 1 & \rho_{14.23} & \rho_{13.24} \\ \rho_{24.13} & \rho_{14.23} & 1 & \rho_{12.34} \\ \rho_{23.14} & \rho_{13.24} & \rho_{12.34} & 1 \end{bmatrix} = \begin{bmatrix} 1 & -0.4661 & 0.1289 & 0.0684 \\ -0.4661 & 1 & -0.0518 & -0.1654 \\ 0.1289 & -0.0518 & 1 & 0.3119 \\ 0.0684 & -0.1654 & 0.3119 & 1 \end{bmatrix}$$

5. Normality Test of Multivariate Correlation Coefficients

In this part, we prove that if the variables are normally distributed, then the relationships between them in terms of correlation coefficients are also normal distributions. Here since the sample size for each index (variable) is 25, the four variables gives 100. It was already assumed to be normal by the law of large number, now let's test for the normality of the correlation coefficients.

Theorem three: If $X \sim N_p(u, v)$, then $\rho_{ij} \sim N_p(v_\rho, \sigma_\rho^2)$. In this test we recall that if $X \sim N(u, \sigma^2)$ with

$$f(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp \frac{-1}{2\sigma^2}(x-\mu)^2$$

Let $\rho = (\rho_{i1}, \rho_{i2}, \dots, \rho_{ij})$ be the correlation coefficients of $X = (X_{i1}, X_{i2}, \dots, X_{ij})$ then the joint probability density function is given by

$$f(\rho, \mu_\rho, \sigma_\rho^2) = \prod_{i=1}^n f(\rho_i, \mu_\rho, \sigma_\rho^2) = \frac{1}{(2\pi)^{\frac{n}{2}\sigma_\rho^2}} \exp \frac{-\frac{1}{2}\sum(\frac{\rho_{ij}-\mu_\rho}{\sigma_\rho})^2}{\sigma_\rho^2}$$

From the above, we can get our test statistic to be $Z_i = (\rho_{ij} - \mu_\rho)/\sigma_\rho$

Table 1. Test for normality of the Correlation Coefficients.

ρ_{ij}	F	$\rho_{ij} - \mu_\rho$	$(\rho_{ij} - \mu_\rho)^2$	$(\rho_{ij} - \mu_\rho)^2 / \sigma_\rho$	$p(\rho_{ij} - \mu_\rho) / \sigma_\rho$	$n p_i(z_i)$	$(1 - e_i)$	$(1 - e_i)^2$	$(1 - e_i)^2 / e_i$
-0.4657	1	-0.4375	0.1911	-1.16349	0.0516	0.3096	0.6904	0.4767	1.5396
-0.1650	1	-0.1364	0.0189	-1.5101	0.0253	1.5204	-0.5204	0.2708	0.1781
-0.0420	1	-0.0137	0.0002	-0.0513	0.1751	1.0506	-0.0506	0.0026	0.0024
0.0661	1	0.0947	0.0090	0.3542	0.1567	0.9402	0.0598	0.0036	0.0038
0.1247	1	0.1533	0.0235	0.5734	0.0789	0.4734	0.5266	0.2773	0.5858
0.3107	1	0.3393	0.1151	1.2688	0.2843	1.7058	-0.7058	0.4982	0.2620
-0.1715	6		0.3575		1	6			2.6018

Where ρ_{ij} = the corresponding correlation coefficient
 F = the frequency of the correlation coefficient which is one per outcome
 μ_ρ and σ_ρ are mean and standard deviation of the correlation coefficients respectively.

The above analysis is grounded under the following hypothesis

$$H_0: \mu_{\rho_1} = \mu_{\rho_2} = \dots = \mu_{\rho_i}; H_1: \mu_{\rho_1} \neq \mu_{\rho_2} \neq \dots \neq \mu_{\rho_i}$$

$$\sigma_\rho^2 = \frac{\sum(\rho_{ij} - \mu_\rho)^2}{n - 1} = \frac{0.3575}{5} = 0.0715$$

$$\sigma_\rho = \sqrt{0.0715} = 0.2674$$

$$\mu_\rho = \frac{\sum \rho_i}{n} = \frac{-0.0286}{6} = -0.0286$$

Since $X_{cal}^2 (2.6018) < X_{(0.950, 4)}^2$ tabulated (9.4880) we therefore accept H_0 and conclude that $\rho_{ij} \sim N_p(v_\rho, \sigma_\rho^2)$

6. Interpretation

The numerical matrix of correlation coefficients displayed all the relationships between any selected sets of two variables, covariance (X_{ij}), under consideration. If both variables increase at the same time, then positive correlation exists between them. This was seen in the case of $\rho_{12}, \rho_{14}, \rho_{24}$ and so on. However, the negative correlation coefficient exist in the cases where one variable tends to increase as the other tends to decrease as seen in $\rho_{13}, \rho_{23}, \rho_{43}$. The result of the first partial correlation coefficient measured the relationship between any two variables when any other

variables are held constant. At the second partial correlation coefficient matrix which examines the correlation at any two variables keeping the rest two variables stable, it was discovered that is $\rho_{ik.m}\rho_{jk.m} = \rho_{im.k}\rho_{jm.k} = \rho_{ij.km}$. This implies, $\rho_{34.12} = \rho_{34.12}$. It was confirmed that there exist perfect diagonal matrices among the coefficients of simple and the second partial correlations. Finally, Table 1 showcased what happened when the correlation coefficients were subjected to normality test. The essence was to prove that correlation coefficients obey the normal probability law of $\rho_{ij} \sim N_p(v_\rho, \sigma_\rho^2)$ when the sample variable is normal $X \sim N_p(u, v)$.

7. Conclusion

In this research, it was shown that multivariate technique can be used to easily compute the correlation coefficient matrices through variance covariance matrices (Kendall et al 1973). Theorem one proved that partial correlation coefficients can be derived by adequately partitioning the covariance matrix. Theorem two established the relationship which exists when two variables are kept constant. In such second partial correlation, it was verified that irrespective of the covariance or correlation coefficient terms used, second partial correlation (in a square matrices) must give a corresponding value to show the diagonal attributes of the correlation coefficients. Finally, the normality nature of multivariate correlation confidants was attested. In other

words, every normally distributed set of observation give correlation coefficient matrix which is equally normally distributed.

Theorems

1. $\rho_{12.3} = \frac{\rho_{12} - \rho_{13}\rho_{23}}{[(1 - \rho_{13}^2)(1 - \rho_{23}^2)]^{1/2}}$
2. $\rho_{12.34} = \frac{\rho_{12.3} - \rho_{14.3}\rho_{24.3}}{[(1 - \rho_{14.3}^2)(1 - \rho_{24.3}^2)]^{1/2}} = \frac{\rho_{12.4} - \rho_{13.4}\rho_{23.4}}{[(1 - \rho_{13.4}^2)(1 - \rho_{23.4}^2)]^{1/2}}$
3. If $X \sim N_p(u, v)$, then $\rho_{ij} \sim N_p(v_\rho, \sigma_\rho^2)$.

References

- [1] Anderson, T. W., (2003); An Introduction to Multivariate Statistical Analysis 3rd Edition. Wiley-Interscience, New York
- [2] Francois Husson, Sebastien Le and Jerome Pages (2010); Exploratory Multivariate Analysis. Chapman & Hall Books, retrieved from CRC Press www.crcpress.com
- [3] Kendall, M. G., and Stuart, A. (1973). The Advanced Theory of Statistics (Inference and Relationship). Vol. 2 (Third ed.). Hafner Publishing Co., New York.
- [4] Morison D. F., (2007); Multivariate statistical Methods. McGraw-Hill Book Company, New York. Retrieved from onlinelibrary.wiley.com.
- [5] Neter, John, Michael H. Kutner, Christopher Nachtsheim, and William Wasserman (1996). Applied Linear Statistical Models. McGraw Hill, Boston.
- [6] Partial Correlation Analysis retrieved from <https://explorable.com/partial-correlation-analysis>
- [7] Rao, C. R., (2008); Linear Statistical Inference and its Application: Second Edition. John Wiley & Sons Inc. Retrieved from onlinelibrary.wiley.com.
- [8] Regression Tutorial Menu Dictionary, STATS @ MTSU retrieved from <http://mtweb.mtsu.edu/stats/regression/level3/multicorrel/multicorrcoef.htm>
- [9] Steiger, J. H., and Browne, M. W. (1984). The Comparison of Interdependent Correlations between Optimal Linear composites. Psychometrika, 49, 11-24.
- [10] Stockwell Ian (2008); Introduction to Correlation and Regression. SAS Global Forum, Batimore.
- [11] Tabachnick, B., and Fidell L. (1989); Using Multivariate Statistics. Harper & Row Publishers, New York.