Department of Statistics

Faculty of Science MSU

Multivariate Analysis -1 (M.Sc. Sem 1)

Practical 3 - Wishart Distribution Date: 7/8/2025

Q 1. In a study of the market structure of firms, data were collected on four variables that determine market structure for a sample of N=231 large U.S. firms during 1960–1961.

The variables are:

- X_1: Market share at a given time (%)
- X_2: Firm size (natural logarithm of net total assets)
- X_3: Rate of return on capital (%)
- X_4: Growth (change in total revenue over a given period, %)

The sample covariance matrix is given as:

$$S = \begin{pmatrix} 26.626 & 2.233 & -0.019 & 0.260 \\ 2.233 & 1.639 & 0.156 & -0.080 \\ -0.019 & 0.156 & 1.513 & 0.222 \\ 0.260 & -0.080 & 0.222 & 1.346 \end{pmatrix}$$

Assume the data follows a multivariate normal distribution $N4(\mu, \Sigma)$. Verify the linear transformation property of the Wishart distribution, which states that if $A \sim Wp\ (\Sigma, n)$ and

B is a $q \times p$ of rank $q \leq p$, then $B\ A\ B^T \sim W_q\ (B\ \Sigma\ B^T, n)$ Specifically:

1. Use the sample covariance matrix to form $A=(N-1)S$, assuming it follows $W_4(\Sigma, N-1)$.
2. Choose a transformation matrix $B=[0.25,0.25,0.25,0.25]$ (representing an equally weighted average of the variables).
3. Compute BAB^T and verify that it follows $W_1(B\Sigma B^T, N-1)$ by checking the mean and distribution properties, approximating $\Sigma \approx S$.

Provide all calculations manually in matrix form and interpret the results in the context of the market structure study

Q 2) Using the Market Structure data, Verify the determinant property of the Wishart distribution, which states that if A~ Wp(Σ, n) ,then the ratio of the determinants $|A|/|\Sigma|$ is distributed as the product of p independent chi-squared random variables with degrees of freedom $n, n-1, \ldots, n-(p-1)$.

Q 3) In the study of financial performance of mid-sized technology forms data were collected on three variables for two independent samples of firm in 2024. X1 is annual revenue growth (in %), X2 is research and development expenditure, X3 is debit to equity ratio and N1 is 151 firms and N2 is 101 firms.

$$S_A = \begin{matrix} 15.750 & 1.20 & -0.150 \\ 1.200 & 2.500 & 0.300 \\ -0.150 & 0.300 & 1.800 \end{matrix} \qquad S_B = \begin{matrix} 16.200 & 1.300 & -0.100 \\ 1.300 & 2.600 & 0.250 \\ -0.100 & 0.250 & 1.900 \end{matrix}$$

Assume that the following MVN distribution with the same population variance covariance matrix Σ. Let $A=(N_1\text{-}1)\, S_A \sim W_3(\Sigma, m)$ and $B=(N_2\text{-}1)\, S_B \sim W_3(\Sigma, n)$

Compute the determinant ratio $|B|/|A|$. According to the Wishart distribution property, this ratio is proportional to the product of p independent F-distributed random variables. Specify the degrees of freedom for these F-distributions and simulate one realization of their product.

Q4) Given the sample covariance matrices S_A and S_B for the four health and fitness variables, compute $\phi=|A^{-1}B|$ | using the Wishart property $|A^{-1}B|=|B|/|A|$, where $A=230S_A$ and $B=179S_B$. Interpret the result in the context of comparing the covariance structures of fitness metrics between 2023 and 2024. Discuss how this can help the health science club assess the consistency of the wellness program's impact on student fitness.

Group A sample variance covariance matrix

$$S_A = \begin{bmatrix} 64.500 & -2.300 & 1.200 & 3.150 \\ -2.300 & 4.800 & -.250 & 0.400 \\ 1.200 & -.250 & 6.750 & .320 \\ 3.150 & .400 & 0.320 & 8.200 \end{bmatrix}$$

Group B sample variance covariance matrix

$$S_B = \begin{bmatrix} 62.400 & -2.100 & 1.1 & 2.900 \\ -2.100 & 4.600 & -0.230 & 0.380 \\ 1.100 & -0.230 & 6.500 & 0.300 \\ 2.900 & 0.380 & 0.300 & 7.900 \end{bmatrix}$$

Q 5) Let $A \sim W_p(I,m)$ and $B \sim W_p(I,n)$ be independent Wishart-distributed matrices with p=3 m=8, and n=4, where I is the 3×3 identity matrix. Compute the Wilks' Lambda statistic $\Lambda = |A|/|A+B| = |I+A^{-1}B|^{-1}$ for a given pair of matrices A and B, and verify that it follows the Wilks' Lambda distribution $\Lambda(3,8,4)$. Use this to perform a hypothesis test for equality of means in a multivariate setting, interpreting the result.

$$A = \begin{bmatrix} 20 & 3 & 1 \\ 3 & 15 & 2 \\ 1 & 2 & 18 \end{bmatrix} \qquad B = \begin{bmatrix} 5 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$

Suppose we are testing whether two groups with p=3 variables have equal mean vectors in a MANOVA setting.

A is the within-group sum of squares and cross-products (SSCP) matrix, $A \sim W_3(I, m=8)$ (e.g., from a group with 9 observations, $m = n_1 - 1 = 8$.

B is the between-group SSCP matrixrum, $B \sim W_3(I, n=4)$ (e.g., from 5 groups, $n = k - 1 = 4$.

****