UNIVERSITY OF THE PUNJAB

Roll No. ....

M.A./M.Sc. Part – II Annual Examination – 2020

Subject: Statistics | Paper: I (Statistical Inference) | Time: 3 Hrs. | Marks: 100

NOTE: Attempt any FOUR questions. All questions carry equal marks.

Q.1 a) How would you decide that a estimator is unbiased? Write a comprehensive note by mentioning all the methods of unbiasedness in detail. (08)

b) Let $\underline{X}$ be a random sample of size n from Bernoulli distribution, and let $T=\sum_{i=1}^{n} X_i$, then show that i)T/n is unbiased estimator of θ ii)$[T(T-1)]/[n(n-1)]$ is unbiased estimator of θ^2 (09)

c) Let y_1 and y_2 be two stochastically independent unbiased estimators of θ and if $V(y_1)$ is twice the $V(y_2)$, where V stands for variance, then find the constants k_1 and k_2 so that $k_1y_1+k_2y_2$ is an unbiased estimator of θ with smallest possible variance. (08)

Q.2 a) What do you mean by regularity conditions? What is their role in efficiency? (2+4)

b) Let 'L' denote the likelihood function and 'ln' denote the natural log then prove that $E[(\partial \ln L)/(\partial\theta)] = 0$ and also prove that $E[(\partial \ln L)/\partial\theta]^2 = -E[(\partial^2 \ln L)/\partial\theta^2]$ for θbeing parameter. (11)

c) Discuss how to attain the equality (=) in Cramer- Rao's minimum variance bound. (08)

Q.3 a) Find the MLEs of the parameters of two parameter gamma distribution, also find the variance co-variance matrix of the parameters. (13)

b) Let $Y_1<Y_2<Y_3\ldots<Y_n$ be the order statistic of a random sample of size `n` from the Uniform distribution of continuous type over the closed interval $[\theta-p, \theta+p]$. Find the MLE of$\theta and p$. (12)

Q.4 a) Find the most general form of the distribution for which Harmonic mean is the MLE of its parameter. (10)

b) Find the moment estimator of the parameter `p` of the binomial distribution if the parameter 'n` is known. (09)

c) Compare the properties of MLE and moment estimators. (06)

Q.5 a) Compare the minimum chi-square estimator and MLE. (05)

b) Find the minimum chi-square estimator of $\theta$from Binomial distribution. (09)

c) If X is a poisson variate with parameter λ and if $g(\lambda) = \frac{1}{m!}\left(\frac{m+1}{\lambda_0}\right)^{m+1} \lambda^m e^{\frac{-(m+1)\lambda}{\lambda_0}}$ is prior density for unknown parameter λ, given a random sample of size n, obtain the Baye's estimator for λ (11)

Q.6 a) In Sequential Probability Ratio Test (SPRT), find the approximate values of K_1 and K_2 (the limits) (10)

b) Let $\alpha \sim N(\theta, 100)$. Perform SPRT to test the hypothesis $H_o{:}\theta = 75, H_1{:}\theta = 78$ such that α and β are approximately equal to 0.1. (10)

c) Explain why do we need to use sequential sampling? (05)

Q.7 a) Let $x_1, x_2, \ldots\ldots, x_n$ denote a random sample from a distribution which has p.d.f. $f(x_i)$ that is positive on only non-negative integers. It is desired to test the simple hypothesis $H_o{:} f(x) = e^{-1}/x!, x = 0,1,2,\ldots\ldots$ against alternative simple hypothesis $H_1{:} f(x) = (1/2)^{x+1} x = 0,1,2,\ldots\ldots$ Derive the expression for BCR (Best critical region). Consider the case of n=1 and k=1, k being any positive integer in the expression $(L(\theta', x_1, x_2, \ldots\ldots, x_n)/L(\theta'', x_1, x_2, \ldots\ldots, x_n)) \leq k$ where $H_o{:}\theta = \theta', H_1 = \theta = \theta''$. Find the power of the test for this combination of n and k when (13)

(i) H_o is true(ii) H_1 is true

b) Explain the statistical method for the construction of confidence intervals. Also differentiate between confidence intervals and confidence region. (8+4)

UNIVERSITY OF THE PUNJAB

M.A./M.Sc. Part – II Annual Examination – 2020

Roll No. ....

Subject: Statistics   Paper: II (Regression Analysis and Econometrics)   Time: 3 Hrs.   Marks: 100

NOTE: Attempt any FOUR questions. All questions carry equal marks.

Q.1.a) Define the type of Econometrics and discuss its methodology.

b) Consider Simple linear regression $Y = \alpha + \beta X + \epsilon$. Such that $\epsilon' s$ are distributed Independently & Identically with zero mean and σ^2 variance. Derive unbiased estimator of σ^2.

Q.2.a) Consider the model $Y = \beta_1 + \beta_2 X_2 + \ldots\ldots + \beta_K X_K + \epsilon$. Which fulfills the OLS assumptions. Develop the testing procedure for $H_0: \beta_1 = \beta_{10}, \beta_2 = \beta_{20}, \ldots\ldots = \beta_K = \beta_{K0}$

b) Given the following calculations deviated from means for linear regression Y on X_1 and X_2 $n = 10, \Sigma y^2 = 48.2, \Sigma x_1^2 = 2\ \Sigma x_2^2 = 3,\ \Sigma x_1 y = -1,\quad \Sigma x_2 y = 8,\quad \Sigma x_1 x_2 = -1$ What is the least square estimate of β_1 and its standard error after imposing the restriction that $\beta_1 + \beta_2 = 4$

Q.3.a) Differentiate between Collinearity and Multicollinearity. When $X'X$ matrix becomes correlation matrix? Discuss the role of correlation matrix in Multicollinearity.

b) Following are the values of a variable and its estimated values obtained by estimating a linear regression.

Y:	2.81	2.88	2.90	3.07	3.16	3.22	3.38	3.53	3.74	3.98
$\hat{Y}$:	2.61	2.76	2.92	3.03	3.22	3.37	3.53	3.68	3.83	3.85
Y:	4.18	4.30	4.52	4.69	4.77	4.83	4.89	4.94		
$\hat{Y}$:	4.13	4.27	4.55	4.59	4.74	4.90	4.97	4.99		

Test the autocorrelation by Geary test. Estimate first order autocorrelation.

Q.4.a) Discuss the assumption and procedure of Goldfeld – Quandt test for heteroskedasticity.

b) For partitioned G.L. regression model, estimate stepwise regression estimates of the parameters and compare with full regression estimates.

Q.5.a) Define the rules / Conditions of identification.

b) Consider the mode; $y_1 = \alpha_1 y_2 + \alpha_2 x_1 + \alpha_3 x_2 + u_1, y_2 = \beta_1 y_1 + \beta_2 x_3 + u_2$ .Obtain consistent estimates of structural parameters, where possible, by appropriate method(s).

Q.6.a) Differentiate between distributed lag model and autoregressive model. Describe the Koyck distributed lag model. What problem arises in the estimation of this model?

b) Define orthogonal polynomials and discuss their use in polynomial regression analysis.

Q.7. Discuss the following:

i) Income Distribution Analysis.
ii) Instrumental Variables.
iii) Ridge Regression.
iv) Consequences of Autocorrelation.

UNIVERSITY OF THE PUNJAB

M.A./M.Sc. Part – II Annual Examination – 2020

Roll No. ....

Time: 3 Hrs. Marks: 75

Subject: Statistics

Paper: III (Part-A) (Data Processing and Computer Programming)

NOTE: Attempt any FOUR questions.

Q.1. (a) Describe the functions of the following components of a digital computer:
(i) Main Memory (ii) Low-level and High-level Languages
(iii) Input and Output Devices

(b) Describe the usage of following functions of
(i) dir (ii) rename (iii) copy (iv) del

(c) Write algorithm and flow chart to print out the even numbers between 1 to 500 and their squares. **(6+8+5)**

Q.2. (a) Explain the FORTRAN worksheet. Which command is used to clear the contents from output screen.?

(b) Write the following mathematical expression into FORTRAN expressions.

(i) $e^{|x-y|} - \frac{2}{3}(x+y)^{3/4}$ (ii) $\frac{e^{\frac{1}{2}(x^2 - 2\sqrt[3]{yz})}}{x+y+z}$

(iii) $a^x + \frac{b}{c}x$ (iv) $tanx - \sqrt{|cos(a-nb)|}$

(v) $\frac{1}{2\sqrt{\pi}}\frac{x^5}{5!}$

(c) Suppose J=5 and K=10. Find the value of J if each of the following is executed.

(i)
```
IF (J.GT.0) then
      K = K + 5
      J = J * K
ENDIF
J = J * K
J = 2 * J
```

(ii)
```
IF (J.LE.10) then
      K=K+5
      J=J*K
ELSE
      J=J*K
ENDIF
J=2*J
```

(d) Write a FORTRAN programme that generates and finds the sum of series.

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots \ldots\ldots\ldots\ldots\ldots - \frac{1}{50}$$ **(4+5+4+6)**

Q.3. (a) Write down the general form of the arithmetic IF and Computed GOTO statements.

(b) Write a FORTRAN program which reads the amount and print the income tax for the individuals according to following slab rates.

Upto Rs. 50000	Nil
Between Rs. 50000 & Rs. 100000	10% of the excess over Rs. 50000
Between Rs. 100000 & Rs. 200000	Rs. 5000+25% of the excess over Rs. 100000
Between Rs. 200000 & Rs. 300000	Rs. 30000+30% of the excess over Rs. 200000
Above Rs. 300000	Rs. 60000+40% of the excess over Rs. 300000

(c) Write a FORTRAN program, which compute and prints multiplication tables for n (say between 2 and 10) **(4+9+6)**

Q.4. (a) Define the functions of the following FORTRAN statements. Give at least two examples in each case.
(i) TYPE STATEMENT
(ii) DATA STATEMENT
(iii) ARITHMETIC STATEMENT

(b) Write a FORTRAN program using arrays to find the correlation coefficient, Simple Linear Regression line between N pairs of real numbers.

(c) Write a programme to find the sum of elements above, below and of diagonal elements of a matrix. **(6+6+7)**

Q.5. (a) Define Loop? What are the advantages of using Loop?

(b) Write a FTN program to read in an integer n > 2 and determine if n is a prime number.

(c) Write a FTN program to find the roots of the quadratic equation. **(5+6+8)**

Q.6. (a) Explain with examples the format and function of the following C statements.
(i) for statement() (ii) do-while statement

(b) Write the following expressions into C++ language:

(i) $\frac{-b+\sqrt{b^2-4ac}}{2a}$ (ii) $\frac{x^2\log_{10}(x)}{\sqrt{a^2-b^2}}$

(c) Write a C++ program that generates Fibonacci numbers using variable length array.

(d) Write and run a C++ program using 'switch() statement' to make a calculator that reads in two operands, an operator and displays the computed result. **(4+4+6+5)**

Q.7. (a) Define variable. Describe basic data types in C++ language.

(b) Write a program in C which read all elements of two matrices A and B of any order, find the product of these two matrices and prints their product.

(c) Write and run a C++ program which finds the greatest of three numbers input from the user. **(4+8+7)**

UNIVERSITY OF THE PUNJAB

Roll No. ....

M.A./M.Sc. Part – II Annual Examination – 2020

Subject: Statistics **Paper: VI (i) [Statistical Quality Control]** **Time: 3 Hrs.** **Marks: 100**

NOTE: Attempt any FOUR questions. All questions carry equal marks.

Q#1 (a)	Define Statistical Quality Control (SQC). Also describe the purposes of SQC.	10
(b)	Show how assignable causes of variation are identified on $\bar{x}$ and R chart?	15
Q#2	Sub-groups of 4 items each are taken from a manufacturing process at regular intervals. A certain quality characteristic is measured and $\bar{x}$ and R values are computed for each subgroup. After 25 groups, $$\sum_{i=1}^{25}\bar{x}_i = 15350, \quad \sum_{i=1}^{25}R_i = 411.4$$ a. Compute the control limits for $\bar{x}$ and R control charts. b. Assume that all points on both charts plot within the control limits. What are the natural tolerance limits of the process? c. All points on the $\bar{x}$ and R charts fall within the control limits. The specification requirements for this particular quality characteristic are 610 ± 15. If the quality characteristic is normally distributed, what are your conclusions regarding the ability of the process to produce items conforming to specifications? d. Assuming that if an item exceeds the upper specification limit it can be reworked, and if it is below the lower specification limit it must be scrapped, what percent scrap and rework is the process now producing?	25
Q#3 (a)	Suppose $\bar{x}$ chart is used with usual 3-sigma limits. The sample size is 5. Find the probability of detecting a shift to $\mu_1 = \mu_0 \pm 2\sigma$ on the first sample following the shift?	10
(b)	A textile manufacturer initiates the use of c-chart to monitor the number of imperfections found in bolt of cloth. Each bolt is same in length, width, weave, and fiber composition. A total of 145 imperfections were found in the last 25 bolts inspected. The four highest and lowest counts were: (see table below) a. Calculate 3-sigma control limits for this process. b. Is this process in control? If not what aimed at value of c_0'. c. Make revised control limits if necessary.	15
Q#4(a)	The following fraction non-conforming control chart with $n = 100$ is used to control the process: UCL=0.0750, CL=0.0400, LCL=0.0050 a. What is the probability of type-I error? b. Find the probability of type-II error if the true process fraction non-conforming is 0.0600. c. Find the ARL when the process is in control and the ARL when the process fraction non-conforming is 0.0600.	10
(b)	Differentiate an item-by-item and group sequential Sampling Plans.	15

Table for Q#3 (b):

Count of imperfections	
Highest	Lowest
20	4
16	4
10	5
9	5

P.T.O

Q#5	Suppose that a supplier ships components in lots of size 5000. A single sampling plan with $n = 50$ and $c = 2$ is being used for receiving inspection. Rejected lots are screened, and all defective items are reworked and returned to the lot. a. Draw the OC-curve for this plan. b. Find the level of lot quality that will be rejected 90% of the time. c. Management has objected to use the above sampling procedure and wants to use a plan with an acceptance number $c = 0$, arguing that this is more consistent with their zero-defects program. What do you think of this? d. Design a single sampling plan with $c = 0$ that will give a 0.90 probability of rejection of lots having the quality level found in part (b). Note that the two plans are now matched at the LTPD point. Draw the OC-curve for this plan and compare it to the one for $n = 50, c = 2$ in part (a).	25
Q#6(a)	An acceptance sampling plan for life testing requires that a sample of 19 items be tested with replacement for 1000 hours. If not more than 7 failures occur, the lot is accepted, otherwise it is rejected. Assume that the probability of failure is constant. Compute the mean life for which: a. The producer's risk of a lot rejection is 0.05. b. The consumer's risk of a lot acceptance is 0.10	10
(b)	Take a sampling plan with $n_1 = 50,\quad c_1 = 2,\quad n_1 + n_2 = 150,\quad c_2 = 5$ If the incoming lots have fraction nonconforming $p = 0.05$ then what is the probability of final acceptance? Calculate the probability of rejection on the first sample	15
Q#7	Write a short note on any **Five** of the following: i. Reliability and Life Testing ii. Process Capability iii. Bath-tub Curve iv. Producer's and Consumer's Risk v. CUSUM control chart vi. Warning Limits	5 each

UNIVERSITY OF THE PUNJAB

M.A./M.Sc. Part – II Annual Examination – 2020

Roll No. ....

Subject: Statistics | Paper: VI (iii) [Operations Research] | Time: 3 Hrs. | Marks: 100

NOTE: *Attempt any FOUR questions. All questions carry equal marks.*

Q.1 Explain your understanding about the following terms
(i) Decision variable (ii) Infeasible solution (iii) Minimax and Maximin (iv) Dummy Sources and destinations (v) Strategy (vi) Alternative optima (vii) Networking (viii) Queueing 25

Q.2 (a) A firm manufactures two types of products A and B and sells them at a profit of Rs. 2.00 on type A and Rs. 3.00 on type B. Each product is processed on two machines G and H. Type A requires one minute of processing time on G and two minutes on H; type B requires one minute on G and one minute on H. The machine G is available for not more than 6 hours 40 minutes while machine H is available for 10 hours during any working day. Formulate the problem as a linear programming problem. 10
(b) What is Linear Programming? How it is applicable in our daily life? 15

Q.3 (a) How many forms of linear programming have you browsed? Illustrate their properties with examples. 15
(b) Explain the optimality and feasibility conditions of dual simplex methods. 10

Q.4 (a) Explain what is degeneracy? Discuss with the following LP model 10
Max $X_o = 3X_1 + 9X_2$ Subject to $X_1 + 4X_2 \leq 8$: $X_1 + 2X_2 \leq 4$; $X_1, X_2 \geq 0$
(b)Temporarily degeneracy is an outcome of the LP-model.
Show with the following LP-model. Max $X_o = 3X_1 + 2X_2$ 15
Subject to $4X_1 + 3X_2 \leq 12$; $4X_1 + X_2 \leq 8$; $4X_1 - X_2 \leq 8$; $X_1, X_2 \geq 0$

Q.5 (a) Explain the role of slack, artificial and surplus variables in providing the solution of linear programming. 10
(b) Find optimal solution by Dual Simplex Method. Max $X_o = 2X_1 - X_2 + X_3$
Subject to $2X_1 + 3X_2 - 5X_3 \geq 4$; $-X_1 + 9X_2 - X_3 \geq 3$; $4X_1 + 6X_2 + 3X_3 \leq 8$;
$X_1, X_2, X_3 \geq 0$ 15

Q.6 (a) What is Transportation model? Explain its components. 5
(b) Consider the Transportation problem having the following cost table

		Destination				
		1	2	3	4	Supply
	1	10	0	20	11	15
Source	2	12	7	9	20	25
	3	0	14	16	18	5
Demand		5	15	15	10	

Solve by (i) North-West Corner Rule (ii) Least Cost Method (iii) VAM 20

Q.7 (a) Explain the Dominance property method to solve a game. 5
(b) Gave a graphical representation of the following game. Determine the value of the game and the optimal mixed strategy for the player who has two strategies.

	B			
	-4	8	2	-10
A	6	-2	0	5

20

UNIVERSITY OF THE PUNJAB

Roll No. ....

M.A./M.Sc. Part – II Annual Examination – 2020

Time: 3 Hrs. Marks: 50

Subject: Statistics Paper: VI (iv) (Part-A-Survey and Report Writing)

NOTE: Attempt any FOUR questions. All questions carry equal marks.

Q.1. Differentiate Experimental study and a Sample Survey. Describe the important steps in conducting a sample survey.

Q.2. What is the importance of accurate data in a study? Discuss the various types of errors that may affect the accuracy of data.

Q.3. What do you understand by determination of sample size? Which factors should be considered in making the decision?

Q.4. What is meant by ethics in research? Discuss the various aspects of ethical considerations in doing a research.

Q.5. Define validity and reliability. Also discuss their important types.

Q.6. Discuss the important factors in constructing a questionnaire.

Q.7. Discuss the various components of survey report.

UNIVERSITY OF THE PUNJAB

M.A./M.Sc. Part – II Annual Examination – 2020

Roll No. ....

Subject: Statistics **Paper: VII (i) (Time Series Analysis)** **Time: 3 Hrs. Marks: 100**

NOTE: Attempt any FOUR questions.

Q.1.a) Describe the useful transformations commonly used in time series. **(5)**

b) Define the following. **(10)**

i. Backward shift operator
ii. Stationarity
iii. Differencing
iv. SARIMA model

c) Find the mean, variance, autocovariance function and autocorrelation function of an MA(1) process, $Y_t = \theta Z_{t-1} + Z_t$. **(10)**

Q.2.a) Show that $Y_t = \phi Y_{t-1} + Z_t$ is stationary when $|\phi|<1$. **(5)**

b) Show that the autocorrelation function of $Y_t = Y_{t-1} - \frac{1}{2}Y_{t-2} + Z_t$ is given by **(12)**

$$\rho_k = \left(\frac{1}{\sqrt{2}}\right)^k \left(\cos\frac{\pi k}{4} + \frac{1}{3}\sin\frac{\pi k}{4}\right), \quad k = 0,1,2,3,\ldots$$

where $\{Z_t\}$ is a purely random process having zero mean and finite variance.

c) Derive the Yule-Walker equations for an AR(p) process. **(8)**

Q.3.a) Define the autoregressive process. Find the mean, variance and autocorrelation function of an AR(1) process. **(12)**

b) Show that the AR(2) process given by **(8)**

$$Y_t = Y_{t-1} + cY_{t-2} + Z_t$$

is stationary provided $-1 < c < 0$, where $\{Z_t\}$ is a purely random process having zero mean and finite variance.

c) Given r1=0.5 and r2=0.2, find the partial autocorrelation at lag 1and 2. **(5)**

Q.4.a) Show that the autocorrelation function of the qth order moving average process given by **(10)**

$$Y_t = \frac{1}{q+1}\sum_{i=0}^{q} Z_{t-i},$$

where $\{Z_t\}$ is a purely random process having zero mean and finite variance.

b) Show that the infinite order MA process $\{Y_t\}$ defined by **(10)**

$$Y_t = Z_t + \theta\sum_{i=1}^{\infty} Z_{t-i}$$

is non-stationary. Also show that $W_t = Y_t - Y_{t-1}$ is stationary. Find the autocorrelation function of W_t.

c) Show that the autocorrelation at lag 'k' lies between -1 and +1 i.e $-1 < \rho_k < 1$ **(5)**

Q.5 a) Describe the stages of Box-Jenkins model building approach. Discuss the tools for identification of appropriate candidate ARMA models. **(10)**

b) Discuss the difficulties in estimation of parameters of a Moving Average process. Provide the iterative least squares estimation procedure for the two parameters in a non-zero MA(1) process. Also discuss how the starting values can be provided for the iterative procedure. **(10)**

Show that for an AR(p) process $Y_t = \sum_{i=1}^{p} \varphi_j Y_{t-j} + Z_t$, **(5)**

$$\sigma_Y^2 = \frac{\sigma_Z^2}{1 - \sum_{i=1}^{p} \varphi_i \rho_i}$$

Where ρ_k is the autocorrelation at lag k.

Q.6.a) If $\{Y_t\}_{t=1}^{n}$ follows an AR(2) process $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t$ where $\{Z_t\}$ is independently and normally distributed process with zero mean and finite variance σ_z^2 then show that log-likelihood function is **(15)**

$$\ln L = const. - \frac{n}{2}\ln\sigma_z^2 + \frac{1}{2}\ln M_2 - \frac{1}{2\sigma_z^2}\left(\sum_{t=3}^{n}\left(Y_t - \sum_{i=1}^{2}\phi_i Y_{t-i}\right)^2 + X^T M_2 X\right)$$

where $(Y_1, Y_2) \sim N_2(0, \Sigma_2)$ and $M_2 = \sigma_z^2 \Sigma_2^{-1}$. Also suggest procedure to find the approximate maximum likelihood estimates of AR parameters.

b) Following are the autocorrelations obtained for the residuals by fitting an ARMA(1,2) process to an observed time series with n = 100. **(10)**

k	1	2	3	4	5	6	7	8	9	10
r_k	0.25	0.18	0.05	-0.01	0.03	0.01	0.05	-0.04	0.01	0.02

Calculate the value of Ljung-Box portmanteau test for m= 5 and see if the fitted model is a good fit for this observed time series. Use 5% level of significance.

Q.7.a) Show that for an ARIMA(1,1,1) process: $(1-\phi B)\nabla Y_t = (1+\theta B)Z_t$, the forecast at origin t with lead time l is given by **(10)**

$$Y_t(l) = Y_t + \frac{\phi(1-\phi^l)}{1-\phi}(Y_t - Y_{t-1}) + \frac{\theta(1-\phi^l)}{1-\phi}Z_t$$

b) An AR(1) model $(1-0.6B)(Y_t - 9) = Z_t$ is fitted to an observed time series of 100 observations. Now a practitioner wants to use this model to forecast Y_{101} and Y_{102}. **(15)**

i. Compute the MMSE forecasts of Y_{101} and Y_{102}, if $Y_{99} = 9, Y_{100} = 8.9$.

ii. Obtain 95% forecast limits for Y_{101}, if $\sigma_z^2 = 0.1$.

Suppose that Y_{101} comes out to be 8.8, update the forecast of Y_{102} by shifting forecast origin to $t = 101$.

UNIVERSITY OF THE PUNJAB

M.A./M.Sc. Part – II Annual Examination – 2020

Roll No. ....

Subject: Statistics Paper: VII (ii) (Multivariate Analysis) Time: 3 Hrs. Marks: 100

NOTE: Attempt any FOUR questions. All questions carry equal marks.

Q1. Let (12+5+8)

$$A = \begin{vmatrix} 9 & -2 \\ -2 & 6 \end{vmatrix}$$

a) Determine the Eigen values and Eigen vectors of **A**.
b) Write down the spectral decomposition of **A** and verify your results.
c) Using (a), write down the spectral decomposition of A^{-1}.

Q2. a) Show that the matrix for the quadratic form, $3x_1^2 + 2x_2^2 - 2\sqrt{2}x_1x_2$, is positive definite.

b) Consider, $\Sigma = \begin{vmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{vmatrix}$. Find ρ_{13}. (17+8)

Q3. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be a random sample from a joint distribution that has mean $\boldsymbol{\mu}$ and covariance matrix Σ. If $\boldsymbol{x}_i$'s are independent of each other, derive $E(\overline{\boldsymbol{x}})$ and $\text{cov}(\overline{\boldsymbol{x}})$. Is the sample covariance matrix $\boldsymbol{S}_n$ an unbiased estimator of Σ? (25)

Q4. a) Let **x** follows $N_3(\mu, \Sigma)$. Find the distribution of $\begin{vmatrix} X_1 - X_2 \\ X_2 - X_3 \end{vmatrix}$. (8+17)

b) Derive the moment generating function of **x** that follows $N_p(\mu, \Sigma)$.

Q5. a) Show that linear combination of Wishart matrices follow Wishart distribution.
b) Derive the additive property of Wishart matrices. (15+10)

Q6. Data for two variables give the summary: (7+11+7)

$$n = 42, \overline{\mathbf{x}} = \begin{bmatrix} 0.564 \\ 0.603 \end{bmatrix}, \mathbf{S} = \begin{bmatrix} 0.0144 & 0.0117 \\ 0.0117 & 0.0146 \end{bmatrix}$$

a) Test the hypothesis $H_0 : \boldsymbol{\mu} = [0.562 \quad 0.589]'$.
b) Find 95% confidence ellipse for $\boldsymbol{\mu}$. Also, obtain the lengths of major and minor axes.
c) Obtain T^2 simultaneous confidence intervals for the components of $\boldsymbol{\mu}$.

Q7. a) Show that the total variance explained by Principal Components is equivalent to the total variance of the data. (8+17)

b) The Eigen-values and Eigen-vectors of a correlation matrix are given below.

$$\lambda_1 = 2.85, \ \boldsymbol{e}_1 = [0.33 \quad 0.46 \quad 0.38 \quad 0.56 \quad 0.47]^T$$
$$\lambda_2 = 1.81, \ \boldsymbol{e}_2 = [-0.61 \quad 0.39 \quad -0.56 \quad 0.08 \quad 0.40]^T$$
$$\lambda_3 = 0.20, \ \boldsymbol{e}_3 = [0.10 \quad 0.74 \quad 0.17 \quad -0.60 \quad -0.22]^T$$
$$\lambda_4 = 0.10, \ \boldsymbol{e}_4 = [0.14 \quad -0.28 \quad 0.12 \quad -0.57 \quad 0.75]^T$$
$$\lambda_5 = 0.03, \ \boldsymbol{e}_5 = [0.70 \quad 0.07 \quad -0.71 \quad 0.00 \quad 0.01]^T$$

Assuming two-factor model, calculate and interpret common factor loadings, communalities and specific variances.