UNIVERSITY OF THE PUNJAB

Part-II : Supplementary Examination 2018
Examination:- M.A./M.Sc.

Roll No. ....

Subject: Statistics
PAPER: I (Statistical Inference)

MAX. TIME: 3 Hrs.
MAX. MARKS: 100

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

Q.1 a) Write the requirements by which any inference can be regarded as Statistical Inference. (05)

b) Let $x_1, x_2, x_3 .... x_n$ be a random sample, n being fixed, from a uniform distribution over the range (0,a). Check $\hat{M} = \max(x_1, x_2 ... x_n)$ *and* $\hat{Q} = \min(x_1, x_2 ... x_n)$ for possible unbiasedness and consistency. Where $\hat{M}$ and $\hat{Q}$ are order statistics. (12)

c) Let x_1, x_2, ... x_n be a random sample from N(0, θ) where $0 < \theta < \infty$ show that $\sum_{i=1}^{n} x_i^2 / n$ is an unbiased estimator of θ and has variance $2\theta^2 / n$ (08)

Q.2 a) State and prove Neyman Factorization theorem. (12)

b) Compare Cramer Rao inequality and Rao Blackwell's theorem. (07)

c) How we can obtain an efficient estimator? Write the names of the procedures. (06)

Q.3 a) Find the approximate ML estimator of parameter θ in the Cauchy distribution. Also find the asymptotic variance of this estimator. (13)

b) Obtain the MLE of the parameters of the distribution $\frac{1}{\Gamma(p)\theta^p} x^{p-1} e^{-x/\theta}$, $x \geq 0$. (12)

Q.4 a) Compare the properties of ML, Least square and Moment estimators. What is the history of Moment estimators?. (12)

b) For the Double Poisson Distribution
Find the moment and ML estimates of parameters (13)

Q.5 a) Let $x_1, x_2 .... x_n$ denote a random sample from Bernoulli distribution. Assume that the prior distribution of θ is given by g(θ) = $I_{(0,1)}\theta$. Find the posterior Bayes estimators of θ and $\tau(\theta) = \theta(1-\theta)$. (07)

b) What is the difference between Baysian and Classical inference? explain. What is Baye's estimator? Differentiate between prior and posterior density. (08)

P.T.O.

c) Based on a random sample of size ‘n’ from the density $f(x/\theta) = 1/\theta$ with prior distribution as $g(\theta) = 1, \quad 0 < \theta < 1$ obtain the Baye's estimator for θ with respect to the loss function $\ell(\theta, t) = (t - \theta)^2 / \theta^2$ **(10)**

Q.6 a) Let there are 3n observations $x_1, x_2 \ldots x_n$, $y_1, y_2 \ldots y_n$, $z_1, z_2 \ldots z_n$ with same unknown variance σ^2, the mean values of observations are given by $E(x_i) = 0.2\theta_1 + 0.3\theta_2 + 0.5\theta_3$, $E(y_i) = 0.3\theta_1 + 0.5\theta_2 + 0.2\theta_3$, $E(z_i) = 0.5\theta_1 + 0.2\theta_2 +$ where $\theta_1, \theta_2, \theta_3$ are unknown parameters. Apply the least square method to derive the estimates of the contrasts $(\theta_1 - \theta_2), (\theta_2 - \theta_3)$ and $(\theta_3 - \theta_1)$ by using $\theta_1 + \theta_2 + \theta_3 = 1$ and obtain the unbiased estimate of σ^2. Also find the variance of each parameter θ_i and the variance of each contrast. **(13)**

b) Define and explain, 95% confidence Interval, confidence belt and confidence region. **(06)**

c) Construct large sample confidence interval for unknown parameter of the distribution $f(x;\theta) = \lambda e^{-\lambda x}, 0 < x < \infty$. **(06)**

Q.7 a) Define sequential sampling, what is their importance? **(04)**

b) Let $\bar{x}$ and $\bar{y}$ be the sample means of random samples, each of size n drawn independently from $N(\mu_1, 400)$ and $\mu(\mu_2, 225)$. The null hypothesis H₀: $\theta = 0$ against H₁: $\theta > 0$ is rejected if and only $\bar{x} - \bar{y} \geq K$, where $\theta = \mu_1 - \mu_2$. Find n and k so that $\pi(0) = 0.05$ and $\pi(10) = 0.90$, where $\pi(\theta)$ is the power function. **(12)**

c) Let $x \sim N(0,1)$ and if you are interested to test $H_O : \theta = \theta_O = 0$ against $H_1 : \theta = \theta_1 = 1$, obtain Average sample number (ASN) for both states of nature. Assume $\alpha_a = 0.01, \beta_a = 0.01$. **(09)**

UNIVERSITY OF THE PUNJAB

Part-II : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. ....

Subject: Statistics
PAPER: II (Regression Analysis and Econometrics)

MAX. TIME: 3 Hrs.
MAX. MARKS: 100

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

Q.1.a) What is Econometrics? Discuss its methodology.

b) Discuss the procedure and assumptions of CHOW Test for comparison of simple linear regressions.

Q.2.a) Consider GLR model $\underline{Y} = X\,\underline{\beta} + \underline{\in}$ *Such that* $E\underline{\in} = 0,\ E\underline{\in}\underline{\in}' = \sigma^2 I$. Derive the best linear unbiased estimator of $C'\,\underline{\beta}$, linear combinations of parameters.

b) The equation $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \in_t$ is estimated with quarterly data from 2001 to 2006 inclusive yielding the explained sum of squares 114.8 and unexplained sum of squares 19.6. compute co-efficient of determination. When three seasonally dummy variables were added in the above equation and the equation was re-estimated, the explained sum of squares now become 120.7. Test for the presence of seasonality.

Q.3.a) What is ridge regression estimates? When these estimate are used. Find mean and variance of the estimate and nature of bias, if any.

b) Let $\underline{Y} = X\,\underline{\beta} + \underline{\in}$ *Such that* $\underline{\in} \sim N(0,\ \sigma^2 I)$.and elements of $\underline{\beta}$obey the restriction $R\beta = \underline{\gamma}$. Obtain the variance covariance matrix of the restricted least squares estimator of $\underline{\beta}$.

Q.4.a) Let $\underline{Y} = X\,\underline{\beta} + \underline{\in}$ *Such that* $\underline{\in} \sim N(0,\ \sigma^2 V)$.Obtain MLE of$\sigma^2$and discuss its sampling distribution.

b) For GLR model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \in$,when error terms are non-spherical, show that generalized least squares estimators of β's are BLUE.

Q.5.a) What is Heteroskedasticity? Discuss the assumption and procedure of Gold feld-Quandt test for heteroskedasticity.

b) Differentiate between perfect and imperfect multicollinearity. Discuss the practical consequences of multicollinearity.

Q.6.a) If in regression model error terms follow AR(I) of the form $\in_t = \rho \in_{t-1} + u_t$ such that u_t's fallow OLS assumptions, then show that $\in_t$'s are non-spherical.

b) Differentiate between distributed lab and autoregressive models. Discuss the features of Koyck transformation for distributed lag models.

Q.7. For the model $y_1 = \alpha_1 y_2 + \alpha_2 x_2 + u_1, y_2 = \beta_1 y_1 + \beta_2 x_1 + \beta_3 X_3 + u_2$ you are given the following information:

$\hat{y}_1 = 5x_1 + 10x_2 + 2x_3$; $\hat{y}_2 = 10x_1 + 10x_2 + 5x_3$

$\Sigma x_1^2 = 2,\ \Sigma x_2^2 = 4,\ \Sigma x_3^2 = 20,\ \Sigma x_1 x_2 = \Sigma x_2 x_3 = \Sigma x_1 x_3 = 0$, Estimate the parameters, where possible, by appropriate method.

UNIVERSITY OF THE PUNJAB

Part-II : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. ....

Subject: Statistics
PAPER: VI (i) [Statistical Quality Control]

MAX. TIME: 3 Hrs.
MAX. MARKS: 100

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

Q#1 (a)	Differentiate the role of Action and Warning limits in statistical process control.	10
(b)	Define attribute control charts? Also discuss some advantages and disadvantages of attribute control charts.	15
Q#2	Samples of $n = 8$ items each are taken from a manufacturing process at regular intervals. A quality characteristic is measured, and $\bar{x}\, and\, R$ values are calculated for each sample. After 50 samples we have $$\sum_{i=1}^{50} \bar{x}_i = 2000\, and \sum_{i=1}^{50} R_i = 250$$ Assume that the quality characteristic is normally distributed. a. Compute control limits for the $\bar{x}\, and\, R$ control charts. b. All the points on both control charts fall between the control limits computed in part (a). What are the natural tolerance limits of the process? c. If the specification limits are 41 ± 5.0 what are your conclusions regarding the ability of the process to produce items within these 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 producing? e. Make suggestions as to how the process performance could be improved.	25
Q#3 (a)	What to do if we want to deal with low defect levels?	10
(b)	A process produces rubber belts in lots of size 2500. Inspection records on the last 20 lots reveal the following data.	15

Lot Number	Number of Nonconforming Belts	Lot Number	Number of Nonconforming Belts
1	230	11	456
2	435	12	394
3	221	13	285
4	346	14	331
5	230	15	198
6	327	16	414
7	285	17	131
8	311	18	269
9	342	19	221
10	308	20	407

i. Compute trial control limits for a fraction nonconforming control chart. Also make a decision.
ii. If you wanted to set up a control chart for controlling future production, how would you use these data to obtain the center line and control limits for the chart?

Q#4(a)	Define acceptance sampling. What are the purposes of acceptance sampling procedure?	10
(b)	Draw the Type-B OC curve for the single sampling plan $n = 50,\ c = 2$.	15
Q#5 (a)	i. Derive an item-by-item sequential Sampling Plan for which $p_1 = 0.01,\ \alpha = 0.05,\ p_2 = 0.06,\ \beta = 0.10$ ii. Draw the OC curve for this plan.	15
(b)	Discuss rectifying inspection in sequential Sampling Plan.	10
Q#6(a)	State some modern definitions of reliability and life testing.	10
(b)	Take a sampling plan with $n_1 = 50, c_1 = 2, n_2 = 100, c_2 = 5$ If the incoming lots have fraction non-conforming $p = 0.05$, what is the probability of acceptance on the first sample? What is the probability of final acceptance? Also calculate the probability of rejection on first sample.	15
Q#7	Write a short note on any **Five** of the following: i. Cumulative Sum (CUSUM) Chart ii. ISO-14000 principles iii. Fast Initial Response (FIR) iv. Producer's and Consumer's Risk v. OC-Curve vi. Modified Control Chart	5 each

UNIVERSITY OF THE PUNJAB

Part-II : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. ....

Subject: Statistics **MAX. TIME: 3 Hrs.**

PAPER: VI (iv) [Part-A-Survey and Report Writing] **MAX. MARKS: 50**

NOTE: Attempt any FOUR questions. All questions carry equal marks. (12.5×4)

Q.1. Explain the difference between survey and census. Explain the uses and advantages of sample survey in detail.

Q.2.
- a) Define the term "Literature Review". Its advantages and importance in a survey research.
- b) What are the qualities of a "Good Survey"?

Q.3. How should validity and reliability of an instrument be tested?

Q.4.
- a) Differentiate between survey and observation
- b) Explain in detail the concept of data analysis.

Q.5. Define the terms Population, Sample and Sample Size. How will you determine the optimal sample size?

Q.6. How is primary data different from secondary data? Explain the methods of primary data collection in detail.

Q.7. Write notes on any three of the following:
- a) Purpose of coding
- b) Purpose of Literature review
- c) Pretesting and its advantages
- d) Coverage error
- e) Editing of data

UNIVERSITY OF THE PUNJAB

Part-II : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. ....

Subject: Statistics
PAPER: VII (i) [Time Series Analysis]

MAX. TIME: 3 Hrs.
MAX. MARKS: 100

NOTE: Attempt any FOUR questions.

Q.1.a) Define stationarity and describe the transformations commonly used to transform a non-stationary time series into a stationary time series. (5)

b) Describe the following. (10)

i. Components of time series
ii. Relation between Backward shift operator and Differencing operator
iii. Stochastic process
iv. Mixed autoregressive moving average process

c) Express an AR(2) process: $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t$ in Moving average representation: (10)

$Y_t = \sum_{i=0}^{\infty} \psi_i Z_{t-i}$. Express ψ_i in terms of ϕ_1 and ϕ_2.

Q.2.a) Show that random walk is a non-stationary process. Transform the random walk into a stationary process and discuss. (5)

b) Show that for an AR(2) process (10)

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

the autocorrelation at lag k is given by

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

c) Derive the stationarity conditions for an AR(2) process. (10)

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

b) Derive the equivalent MA and AR representation of an ARMA(1,1) process. (8)

c) 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_i is the ith lag autocorrelation.

Q.4.a) Define moving average process and obtain mean, variance and autocorrelation function of an MA(q) process given by (12)

$$Y_t - \mu = \sum_{i=1}^{q} \theta_i Z_{t-i}$$

b) Show that the infinite order MA process given below is non-stationary; (6)

$X_{t=}Z_t + c\sum_{i=0}^{\infty} Z_{t-i}$,
where c is a constant. Also show that the first order differences have a first order MA process and is stationary process. Derive the autocorrelation function of this differenced series.

c) Show that the autocorrelation at lag 'k' lies between -1 and +1 i.e $-1 < \rho_k < 1$ (7)

(P.T.O)

Q.5.a) Describe the four stages of Box-Jenkins model building approach. **(5)**

b) Describe the iterative procedure of obtaining the least squares estimates of an MA(1) process with non-zero mean. Write the initial iteration and the stopping rule. Also suggest some starting values of the parameters. **(10)**

c) If $\{Y_t\}_{t=1}^{n}$ follows an AR(1) process $Y_t = \phi Y_{t-1} + z_t$ 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 given by **(10)**

$$\ln L = const. - \frac{n}{2}\ln\sigma_z^2 + \frac{1}{2}\ln\left(1-\phi^2\right) - \frac{1}{2\sigma_z^2}\left(\sum_{t=1}^{n} Y_t^2 - 2\phi\sum_{t=2}^{n} Y_t Y_{t-1} + \phi^2\sum_{t=2}^{n-1} Y_t^2\right).$$

Also show that the maximum likelihood estimate of AR parameter is

$$\hat{\phi} = \frac{\sum_{t=2}^{n} Y_t Y_{t-1}}{\sum_{t=2}^{n-1} Y_t^2}.$$

Q.6.a) Show that the minimum mean squared error forecast with origin at n and lead time l is given by **(15)**

$$Y_n(l) = \frac{1}{\psi(B)}\left[\frac{\psi(B)}{B^l}\right]_+ Y_n$$

where $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$ and $\{\psi_j : j=1,2,\ldots\}$ are the weights of moving average representation. Moreover, using this rule, find the forecasts of Y_{100+l} : $l = 1,2$ for a non-zero AR(1) process given by

$$Y_t - 53 = 0.65(Y_{t-1} - 53) + z_t.$$

b) Show that the covariance between forecast errors for different lead times but with the same forecast origin is given by **(10)**

$$cov\left(e_n(l), e_n(l+j)\right) = \sigma_z^2 \sum_{i=0}^{l-1} \psi_i \psi_{i+j}$$

Q.7.a) Derive the rule of updating forecast i.e. $Y_{n+1}(l) = Y_n(l+1) + \psi_l\left(Y_{n+1} - Y_n(1)\right)$ **(10)**

b) Using the weighted sum of past observations, find the forecasts of Y_{93} and Y_{94}. Also find 95% forecast interval for an AR(2) process given by **(15)**

$$(1 - 0.6B - 0.2B^2)(Y_t - 50) = z_t$$

where $n = 92, Y_{91} = 56.5, Y_{92} = 48.3, \sigma_z^2 = 9$

UNIVERSITY OF THE PUNJAB

Part-II : Supplementary Examination 2018

Examination:- M.A./M.Sc.

Roll No. ....

Subject: Statistics
PAPER: VII (ii) [Multivariate Analysis]

MAX. TIME: 3 Hrs.
MAX. MARKS: 100

NOTE: Attempt any FOUR questions.

Q1. (a) Explain the following: *(4,4,2)*

a) Positive Definite and Semi-Positive Definite Matrices
b) Eigen Values and Eigen Vectors
c) Generalized Variance.

(b) The random vector $\mathbf{X}' = [X_1 \quad X_2 \quad X_3 \quad X_4]$ has a Multivariate Normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ given by: *(15)*

$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 0 \\ 3 \\ 7 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 7 & -1 & 0 & 3 \\ & 8 & -2 & 6 \\ & & 12 & 9 \\ & & & 3 \end{bmatrix}$$

Suppose $\mathbf{Y}_1' = [X_1 \quad X_2]$ and $\mathbf{Y}_2' = [X_3 \quad X_4]$ are the sub-vectors of $\mathbf{X}$ then find

a) $\mathbf{E}\left({}^{\mathbf{Y}_1}\!/\!{}_{\mathbf{Y}_2} \right)$ b) $\mathbf{Cov}\left({}^{\mathbf{Y}_1}\!/\!{}_{\mathbf{Y}_2} \right)$

Q2. (a) If $\mathbf{X}$ denotes a $(p \times 1)$ column vector of random variables, $\boldsymbol{\mu}$ is a column vector of constants and Σ is a positive definite matrix then find the value of k such that $\mathbf{f(X)} = \mathbf{k} \exp\left[-\frac{1}{2}(\mathbf{X} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \right]$ is a pdf $\mathbf{X}$. Also find $E(\mathbf{X})$. *(18)*

(b) Show that $(\mathbf{X} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2_{(p)}$. *(07)*

Q3. (a) Let $\mathbf{W} \sim \mathbf{W}_p(\mathbf{f}, \Sigma, \mathbf{M})$. If $\mathbf{C}$ is any $(p \times q)$ matrix of constants, then show that $\mathbf{C'WC} \sim \mathbf{W}_q(\mathbf{f}, \mathbf{C'\Sigma C}, \mathbf{MC})$. *(10)*

(b) Data for two variables give the summary: *(15)*

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

Test the hypothesis $H_0 : \boldsymbol{\mu} = [0.562 \quad 0.589]'$. Find 95% confidence ellipse for μ consisting of all values (μ_1, μ_2). Also draw the ellipse.

Q4 (a) Let X_1 and X_2 be two random variables with covariance matrix: *(18)*

$$\Sigma = \begin{bmatrix} 9 & \sqrt{6} \\ \sqrt{6} & 4 \end{bmatrix}$$

i) Obtain the Principal Components and find the percentage of variation explained by each.
ii) Change Σ into correlation matrix ρ and find Principal components using ρ.
iii) Are Principal components from (i) and (ii) same? If yes, why?

(b) What is Principal Compnonet Analysis? How it is different from Factor Analysis? *(07)*

P.T.O.

Q5.	**(a)**	Explain the method of factor analysis , indicating the assumptions involved.	*(15)*
	(b)	The eigen values and eigen vectors of the covariance matrix $$\Sigma = \begin{bmatrix} 1.00 & 0.63 & 0.45 \\ & 1 & 0.35 \\ & & 1.00 \end{bmatrix}$$ are $\lambda_1 = 1.96; \lambda_2 = 0.68; \lambda_3 = 0.36$ and $\beta_1 = \begin{bmatrix} 0.625 \\ 0.593 \\ 0.567 \end{bmatrix}, \beta_2 = \begin{bmatrix} -0.219 \\ -0.491 \\ 0.843 \end{bmatrix} \& \beta_3 = \begin{bmatrix} 0.749 \\ -0.638 \\ -0.177 \end{bmatrix}$ i) Assume m = 1 factor model, calculate the loading matrix and matrix of specific variances using principal component solution method. ii) What portion of the total population variance is explained by the first common factor? iii) What percentage of the variation is explained by specific factor?	*(10)*
Q6.	(a)	Suppose that population 1 $P_1 \sim N(\mu_1, \sigma_1^2)$ and population 2 $P_2 \sim N(\mu_2, \sigma_2^2)$. Discuss the maximum likelihood discriminant rule.	*(10)*
	(b)	The following mean vector and covariance matrices are based on $n_1 = n_2 = 100$ observations: $\bar{\mathbf{X}}_1 = \begin{bmatrix} 6.213 \\ 3.133 \end{bmatrix}, \bar{\mathbf{X}}_2 = \begin{bmatrix} 7.412 \\ 5.321 \end{bmatrix}, \mathbf{S}_1 = \begin{bmatrix} 1.813 & 0.321 \\ & 0.937 \end{bmatrix}, \mathbf{S}_2 = \begin{bmatrix} 2.193 & 1.654 \\ & 3.789 \end{bmatrix}$ Find the Fisher's Linear Discriminat function and Discriminant rule. Also allocate the new observations $\bar{\mathbf{X}}' = [7.2 \quad 3.1]$ to any of these populations.	*(15)*
Q7.		What is Canonical Correlation? Derive the canonical correlations and canonical variables.	*(25)*