CENTRAL UNIVERSITY OF HARYANA

Second Semester Term End Examinations August- September 2022

Programme : M.Sc. Statistics | **Session: 2021-22**

Semester : Second | *Max. Time: 3 Hours*

Course Title : Time Series and Statistical Quality Control | *Max. Marks: 70*

Course Code : SBS ST 01 201 DCE 3104

Instructions:

1. Question no. 1 has seven parts and students need to answer any four. Each part carries three and half Marks.
2. Question no. 2 to 5 have three parts and student need to answer any two parts of each question. Each part carries seven marks.

Question No. 1. **(4x3.5=14)**

a. What are chance and assignable causes of variability? What part do they play in the operation and interpretation of a Shewhart control chart?

b. Describe the mathematical models of a time series. How can an additive model be considered as a particular type of multiplicative model?

c. Define auto correlation and partial correlation.

d. Explain AOQL and LTPD.

e. What do you understand by seasonal variation in a time series? Give one example.

f. Define process control and product control with suitable example for each.

g. Distinguish between Producer's risk and Consumer's risk.

Question No. 2 **(2x7=14)**

a. In the Usual notations, prove that

$$\frac{1}{m}[m]y_0 = [y_0 + \frac{m^2-1}{24}\delta^2 y_0]$$

Where $\frac{1}{m}[m]$ stands for the simple average of m terms. Further show that

$$\frac{1}{m_1 m_2 \dots\dots m_r}[m_1][m_2] \dots\dots [m_r]y_0 = [y_0 + \frac{m_1^2 + m_2^2 + \cdots + m_r^2 - r}{24}\delta^2 y_0]$$

Hence deduce Spancer's 15-point formula.

b. What do you understand by time series? Explain in detail the trend and irregular components of a time series.

c. Explain variate-difference method for trend analysis. How is the appropriate order of differencing determined?

Question No. 3 **(2x7=14)**

a. Explain ARIMA and obtain the estimates of its parameters.

b. Let $\varepsilon_t, \varepsilon_{t+1}, \ldots\ldots$ and ξ be independent variables with zero mean and unit variance. Making $y_t = a\xi + \varepsilon_t, -\infty < t < \infty$,
Show that the process is stationary with correlation.

c. Define moving average process and obtained correlogram of MA (2) Process.

Question No. 4 **(2x7=14)**

a. What are modified control limits? Explain how they are derived. Why is it required to construct modified control limits?

b. What are control charts for attributes? Derive control chart for controlling the proportion defectives when sample size is not fixed, giving clearly the statistical concept used.

c. What is the statistical justification for using the three sigma limits in the control charts irrespective of the actual probability distribution of the quality characteristic?

Question No. 5 **(2x7=14)**

a. Distinguish between single sampling plan and double sampling plan. Obtain an expression for AOQL in single sampling plan.

b. What is an OC curve? Obtain its expression in sequential sampling plans for attributes. Also give five points of OC Curve.

c. Explain various types of variables sampling plans. Give advantages and disadvantages of variables sampling.

CENTRAL UNIVERSITY OF HARYANA

Second Semester Term End Examinations August-September 2022

Programme: M.Sc. Statistics | **Session: 2021-22**

Semester: Second | **Max. Time: 3 Hours**

Course Title: Regression Analysis | **Max. Marks.: 70**

Course Code: SBS ST 01 202 C 3104

Instructions:

1. Question no. 1 has seven parts and students need to answer any four. Each part carries three and a half Marks.
2. Questions no. 2 to 5 have three parts and students need to answer any two parts of each question. Each part carries seven marks.

Question No. 1. **(4x3.5=14)**

a. Explain the direct regression model or ordinary least square Model.

b. Obtain ordinary least squares estimators for the parameters in the case of a simple linear regression model and discuss their properties.

c. Discuss the assumption used in the multiple linear regression model.

d. Find the covariance matrix of the parameter vector of a multiple linear regression model.

e. Define the Predicted residual error sum of square (PRESS) statistic.

f. What is the usefulness of orthogonal polynomials?

g. What is the explanatory variable selection procedure?

Question No. 2 **(2x7=14)**

a. Discuss the testing of hypothesis and confidence interval estimation for intercept parameter in the simple linear regression model.

b. Describe the coefficient of determination R^2 and adjusted R^2 in a multiple linear regression model $y = \beta X + \varepsilon$.

c. Prove that the ordinary least squares estimator (OLSE) is the best linear unbiased estimator (BLUE) in the case of multiple linear regression model.

Question No. 3 **(2x7=14)**

a. Give some realizations of the residual plot for checking the homogeneity of variances in a regression model.

b. Discuss the variance stabilizing transformations to correct model inadequacies.

c. What do you mean by leverage and influence points in regression analysis? Explain in detail.

Question No. 4 **(2x7=14)**

a. Discuss the generalized least square estimation method for the multiple linear regression model.

b. Discuss the estimation of unknown parameters using orthogonal polynomials and related analysis of variance.

c. Explain the statistical consequences of the exclusion of relevant variables.

Question No. 5 **(2x7=14)**

a. Why logistic regression is useful? What are the link functions, logistic functions and odds ratio?

b. Discuss the estimation problem in logistic regression using maximum likelihood estimation.

c. Discuss the problem of testing of hypothesis in logistic regression.

CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, August- September, 2022

Programme : M. Sc. Statistics
Semester : II
Course Title : Statistical Inference-I
Course Code : SBS ST 01 201 C 3104

Session: 2021-2022
*Max. Time : **3 Hours***
*Max. Marks : **70***

Question no. 1 has seven sub parts and students need to answer any four. Each sub part carries three and half Marks.
Question no. 2 to 5 have three sub parts and students need to answer any two sub parts of each question. Each sub part carries seven marks.

Question No. 1. **(4X3.5=14)**

a) Define Minimal Sufficient Statistic.

b) State and prove Rao Blackwell Theorem.

c) Explain the methods of maximum likelihood estimator and write down its small sample properties

d) What do you mean by the family of distribution having a monotone likelihood ratio?

e) Find the *maximum likelihood estimator* of θ if sampling from double exponential distribution.

f) To test $H_0 : \mu = 100$, $H_A : \mu \neq 100$. A random sample of size 50 is drawn from a normal population with unknown mean μ and variance 200. If $\alpha = 0.05$, specify the critical value.

g) Discuss the terms "Point estimation", "Interval estimation" and testing of hypothesis with at least one example each.

Question No. 2. **(2X7=14)**

a) Let $X_1, X_2, \cdots, X_n$ be *iid* random variables with *pdf*
$f(x;\theta_1,\theta_2) = c(\theta_1,\theta_2)\, m(x)$, $\theta_1 < x < \theta_2$. Find sufficient statistic for (θ_1,θ_2).

b) If T_1 and T_2 be two unbiased estimators of a parameter θ with variance σ_1^2, σ_2^2 and correlation ρ. Find best unbiased linear combination of the two, also find its variance.

c) Let $X_1, X_2, \cdots, X_n$ be a random sample drawn from $U(0, \theta)$. Show that the estimator $T_n(X) = \left(\prod_{i=1}^{n} X_i\right)^{1/n}$ is consistent for θe^{-1}.

Question No. 3. **(2X7=14)**

a) Let $X_1, X_2, \cdots, X_n$ be *iid* from Poisson variate $P(\theta)$ find minimum variance unbiased estimator (MVUE) for $\frac{e^{-\theta}\theta^y}{y!}$.

b) Let X_1 and X_2 be a sample from a distribution with the *pdf*
$f(x;\theta) = \frac{2}{\theta^2}(\theta - x)$, $0 < x < \theta$.

Find the maximum likelihood estimator (MLE) of θ. Based on single observation, show that $2X$ is the MLE of θ but not unbiased.

c) Let $x_1, x_2, \cdots, x_n$ be a random sample drawn from $c(1, \theta)$ with *pdf* $f(x) = \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}$, $-\infty < x < \infty$. Show that the Cramer Rao lower bound (CRLB) for the variance of an unbiased estimator of θ is $2/n$.

Question No. 4. **(2X7=14)**

a) Show that double exponential family of distribution has monotone likelihood ration when α is un known and β is known.

b) Let $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_n$ be independent random sample from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively. Obtain likelihood Ratio Test of $H_0 : \mu_1 = \mu_2$, $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ against $H_1 : \mu_1 \neq \mu_2$, $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ when σ_1^2 and σ_2^2 are known.

c) Let $X_1, X_2, \cdots, X_n$ be *iid* $U(\theta - 1/2, \theta + 1/2)$, $\theta \in \Re^1$. Show that the statistic $T(X) = (X_{(1)}, X_{(n)})$ is minimal sufficient but not complete.

Question No. 5. **(2X7=14)**

a) Find the *MP* size α $(= 0.05)$ test for testing $H_0 : \mu = 0$ vs $H_A : \mu = 1$, given a random sample of size 25 from $N(\mu, 16)$ population.

b) Let X be a random sample of size n from a normal population with mean μ and variance σ^2. Find out the likelihood ratio test of

$H_0 : \mu \leq \mu_0$ against $H_1 : \mu > \mu_0$ when σ^2 is known.

c) Let $X_1, X_2, \cdots, X_n$ be a random sample from the $N(\theta, \sigma^2)$ population. Find $(1-\alpha)\%$ level of confidence interval.

CENTRAL UNIVERSITY OF HARYANA

Second Semester Term End Examinations August-September 2022

Programme : M.A/ M.Sc.
Semester : Second
Course Title : Applied Statistics
Course Code : SBS ST 01 201 GE 3104

Session: 2021-22
Max. Time: 3 Hours
Max. Marks: 50

Instructions:

1. Question no. 1 has five parts and students need to answer any four. Each part carries two and half Marks.
2. Question no. 2 to 5 have three parts and student need to answer any two parts of each question. Each part carries five marks.

Note: Scientific Non programmable Calculator is allowed

Question No. 1. **(4x2.5=10)**

a. What are chance and assignable causes of variability? What part do they play in the operation and interpretation of a Shewhart control chart?

b. Describe the mathematical models of a time series. How can an additive model be considered as a particular type of multiplicative model?

c. Fill in the blank of the following table which are marked with ?

X	l_x	d_x	q_x	L_x	T_x
20	693435	?	?	?	35081126
21	690673	-	-	-	-

d. Write the merits and demerits of specific death rate.

e. What do you understand by seasonal variation in a time series? Give one example.

Question No. 2 **(2x5=10)**

a. For the following data fit a second degree parabola by using least square method

Year	2003	2004	2005	2006	2007	2008	2009
Sale	-5	-2	5	16	31	50	73

b. What do you understand by time series? Explain in detail the trend and irregular components of a time series.

c. Using Ratio to Trend method to determine the quarterly seasonal indices for the following data.

Year	Quarterly production in ('ooo tonnes)			
	Quarter 1	Quarter II	Quarter III	Quarter IV
2010	75	60	54	59
2011	86	65	63	80
2012	90	72	66	85
2013	100	78	72	93

Question No. 3 **(2x5=10)**

a. What is the statistical justification for using the three sigma limits in the control charts irrespective of the actual probability distribution of the quality characteristic?

b. The following data are found during the inspection of the first 15 samples of size 100 each from a lot of two-wheelers manufactured by an automobile company

Sample No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Number of Defectives	3	4	6	2	12	5	3	6	3	5	4	15	5	2	3

Calculate the control limits for fraction defectives.

c. As part of an overall quality improvement programme, a textile manufacturer decides to monitor the number of defects found in each inspected bolt (large bundle) of cloth. The data from 20 inspections are recorded in the table given below

Bolt of Cloth	1	2	3	4	5	6	7	8	9	10
Number of Defects	10	19	5	9	2	8	7	13	3	2
Bolt of Cloth	11	12	13	14	15	16	17	18	19	20
Number of Defects	22	4	6	9	7	2	5	12	4	2

Construct the appropriate control chart and Comment whether the process is in Statistical Central.

Question No. 4 **(2x5=10)**

a. Define vital statistics. What are vital events? Describe the usual sources of data collection on vital events.

b. Describe infant mortality rate in detail. Define foetal death and Still Birth.

c. Compute the standardised death rates of the two populations A and B regarding A as standard population from the following data:

Age Group Years	A		B	
	Population	Death	Population	Death
Under 10	20000	600	12000	372
10 – 20	12000	240	30000	660
20 – 40	50000	1250	62000	1612
40 – 60	30000	1050	15000	325
Above 60	10000	500	3000	180

Question No. 5 **(2x5=10)**

a. Describe the structure of a complete life table. Explain the different columns of the life table.

b. From the following data, calculate the gross reproductive rate and net reproductive Rate.

Age group	No. of children born to 1000 woman passing through the age group	Mortality rate (per 1000)
16 – 20	150	120
21 – 25	1500	180
26 – 30	2000	150
31 – 35	800	200
36 – 40	500	220
41 – 45	200	230
46 – 50	100	250

Sex Ratio being males & females is 52:48

c. Calculate the General Fertility Rate and Total Fertility Rate from the following data. Assuming that for every 100 girls, 106 boys are born:

Age of Woman	No. of Woman	Age specific fertility rate (per 1000)
15 – 19	212619	98.0
20 – 24	198732	169.6
25 – 29	162800	158.2
30 – 34	145362	139.7
35 – 39	128109	98.6
40 – 44	106211	42.8
45 - 49	86753	16.9

CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, August- September 2022

Programme: M.Sc (Statistics) **Session: 2021-22**

Semester: II **Max. Time: 3 Hours**

Course Title: Design of Experiments **Max. Marks: 70**

Course Code: SPMS ST 01 203 C 3104

Instructions:

1. Question no. 1 has seven sub parts and students need to answer any four. Each sub part carries three and half Marks.

2. Question no. 2 to 5 have three sub parts and students need to answer any two sub parts of each question. Each sub part carries seven marks.

Question No. 1. (4X3.5=14)

a) Explain with suitable example, how replication and local control enables to increase the efficiency of a design.

b) What is a uniformity trial? Discuss its use in field experiments.

c) What is split-plot design? Discuss the advantages and disadvantages of split-plot design relative to randomised block design.

d) What is a treatment contrast? When are two such contrasts said to be orthogonal?

e) What is meant by confounding in a factorial experiment? Why is confounding used even at the cost of loss of information on the confounded effects?

f) The following are two key-blocks of a lay out plan before randomization for a 2^4 experiment with factors A, B, C and D

Replication I : (1) abc abd cd

Replication II: (1) abc acd bd.

Identify the confounded effects.

g) Define a balanced incomplete block design.

Question No. 2. (2X7=14)

a) Derive the analysis of covariance for a two-way layout with one concomitant variable.

b) Show that the efficiency of latin square design compared to randomized block design, with rows of latin square design as blocks of the randomized block design, is given by $\frac{(m-1)MSC+(m-1)^2 MSE}{m(m-1)MSE}$, and the efficiency of latin square design compared to complete randomized design is given by $\frac{(m-1)MSR+(m-1)MSC+(m-1)^2 MSE}{(m^2-1)MSE}$.

c) Derive the layout and the analysis of an randomized block design with one missing value.

Question No. 3. (2X7=14)

a) Explain in detail the analysis of a 2^3 experiment conducted in randomised block.

b) A 2^5 factorial experiment with factor A, B, C, D and E are arranged in 4 blocks of 8 plots each. If 4 elements of one of the blocks are $(1), ab, cd, e$. What are the block compositions of the four blocks? Find the factorial effects confounded.

c) Construct a 2^{6-2} fractional factorial design using the design generators I = ACE and I = ACDF. Write down the aliases of main effects and resolution of the design.

Question No. 4. (2X7=14)

a) Derive the expression to measure the efficiency of balanced incomplete block design relative to randomized block design.

b) Derive the expected values of different sums of squares in the intra-block analysis of variance of a balanced incomplete block design.

c) Write down the ANOVA table for a split-plot design with two factors A and B, where factor A is main- plot treatment and factor B is sub-plot treatment.

Question No. 5. (2X7=14)

a) Obtain a half fraction of a 2^7 factorial experiment with the defining contrast I = ABCDEFG. Also write down the complete alias structure.

b) Explain the crossover designs. Discuss their advantages and disadvantages.

c) Write sort notes on the following terms: (i) Complete confounding and
(ii) Partial confounding.