(Pages : 2)

7819—III S STA—March 2021

THIRD SEMESTER M.A./M.Sc. DEGREE (C.B.C.S.) EXAMINATION, MARCH/APRIL 2021

Statistics

STCT 3.1—ELEMENTARY STOCHASTIC PROCESSES

Time : Three Hours

Maximum : 75 Marks

Question No. 1 *is* **compulsory.**

Answer any **four** *questions from the remaining.*

Each question carries 15 *marks.*

1. a) What does the Markov property state ? When do the states of a Markov chain are recurrent and transient ?

   b) What is a period of a state ? When is a state periodic and aperiodic ?

   c) Show that sum of two independent Poisson processes is a Poisson process.

   d) Define a Poisson process. Give two examples.

   e) Define residual and spent life times. (5 × 3 = 15 marks)

2. a) Show that the initial distribution and one-step transition probabilities together determine all the finite dimensional distributions associated with a Markov chain.

   b) State and prove the first entrance theorem. (7 + 8 = 15 marks)

3. a) Let $X_i = 0$ and $X_i = 1$ if it rains and if it doesn't rain on the i^{th} day respectively. Suppose $p_{00} = 0.7$ and $p_{10} = 0.4$. Then find the probability that it will rain on Friday given it was raining on Monday.

   b) Describe a one-dimensional random walk and classify its states as recurrent and transient.

   (5 + 10 = 15 marks)

4. a) Show that for any arbitrary stochastic process, if the inter-arrival times are exponential then the underlying process is Poisson process.

   b) State and prove the relationship between Poisson process and Geometric distribution.

   c) Given that only one occurrence of the Poisson process $\{N(t), t \geq 0\}$ has occurred by time instant t, then the distribution of the time interval s during which it has occurred is uniform in $(0, t)$.

   (6 + 6 + 3 = 15 marks)

Turn over

5. a) Describe in brief Birth and Death Process. What are Immigration-Emigration process and Linear growth process.

b) Show that in an One-dimensional Brownian motion the increments are independent and stationary. (5 + 10 = 15 marks)

6. a) What is a renewal function Show that for a renewal process $\{N(t), t \geq 0\}$ with renewal function $m(t)$.

$$m(t) = \sum_{n=1}^{\infty} F_n(t).$$

b) Prove that $\frac{N(t)}{t} \to \frac{1}{\mu}$ as $t \to \infty$ with probability 1 given that $\mu = E[X_n] \leq \infty$.

c) State the Elementary renewal theorem and the Key renewal theorem.

(7 + 4 + 4 = 15 marks)

7. a) Consider the process $X(t) = A\cos(wt) + B\sin(wt)$, where A and B are uncorrelated random variables each with mean zero and variance 1. If w is a positive constant, then show that $X(t)$ is a wide-sense stationary process.

b) Let $\{Z_n, n = 0, 1, 2....\}$ be a branching process and let probability generating function of its offspring distribution $\{p_k\}$ be given by $P(s)$. Then show that the probability generating function of $\{Z_n\}$ is the n-fold composition of $P(s)$ with itself. Further show that

$$P_n(s) = P_{n-1}(P(s))$$
and $P_n(s) = P(P_{n-1}(s))$. (6 + 9 = 15 marks)

8. Write short notes on any *three* of the following :

a) The Gambler's Ruin problem.

b) Yule-Furry birth process.

c) Diffusion process.

d) Auto-regressive process.

7820—III S STA—March 2021

THIRD SEMESTER M.A./M.Sc. DEGREE (C.B.C.S.) EXAMINATION, MARCH/APRIL 2021

Statistics

STCT 3.2—TESTING OF HYPOTHESES

Time : Three Hours | Maximum : 75 Marks

Answer any **five** *questions including question* 1.
Each question carries 15 *marks.*

Question No. 1 *is* **compulsory**. *Answer any* **four** *from the remaining* **seven** *questions.*

1. Answer the following questions :

   (a) Let θ be the probability that a coin will fall head in a single toss in order to test $H_0 : \theta = \frac{1}{2}$ against $H_1 : \theta = \frac{3}{4}$ The coin is tossed 5 times and H_0 is rejected if more than 3 heads are obtained. Find the two error probabilities and power of the test.

   (b) Define MLR property and show that B (n, θ) possesses MLR property.

   (c) Explain asymptotic property of LRTs with an example.

   (d) Explain runs test.

   (e) Derive Walds equation.

(5 + 3 = 15 marks)

2. (a) Define most powerful (MP) test. State and prove the sufficiency part of Neyman-Pearson theorem to obtain MP test of size α $(0 < \alpha < 1)$.

   (b) Based on a random sample size n from *Exp.* (θ), obtain a most powerful test (MP) of level α for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1 \ (> \theta_0)$.

   (c) Based on a random sample of size n from $N\left(\mu, \sigma^2\right)$, obtain a Uniformly Most Powerful (UMP) test of level α for testing $H_0 : \sigma^2 = \sigma_0{}^2$ against $H_1 : \sigma^2 = \sigma_1{}^2 \left(> \sigma_0{}^2\right)$.

(6 + 5 + 4 = 15 marks)

Turn over

3. (a) Obtain a UMP test for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$ based on a random sample of n observations from $B(1, \theta)$. Also find power function of the test.

(b) Define UMP unbiased test. Let $X_1, X_2, \ldots, X_n$ be iid $N(\mu, \sigma^2)$, where μ is known. Obtain a UMPU test of size α for testing $H_0 : \sigma^2 = {\sigma_0}^2$ against $H_1 : \sigma^2 \neq {\sigma_1}^2$.

(6 + 9 = 15 marks)

4. (a) Let $X_1, X_2, \ldots, X_n$ be iid random variables from $N(\mu, \sigma^2)$, where both μ and σ^2 are unknown. Obtain a likelihood ratio test of size α for testing $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$.

(b) Explain the application of LRTs to contingency table.

(7 + 8 = 15 marks)

5. (a) Define UMA and UMAU confidence intervals. Let $X_1, X_2, \ldots, X_{n1}$ be iid random variables from $N(\mu_1, \sigma_1^2)$, and $Y_1, Y_2, \ldots, Y_{n1}$ be iid random variables from $N(\mu_2, \sigma_2^2)$. X and Y are independent and $\sigma_1^2 = {\sigma_2}^2 = \sigma^2$. Derive a LRT at size α for testing $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$. Also obtain UMAU confidence intervals for $\mu_1 - \mu_2$ at confidence co-efficient $1 - \alpha$.

(b) Define confidence set estimator and shortest length confidence interval. Based on a random sample of size n from $N(\theta, \sigma^2)$, σ^2 is unknown. Obtain a confidence interval for θ at confidence co-efficient $1 - \alpha$ under $H_0 : \theta > \theta_0$ against $H_1 : \theta \neq \theta_0$.

(10 + 7 = 15 marks)

6. (a) Describe sign test.

(b) Describe Kolmogorov-Smirnov test for two samples.

(c) Explain Spearman's rank correlation co-efficient.

(5 + 5 + 5 = 15 marks)

7. (a) Define Wald's SPRT and derive the stopping bounds of SPRT in terms of its strength (α, β).

(b) Define OC and ASN functions. Let $X_1, X_2, \ldots, X_n$ be iid random variables from $f(x/\theta) = \frac{1}{\theta} e^{-x/\theta}$, $0 < x < \infty$. Describe Wald's SPRT for testing $H_0 : \theta > \theta_0$ against $H_1 : \theta \neq \theta_1 (> \theta_0)$. Obtain OC and ASN functions of SPRT.

(6 + 9 = 15 marks)

8. Write short note any *three* of the following :

(a) UMPU test for multi-parameter exponential family.

(b) Mann-Whitney -Wilcoxon test for two samples.

(c) Median test.

(d) Sequential analysis.

(3 × 5 = 15 marks)

(Pages : 2) **7821—III S STA—March 2021**

THIRD SEMESTER M.A./M.Sc. DEGREE (C.B.C.S.) EXAMINATION, MARCH 2021

Statistics

STCT 3.3—STATISTICAL ORIENTED R-PROGRAMMING

Time : Two Hours Maximum : 40 Marks

Question No. 1 *is* **compulsory.**

Answer any **three** *from the remaining.*

1. (a) What are variables ? Give an example.

   (b) Explain basic R operators.

   (c) Describe "next" statement with an example.

   (d) Explain "vectors" and "lists". Give one example to each one of them.

   (e) Two matrices A and B are defined as follows :

$$A = \begin{bmatrix} 2 & 6 \\ 3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 3 & 2 \\ 5 & 7 \end{bmatrix}.$$

   What will be the output of matrix $C = A * B$? (5 × 2 = 10 marks)

2. (a) Describe about procedure of using R packages.

   (b) Explain "apply" family of functions with syntax along with an example to each member.

   (4 + 6 = 10 marks)

3. (a) Describe JSON and XML files.

   (b) Discuss about the following tests with R syntax :

      (i) Paired *t*-test.

      (ii) Chi-square test.

      (iii) One-way ANOVA. (4 + 6 = 10 marks)

4. (a) Explain the variable selection procedure in linear regression in R.

   (b) How are the assumptions of Linear Regression Model are verified using R ?

   (6 + 4 = 10 marks)

Turn over

5. Explain any *two* of the following :

   (a) Data type conversions in R.

   (b) Advantages and disadvantages of R language.

   (c) Help system in R.

(5 + 5 = 10 marks)

(Pages : 2) **7822—III S STA—March 2021**

THIRD SEMESTER M.A./M.Sc. DEGREE (C.B.C.S.) EXAMINATION, MARCH 2021

Statistics

STCT 3.4—OPERATIONS RESEARCH

Time : Three Hours | Maximum : 75 Marks

Question No. 1 *is* **compulsory.**
Answer any **four** *from the remaining.*
Each question carries 15 *marks.*

1. (a) Define Linear Programming Problem (LPP). Give an example.

   (b) Discuss about the economic interpretation of duality.

   (c) Describe the applications of transportation problem.

   (d) Define Poisson queue. Give two examples.

   (e) Explain about various costs involved in inventory.

(5 × 3 = 15 marks)

2. (a) Distinguish between standard and canonical forms of LPP. Express the following LPP in both of these forms.

   Minimize $z = 2x_1 + 4x_2 + 8x_3$

   subject to

   $$12x_1 - 8x_2 + 5x_3 \le 12$$
   $$-3x_1 + 7x_2 - 6x_3 = 8$$
   $$x_1 - 2x_2 + x_3 \ge -15$$

   (b) Explain graphical and analytical methods of solving LPP. Illustrate any one of them.

(8 + 7 = 15 marks)

3. (a) Explain two-phase method of solving LPP with an example. Discuss about its advantage over big M method.

   (b) Explain the concept of duality. Outline the dual simplex algorithm.

(8 + 7 = 15 marks)

4. (a) Compare simplex and revised simplex procedures.

   (b) Define transportation problem. Explain and illustrate VAM of obtaining initial bfs.

   (c) Explain Hungarian method of solving an assignment problem.

(4 + 7 + 4 = 15 marks)

Turn over

5. (a) Explain $(M|M|1):(\infty|FIFO)$ queue and obtain its steady state solution.

(b) Obtain the queue length, mean and variance of the above model.

(8 + 7 = 15 marks)

6. (a) State and prove Little's formula. Discuss about its applications.

(b) Distinguish between deterministic and probabilistic inventory models. Stating the assumptions, derive EOQ for economic lot size model with uniform rate of demand, finite rate of replenishment having no shortages.

(5 + 10 = 15 marks)

7. (a) Explain single period model with instantaneous demand.

(b) Explain and illustrate inventory model with two price breaks.

(9 + 6 = 15 marks)

8. Write short notes on any *three* :

(a) Theorem of complementary slackness.

(b) Post optimality analysis.

(c) Characteristics of a queue.

(d) (s,S) policy.

(3 × 5 = 15 marks)

(Pages : 2) **7823—III S STA—March 2021**

THIRD SEMESTER M.A./M.Sc. DEGREE (C.B.C.S.) EXAMINATION, MARCH/APRIL 2021

Statistics

STCT 3.5—ECONOMETRICS (OPTIONAL)

Time : Three Hours Maximum : 75 Marks

Questions No. 1 *is* Compulsory.

Answe any **four** *from the remaining.*

Each question carries 15 *marks.*

1. (a) Define multiple linear regression model with assumptions.

 (b) Outline step-wise regression.

 (c) Define heteroscedasticity, what properties OLS estimators will possess under heteroscedasticity.

 (d) Discuss the problem of identifiability of a equation of system.

 (e) Write about full information maximum likelihood estimation procedure.

(5 × 3 = 15 marks)

2. (a) Derive the OLS estimators of co-efficients and σ^2 in simple linear regression model.

 (b) Show that the estimators of co-efficients obtained in (2*a*) are BLUE.

(9 + 6 = 15 marks)

3. (a) Obtain an unbiased estimator of variance of disturbance term of the multiple linear regression model.

 (b) State the procedures of selection of variables in multiple regression.

(9 + 6 = 15 marks)

4. (a) Describe the method of least absolute residuals to solve the problem of presence of outliers in the data.

 (b) Write about recording errors and errors of measurement of regression model variables.

 (c) Write a note on stochastic regression.

(7 + 4 + 4 = 15 marks)

5. (a) Describe how multicollinearity is detected.

 (b) Explain the use of : (i) Priori information ; and (ii) Principal components in solving multicollinearity.

(7 + 8 = 15 marks)

Turn over

6. (a) Outline the sources of autocorrelation.

   (b) Describe Durbin-Watson test for autocorrelation.

   (c) Explain use of d-statistic in estimating auto correlations parameters.

(4 + 6 + 5 = 15 marks)

7. (a) What are reasons for lags in econometric model.

   (b) Write importance of mean lag and median lag in the study of econometric models.

   (c) Write a procedure of estimating econometric model using Adaptive expectation.

(5 + 4 + 6 = 15 marks)

8. (a) Define : (i) Structural Form ; (ii) Reduced form.

   (b) Describe rank and order conditions for identifiability of a equation of system.

   (c) Describe Three Stage Least Squares estimation procedure.

(4 + 6 + 5 = 15 marks)

[5 × 15 = 75 marks]

(Pages : 3) **7824—III S STA—OE—March 2021**

THIRD SEMESTER M.A./M.Sc. DEGREE (C.B.C.S.) EXAMINATION, MARCH/APRIL 2021

Statistics

STET 3.1—APPLIED STATISTICS

Time : Three Hours

Maximum : 75 Marks

Answer any **five** *questions including question* **one.**
Each question carries 15 *marks.*

1. a) Distinguish between price index numbers and quantity index numbers.

   b) Describe census method of obtaining vital statistics

   c) Explain basic principles of design of experiments

   d) Define systematic sampling and state its advantages

   e) Describe the construction of p-chart

(5 × 3 = 15 marks)

2. a) Write about seasonal and cyclical variations of time series. Explain measurement of trend.

   b) Calculate Fisher's price index number for the following data :

Commodity	Base year		Current year	
	Price	Quantity	Price	Quantity
A	8	25	12	30
B	10	20	18	26
C	7	35	13	40
D	9	22	10	15
E	3	15	5	10

(7 + 8 = 15 marks)

3. a) Define various measures of fertility rates.

   b) Calculate crude death rate and age-specific death rates for the following data :

Turn over

Age (in years)	Population ('000)	Number of deaths
0-4	25	1354
5-14	20	1009
15-24	16	885
25-34	15	649
35-54	12	235
55-70	14	428
70 and above	23	1489

(7 + 8 = 15 marks)

4. a) Define completely randomized design and explain its statistical analysis.

   b) Carry out the statistical analysis for the following data :

Blocks	Treatments					
	1	2	3	4	5	6
I	12.1	18.1	23.4	27.2	32.1	30.1
II	13.2	19.3	24.5	28.1	33.1	29.8
III	12.5	22.2	22.8	27.9	34.2	28.7
IV	12.6	21.5	25.1	27.2	33.7	28.6

(7 + 8 = 15 marks)

5. a) Compare probability and nonprobability sampling. Mention various schemes under both.

   b) Describe simple random sampling and stratified random sampling.

(7 + 8 = 15 marks)

6. a) Explain the analysis of plotted points on control charts

   b) Describe how control limits of $\overline{X}$ and R charts are obtained.

(7 + 8 = 15 marks)

7. a) Producer's risk and consumer's risk.

   b) Explain the operation of double sampling plan.

   c) Compare process control and product control.

(4 + 7 + 4 = 15 marks)

8. Write short notes on any *three* of the following :

   a) Sampling and non-sampling errors.

   b) Factor reversal test.

   c) Mortality rates.

   d) Latin square design.

   e) Cost of living index number.

(3 × 5 = 15 marks)