C 83559 (Pages : 2) Name..................................

Reg. No..................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2020

(CCSS)

STATISTICS

STA 2C 09—REGRESSION METHODS

(2019 Admissions)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (i) List out the properties of Least Square Estimates.

(ii) Define R^2 and adjusted R^2.

(iii) Describe Likelihood Ratio Test.

(iv) Define Prediction Intervals.

(v) Describe the concepts of lack of fit and pure error.

(vi) Write a Short Note on spline smoothing.

(vii) What is the need of normal probability plots ?

(viii) What is Random Explanatory Variable ? Discuss its effect in a Regression Model.

($4 \times 4 = 16$ marks)

Section B

Answer either part A **or** *part B of all questions.*
Each question carries 16 *marks.*

II. A. a) Describe the assumptions underlying the classical linear regression model.

b) Prove that least square estimate and MLE of the parameters of normal linear regression model coincides.

($8 + 8 = 16$ marks)

Or

B. a) Explain the methods of estimation in regression models with linear restrictions.

b) Describe generalized least square method.

($8 + 8 = 16$ marks)

Turn over

III. A. a) Explain the procedure for testing the significance of regression parameters in simple linear regression model.

b) Find out the Confidence Interval for the slope and Intercept in Simple Linear Regression Model.

(8 + 8 = 16 marks)

Or

B. a) Define the Co-efficient of Determination. How do you intercept it ? Is it related in any way with Multiple Correlation Co-efficients ?

b) What is Confidence bands for the Regression Surface ? How they used in linear regression ?

(8 + 8 = 16 marks)

IV. A. a) Define polynomial Regression and discuss its estimation methods.

b) Describe the method of weighted least square for straight line regression.

(8 + 8 = 16 marks)

Or

B. a) Describe the Method of fitting Orthogonal Polynomials.

b) Explain logistic regression model. What are its uses ?

(8 + 8 = 16 marks)

V. A. a) Define Collinearity. Discuss the method of Diagnosing it.

b) Discuss the effect of outliers in a Regression Model.

(8 + 8 = 16 marks)

Or

B. a) Describe how you use residuals for checking the adequacy of the model fitted.

b) Distinguish between Ridge Regression and Principle Component Regression.

(8 + 8 = 16 marks)

C 83558 (Pages : 2) Name..............................

Reg. No..............................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2020

(CCSS)

Statistics

STA 2C 08—DESIGN AND ANALYSIS OF EXPERIMENTS

(2019 Admissions)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (i) What are the Guidelines of Experimental Design. Also note two Applications of the Experimental Design.

(ii) Give a brief note on Regression approach in ANOVA.

(iii) Explain the analysis of Completely Randomized Design.

(iv) What is Youden square - Lattice Design ?

(v) Explain the salient features of Fractional Factorial Design.

(vi) Construct a 2^6 design in blocks of 8 plots confounding ABC, ADE and BCDE.

(vii) Describe the effect components of a 2^3 design into seven mutually orthogonal contrast.

(viii) What are Multiple Responses ?

(4 × 4 = 16 marks)

Section B

Answer either part A **or** *part B of all questions.*
Each question carries 16 *marks.*

II. A. a) What are the Principles of Design ? Explain why RBD is much better than CRD in light of the principles of design.

b) Explain the process of choice of sample size and its importance in the experimental design.

(8 + 8 = 16 marks)

Or

B. a) Distinguish between Fixed effect and Random effect Model.

b) Describe the concept of Two way ANOVA with Interaction and its analysis.

(8 + 8 = 16 marks)

Turn over

III. A. a) Describe the concept and analysis of Latin Square Design.

b) Explain Greaco-Latin Square Design.

(8 + 8 = 16 marks)

Or

B. a) Distinguish between BIBD and PBIBD.

b) Note down the model for BIBD and give its analysis.

(8 + 8 = 16 marks)

IV. A. a) Describe two factor factorial design and give its analysis.

b) Explain the Yate's procedure of obtaining various effects and sum of squares of a 2^k factorial experiments.

(8 + 8 = 16 marks)

Or

B. a) Describe confounding with suitable example.

b) Write down the confounded arrangement of a 3^3 design by confounding the Interactions ABC and BC^2 into blocks. Identify other interactions of any which get confounded in our arrangement. Hence analyse the design if there are two replications of the same type of arrangement of the treatment.

(8 + 8 = 16 marks)

V. A. a) What are Response Surface Designs ?

b) Give some Experimental Designs for Fitting and Analysing Response Surface Designs.

(8 + 8 = 16 marks)

Or

B. a) Explain ANCOVA.

b) Develop the analysis of RBD with one concomitant variable.

(8 + 8 = 16 marks)

C 83557 (Pages : 3) Name..

Reg. No..

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2020

(CCSS)

Statistics

STA 2C 07—STATISTICAL INFERENCE—I

(2019 Admissions)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 marks.

I. (i) Let X_1, X_2 be i.i.d. $P(\lambda)$ r.vs. Show that $T = (X_1 + X_2)/3$ is sufficient but $T_1 = X_1 + 3X_2$ is not sufficient for λ.

(ii) What do you mean by a complete statistic ? Give an example of a statistic which is sufficient but not complete.

(iii) Define exponential family of density. Show that $\left\{N\left(\mu, \sigma^2\right), \mu \in R, \sigma > 0\right\}$ is an exponential family. Obtain a minimal sufficient statistic for $\left(\mu, \sigma^2\right)$.

(iv) Define CAN estimator. Examine whether sample median is a CAN estimator for μ in the Cauchy population $C(\mu, \sigma), \mu \in R, \sigma > 0$.

(v) Obtain the moment estimators of α and β in the Gamma distribution $G(\alpha, \beta)$ based on a sample of size n from this population.

(vi) Lind out the MLE of θ based on a sample of size n from $U(\theta, \theta + 2), \theta \in R$.

(vii) Distinguish between Bayesian and fiducial intervals.

(viii) Let $X_1, X_2, \ldots, X_n$ be a random sample from $U(0, \theta)$. Obtain $(1 - \alpha)$ level UMA confidence interval for θ.

(4 × 4 = 16 marks)

Turn over

Section B

Answer **either** *part A* **or** *part B of* **all** *questions.*
Each question carries 16 *marks.*

II. A (a) State and prove Cramer-Rao inequality.

(b) Examine whether MVB estimator for θ exists for the exponential population with mean θ. If so obtain the MVB estimator.

Or

B (a) State and prove Rao-Blackwell theorem.

(b) Obtain UMVUE of $2\mu^2 + 3$ in $N(\mu, 1), \mu \in R$ based on a sample of size n.

III. A (a) State and prove the invariance property of CAN estimators.

(b) Obtain the MLE and moment estimators of θ in $U(\theta, 2\theta), \theta > 0$ and examine whether these estimators are CAN.

Or

B (a) If T is a consistent estimator of θ show that $1 + 2e^{T}$ is a consistent estimator of $1 + 2e^{\theta}$.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Show that $T = \bar{X} + S^2$ is a CAN estimator of $\mu + \sigma^2$.

IV. A (a) Explain maximum likelihood method of estimation. Show that under some regularity conditions to be stated MLE is a CAN estimator.

(b) Explain Bayes method of estimation. Obtain Bayes estimate of λ under squared error loss, in $P(\lambda), \lambda > 0$, when the prior distribution of λ is an exponential distribution with mean 5.

Or

B (a) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Obtain Bayes estimator under squared error loss for the parameter μ when σ^2 is known and the prior distribution of μ is $N(\mu_0, \sigma_0^2)$.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from $b(1, \theta), \theta \in [1/3, 3/4]$. Find an MLE of θ if it exists.

V. A (a) What do you mean by a large sample confidence interval ? Obtain confidence interval for the $p_1 - p_2$ based on random samples from two independent binomial populations $b(n, p_1)$ and $b(m, p_2)$.

(b) Let X_1 and X_2 be two independent observations from the exponential population with pdf

$$f(x;\theta) = e^{-(x-\theta)}, x > \theta.$$

Let $Y = \min(X_1, X_2)$. Find the confidence co-efficient of the interval $[Y - 1/2, Y + 1/2]$.

Or

B (a) Let $X_1, X_2, \ldots, X_n$ be a random sample from

$$f_\theta(x) = \frac{1}{2\theta} \exp\left(\frac{-|x|}{\theta}\right), x \in R, \theta > 0.$$

Find the shortest length $(1-\alpha)$ level confidence interval for θ, based on the sufficient statistic $\sum_{i=1}^{n} |X_i|$.

(b) Let $\bar{X}$ be the mean of a random sample from $N(\mu, 25)$. Find the sample size n such that $(\bar{X} - 2, \bar{X} + 2)$ is a 0.90 level confidence interval for μ.

(4 × 16 = 64 marks)

C 83556 (Pages : 2) Name.......................................

Reg. No..................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2020

(CCSS)

Statistics

STA 2C 06—PROBABILITY THEORY–II

(2019 Admissions)

Time : Three Hours Maximum : 80 Marks

Part A

Answer any **four** *questions.*
Each question carries 4 marks.

I. 1 State and prove Hellys convergence theorem.

2 Show that characteristic function is real iff the distribution function is symmetric about zero ?

3 For the sequence $\{X_n\}$ of independent random variables with $P\{X_k = \pm\sqrt{k}\} = 1/2$, does the WLLN hold.

4 State Borel Cantelli lemma. Is the converse is true. Justify your answer.

5 State and prove Kolmogorov three series theorem.

6 Show that Liapunov condition implies Lindberg-Feller condition.

7 Define sub-Martingale and super-Martingale. Give examples.

8 For the sequence $\{X_n\}$ of independent random variables with $P\{X_k = \pm 2^k\} = 1/2$, does the SLLN hold.

(4 × 4 = 16 marks)

Part B

Answer **either** *part A* **or** *part B.*
Each question carries 16 marks.

II. A (i) State continuity theorem on characteristic function.

(ii) Show that characteristic function is uniformly continuous over R.

Or

B (i) State and prove Inversion theorem on characteristic function.

(ii) Find the probability density function corresponding to the characteristic function

$$\phi(t) = \begin{cases} 1-|t| \text{ if } |t| \le 1 \\ 0 \text{ other wise} \end{cases}.$$

Turn over

III. A (i) State and prove Helly Bray lemma.

(ii) State and prove Scheffe's theorem.

Or

B (i) Prove that probability of the tail events of a sequence of independent random variable is either 0 or 1.

(ii) State and prove Kolmogorov three series theorem.

IV. A (i) Establish Kolmogorov strong law of large numbers for a sequence of independent random variables.

(ii) For the sequence $\{X_k\}$ of independent random variables with $P\{X_k = \pm 2^k\} = \frac{1}{2^{2k+1}}$ and $P\{X_k = 0\} = 1 - \frac{1}{2^{2k}}$ does the SLLN hold.

Or

B (i) State and prove Liaponov CLT.

(ii) State Lindeberg -Feller form of CLT.

V. A (i) State and prove Kolmogorov inequality.

(ii) Define Martingales. Show that if $\{X_n\}$ is a sequence of independent random variables with $E(X_n) = 0$ then $Y_n = \sum_{k=1}^{n} X_k$ is a martingale sequence.

Or

B (i) Define conditional expectation. State and prove its important properties.

(ii) Define infinite divisibility. Examine the infinite divisibility of Poisson distribution and Normal distribution.

(4 × 16 = 64 marks)

C 83555 (Pages : 3) Name.......................................

Reg. No...................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2020

(CCSS)

Statistics

STA 2C 09—REGRESSION METHODS

(2013 Admissions)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 marks.

I. (a) State Gauss-Markov theorem and explain its importance.

(b) Distinguish between generalized least squares estimate and ordinary least squares estimates for a linear regression model $Y = X\beta + \varepsilon$.

(c) For the linear model $Y = X\beta + \varepsilon$, where X is $n \times p$ of rank p and $\varepsilon \sim N\left(0, \sigma^2 I_n\right)$, obtain the likelihood ratio test for the hypothesis $H_0 : A\beta = C$.

(d) For a straight line regression model $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$; $i = 1, 2, ..., n$, obtain confidence intervals for the slope and intercept.

(e) Construct a $100(1-\alpha)\%$ confidence interval for an estimable linear parametric function $\lambda'\theta$, in a less than full rank model.

(f) Discuss the concept of orthogonal polynomials.

(g) What is collinearity ? Point out its consequences.

(h) Explain the effect of non-constant variance in multiple linear regression models. How will you remove the problem of non-constant variance.

(4 × 4 = 16 marks)

Turn over

Section B

Answer **either** *A* **or** *B of all questions.*
Each question carries 16 *marks.*

II. A. (a) In the linear model $Y = X\beta + \varepsilon$, obtain the least squares estimate of the parameter β and discuss the properties of the estimate.

(b) If $Y \sim N_n\left(X\beta, \sigma^2 I_n\right)$, where X is $n \times p$ of rank p, then prove that :

(i) $\dfrac{\left(\hat{\beta}-\beta\right)' X'X\left(\hat{\beta}-\beta\right)}{\sigma^2} \sim \chi^2_p$,

(ii) $\hat{\beta}$ is independent of S^2.

Or

B. (a) In the linear model $Y = X\beta + \varepsilon$, derive the distributions of (i) $\hat{\beta}$; (ii) $RSS = \left(Y - X\hat{\beta}\right)'\left(Y - X\hat{\beta}\right)$. Show that $\dfrac{RSS}{n-p}$ is an unbiased estimate of σ^2, where p is the rank of X.

(b) Explain the method of Lagrange Multipliers for estimating parameters in a multiple linear regression model with linear constraints.

III. A. (a) For the linear regression model $Y = X\beta + \varepsilon$, to test the hypothesis $H : A\beta = c$, prove that (i) $\dfrac{\left(A\hat{\beta}-c\right)'\left[A\left(X'X\right)^{-1}A'\right]^{-1}\left(A\hat{\beta}-c\right)}{qS^2}$ is distributed as $F_{q,n-p}$; (ii) When $c = 0$, F can be expressed in the form $F = \dfrac{n-p}{q}\,\dfrac{Y'\left(P-P_H\right)Y}{Y'\left(I_n-P\right)Y}$, where P_H is a symmetric idempotent matrix $P_H P = P P_H = P_H$.

(b) Define multiple correlation co-efficients. For the general linear model $Y = X\beta + \varepsilon$, prove that :

(i) $\sum_i \left(Y_i - \bar{Y}\right)^2 = \sum_i \left(Y_i - \hat{Y}_i\right)^2 + \sum_i \left(\hat{Y}_i - \bar{Y}\right)^2.$

(ii) $R^2 = \dfrac{\sum_i \left(\hat{Y}_i - \bar{Y}\right)^2}{\sum_i \left(Y_i - \bar{Y}\right)^2}.$

where R^2 is the co-efficient of determination.

Or

B. (a) In the regression model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$, explain how to test (i) $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4$; and (ii) $H_0 : \beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$.

(b) What are Bonferroni t-intervals ? Explain. Also explain how prediction intervals are constructed for the response in a linear regression model.

IV. A. (a) What are polynomial regression models ? Explain the important considerations that arise while fitting polynomial models in one variable.

(b) Describe the weighted least squares technique for the straight line model with known weights.

Or

B. (a) Explain polynomial regression in several variables.

(b) What is the need for piecewise polynomial fitting ? Discuss the method of splines in this context.

V. A. (a) What is the bias of multiple linear regression models due to underfitting ? Suppose that we postulate the model $E(Y) = \beta_0 + \beta_1 x$ when the true model is $E(Y) = \beta_0 + \beta_1 x + \beta_2 x^2$. If we use observations of Y at $x_1 = -1, x_2 = 0$ and $x_3 = 1$ to estimate β_0 and β_1 in the model postulated, then what bias will be introduced ?

(b) Stating the assumptions of linear regression model, explain how residuals are useful in model assumptions checking.

Or

B. (a) What are outliers ? Explain the effect of outliers in the linear regression model. Explain the leave-one-out diagnostic method.

(b) Explain the methods of detecting collinearity. What are the remedies for collinearity ?

(4 × 16 = 64 marks)

C 83554 (Pages : 2) Name..................................

Reg. No..............................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2020

(CCSS)

Statistics

STA 2C 08—DESIGN AND ANALYSIS OF EXPERIMENTS

(2013 Admissions)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (a) Explain the concepts : Randomization and replication. How do they increase the efficiency of an experiment ?

(b) Differentiate between fixed effect and random effect models. Give example to each model.

(c) Describe how efficiency of RBD is estimated compared to CRD.

(d) Discuss the role of partially balanced incomplete block designs (PBIBD) and define it.

(e) What is confounding in factorial experiment ? Explain its need.

(f) Describe Yate's method of computing all the factorial effects in a 2^n factorial experiment.

(g) What are response surfaces ? Cite some of its applications.

(h) Explain the term analysis of covariance. Write down a model for a RBD with v treatments in r blocks and a covariate x.

(4 × 4 = 16 marks)

Section B

Answer **either** *A* **or** *B of all questions.*
Each question carries 16 *marks.*

II. (A) (a) For the linear model $y_{ij} = \mu + \alpha_i + \epsilon_{ij}, i = 1, 2, ..., v ; j = 1, 2, ..., n_i$, obtain the least squares estimators of the parameters and derive a test of the hypothesis $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_v$. State the underlying assumptions involved.

(b) List various model adequacy checking procedures commonly used in design of experiments. Explain any two techniques.

Or

Turn over

(B) (a) State and prove Gauss-Markov theorem. Discuss the importance of it in the analysis of variance.

(b) Discuss the method of fixing the sample size in experiments. Also explain the regression approach of analysis of variance.

III. (A) (a) What is missing plot technique ? Suppose that in a RBD one observation is missing, how do you estimate it ? What are the modifications to be made then in the ANOVA ?

(b) What is a Latin Square Design ? Stating the basic model explain the analysis of data from such a designs.

Or

(B) (a) Explain the importance of Incomplete Block Designs (ICD) and hence define Balanced Incomplete Block Design (BIBD). Derive the parametric relationships of BIBD.

(b) Write a short note on Lattice design.

IV. (A) (a) In a 2^3 factorial experiments how the analysis is performed in order to get the main effects and interaction effects of treatments.

(b) Express the main effects and interactions in terms of treatment means for 2^4 factorial experiment.

Or

(B) (a) If there are three replications of the experiment with 2^4 factorial set up and in each of which ABC and BCD are completely confounded, outline the analysis of the experiment.

(b) Explain the important features of one-half fraction of the 2^3 design. How will you construct such designs ?

V. (A) (a) What is a nested design ? Explain the statistical analysis of a two-stage nested design.

(b) What is meant by design resolutions ? Give examples to each of Resolutions III, IV and V designs based on two level factorial designs.

Or

(B) (a) Explain the second order rotatable design and its uses. Describe a method for construction of such a design.

(b) Write down the covariance model with LSD layout involving one concomitant variable. Briefly describe the analysis with regard to testing for the homogeneity of treatment effects.

($4 \times 16 = 64$ marks)

C 83553 (Pages : 3) Name...

Reg. No..................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2020

(CCSS)

Statistics

STA 2C 07—STATISTICAL INFERENCE–I

(2013 Admissions)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (i) Define minimal sufficiency. Let $X \sim U(\theta, \theta+1)$ obtain minimal sufficient statistic for θ.

(ii) Define unbiased estimator. Let $X \sim P(\lambda)$. Show that $\frac{1}{\lambda}$ is not estimable.

(iii) Define consistent estimator. Let $X \sim U(0,\theta)$. Show that $X_{(n)} = \max(X_1, \ldots X_n)$ is consistent. Is it unbiased ?

(iv) Let X follows exponential with mean θ. Derive CAN estimator for θ^2.

(v) Let X follows Cauchy distribution with location parameter θ. Derive percentile estimator of θ. Is it unbiased ?

(vi) Let $X \sim N(\mu, \sigma^2)$. Find MLE's of μ, σ^2 and $P(X \leq x)$.

(vii) Let $X_1, \ldots, X_n$ be a random sample from exponential with mean θ. Obtain unbiased confidence interval for θ based on the pivot $\frac{\sum_1^n \times i}{\theta}$.

(viii) Explain the method of construction of Bayesian credible intervals.

(4 × 4 = 16 marks)

Turn over

Section B

Answer **either** *part A* **or** *part B of all questions.*
Each question carries 16 *marks.*

II. A (i) Define (i) ancillary ; and (ii) completeness. State and prove Basu's theorem.

(ii) Let $X_1, X_2, \ldots, X_n$ be i.i.d. exponential with mean θ. Find UMVUE of θ^r and $P(X_1 \le x)$.

Or

B (i) Let $X_1, X_2, \ldots, X_n$ be a random sample from the density $f(x, \theta, \beta) = \frac{1}{\beta} e^{\frac{-(x-\theta)}{\beta}}$ $x > \theta, \beta > 0$.

Show that $\left(X_{(1)}, \sum_{j=1}^{k}\left(X_j - X_{(1)}\right)\right)$ is jointly sufficient and complete for (θ, β). Also find UMVUE's of θ and β.

(ii) Define Fisher information matrix. Let $X \sim N(\mu, \sigma^2)$. Obtain Fisher information matrix of (μ, σ^2).

III. A (i) Derive a sufficient condition for an estimator to be consistent. Let $X \sim U(0, \theta)$. Show that $X_{(n)} = \max_i X_i$ is consistent.

(ii) Distinguish between marginal consistency and Joint consistency. Prove or disprove Joint consistency implies marginal consistency.

Or

B (i) Define CAN estimator. Let $X \sim N(\mu, \sigma^2)$. Derive CAN estimators of μ^2 and σ^2.

(ii) Let X follows inverse Gaussian distribution with parameters λ and μ. Find moment estimators of λ and μ. Show that they are consistent.

IV. A (i) Under certain regularity conditions to be stated, show that MLE's asymptotically normal.

(ii) Let $X \sim P(\lambda)$. Assume that prior distribution of λ is gamma with parameters α and β. Derive Bayes estimator of λ under squared error loss. Find its Bayes risk.

Or

B (i) Find MLE's of parameter p_i, $i = 1, 2, ..., k$ of multinomial distribution.

(ii) Define (i) Bayes risk ; (ii) Bayes estimator. Let $X_1, X_2, ..., X_n$ be i.i.d. $U(0, \theta)$ random variables. The prior distribution of θ is $\Pi(\theta) = \frac{\alpha a^{\alpha}}{\beta^{\alpha+1}}; \theta \geq a$. Find Bayes estimator of θ under squared error loss.

V. A (i) What do you mean by large sample confidence intervals. Let $X \sim P(\lambda)$. Obtain large sample confidence interval for λ. Using the property of M.L.E.

(ii) Define unbiased confidence interval. Let $X_1, X_2, ..., X_n$ be i.i.d. from $U(0, \theta)$. Obtain unbiased confidence interval of θ based on the pivot $\frac{X_{(n)}}{\theta}$.

Or

B (i) Let $X \sim N(\mu_1, \sigma^2)$ and $Y \sim N(\mu_2, \sigma^2)$. Suppose that X and Y are independent. Based on random samples from two populations, derive $100(1-\alpha)\%$ shortest length confidence interval for $\mu_1 - \mu_2$. Show that this is unbiased.

(ii) Let $X_1, X_2, ..., X_n$ be a sample from geometric distribution with parameter θ. Assume that θ has prior density beta with parameters α and β. Find a Bayesian confidence interval for λ.

(4 × 16 = 64 marks)

C 83552 (Pages : 3) Name..................................

Reg. No..................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2020

(CCSS)

Statistics

STA 2C 06—PROBABILITY THEORY–II

(2013 Admissions)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 marks.

I. (a) Define the characteristic function (c.f) of a distribution function. Show that the c.f. of a random variable is non-negative definite.

(b) If $\phi(u)$ is a characteristic function then prove that $\mathrm{Re}(1-\phi(u)) \geq \frac{1}{4}\mathrm{Re}(1-\phi(2u))$.

(c) Define : (i) weak convergence ; and (ii) complete convergence. Show that complete convergence implies weak convergence.

(d) State and prove Levy's continuity theorem.

(e) Show that expectation of a conditional expectation is the unconditional expectation.

(f) Check if the strong law of large numbers hold for a sequence of independent random variables with probability mass function : $P(X_n = 0) = 1 - \frac{1}{n^2}$ and $P(X_n = n) = P(X_n = -n) = \frac{1}{2n^2}, n \geq 1$.

(g) Check whether the central limit theorem holds for the sequence of independent r.vs. with PMF : $P(X_n = n) = P(X_n = -n) = \frac{1}{2n^{2.5}}, P(X_n = 0) = 1 - \frac{1}{n^{2.5}}, n = 1, 2, \ldots\ldots$

(h) When do you say that a distribution function is infinitely divisible (id) ? Show that an exponential distribution is id.

(4 × 4 = 16 marks)

Turn over

Section B

Answer **either** *Part A* **or** *Part B of all questions.*
Each question carries 16 marks.

II. A. (a) Prove that the characteristic function and the distribution uniquely determine each other.

(b) Obtain the characteristic function of the following distribution function :

$$F(x) = \begin{cases} \frac{1}{4} + \frac{3}{4}\left(1 - e^{-\mu x}\right) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}.$$

(10 + 6 = 16 marks)

Or

B. (a) Establish Scheffe's theorem for sequence of distribution functions involving densities.

(b) Let $\{X_n\}$ be a sequence of i.i.d. random varibles with common density function $\int(\cdot)$ and the characteristic function : $\phi(u) = \begin{cases} 1-|u| & \text{if } |u| \leq 1 \\ 0 & \text{if } |u| > 1 \end{cases}$. Check if the sequence of distribution functions of $\left\{Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i, n = 1,2,.....\right\}$ converges weakly. If it does then find the limit distribution.

(8 + 8 = 16 marks)

III. A. (a) State and establish Kolmogoro's inequality.

(b) Let $\{X_n\}$ be a sequence of almost sure bounded random variables and $\sigma_n^2 = \text{var}(X_n)$. Show that $\sum_n (X_n - E(X_n))$ converges almost surely if and only if $\sum \sigma_n^2 < \infty$.

(10 + 6 = 16 marks)

Or

B. (a) Let $\{X_n\}$ be a sequence of i.i.d. r.vs with common d.f. F and $S_n = X_1 + X_2 + ... + X_n$. Then prove that there exists a sequence $\{\mu_n\}$ such that $\frac{S_n}{n} - \mu_n \to 0$ in probability if $\lim .nP(|X_1| > n) = 0$ as $n \to \infty$.

(b) Let $\{X_n\}$ be a sequence of pairwise uncorrelated random variables with $E(X_i) = \mu_i$ and $\text{var}(X_i) = \sigma_i^2, i = 1,2,....$ such that $\sum_{i=1}^{n} \sigma_i^2 \to \infty$ as $n \to \infty$. Show that $\{X_n\}$ obeys weak law of large numbers.

IV. A. (a) State and prove Lindeberg-Levy central limit theorem for i.i.d. random variables with finite variance.

(b) Check whether the central limit theorem holds for the sequence of independent r.vs. with

PMF : $P(X_n = n) = P(X_n = -n) = \frac{1}{2\sqrt{n}}, P(X_n = 0) = 1 - \frac{1}{\sqrt{n}}, n = 1, 2, ...$

(c) Do you agree with the statement "CLT is a generalization of the law of large numbers" ? Justify your argument.

(8 + 4 + 4 = 16 marks)

Or

IV. B. (a) State Lindeberg-Feller CLT. What are its advantages compared to other central limit theorems ?

(b) Let $\{X_n\}$ be a sequence of independent random variables with $E(X_n) = 0$ and $var(X_n) = \sigma_n^2 > 0$. Let $S_n = \sum_{i=1}^{n} X_i$ and $S_n^2 = \sum_{i=1}^{n} \sigma_i^2 < \infty$. If $\frac{\sum_{i=1}^{n} E|X_i|^{2+\delta}}{s_n^{2+\delta}} \to 0$ for some $\delta, 0 < \delta < 1$ then show that $\frac{S_n}{s_n}$ converges in distribution to standard normal random variable as $n \to \infty$. Imposing the required condition prove the converse part also.

(6 + 10 = 16 marks)

V. A. (a) Define a (i) martingale, (ii) submartingale and give an example each. Prove or disprove : If $\{X_n\}$ is a martingale then $\{X_n^2\}$ is a sub-martingale.

(b) Prove that weak limit of a sequence of infinitely divisible distribution functions whenever exists is again infinitely divisible.

(8 + 8 = 16 marks)

Or

B. (a) Define conditional expectation with respect to a sigma field and establish the smoothing property.

(b) Show that the characteristic function of an infinitely divisible distribution function never vanishes.

[4 × 16 = 64 marks]

C 4772 (Pages : 2) Name.....................................

Reg. No.....................................

SECOND SEMESTER M.Sc. DEGREE (REGULAR/SUPPLEMENTARY) EXAMINATION, APRIL 2021

(CBCSS)

Statistics

MST 2C 09—TESTING OF STATISTICAL HYPOTHESIS

(2019 Admissions)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Weightage 2 for each question.

1. Show that there does not exist UMP tests for two sided alternative hypotheses in testing the value of the parameter of distributions belonging to one-parameter exponential family.
2. Show that if a sufficient statistics exists for the family, the Neymann Pearson most powerful test is a function of it
3. Describe LMP tests.
4. Define similar region tests. How they are obtained ?
5. For an SPRT with bounds (A, B) and strength (α, β), obtain the inequalities connecting these.
6. Show that likelihood ratio tests are asymptotically unbiased.
7. Derive the formula for Spearman's rank correlation co-efficient.

(4 × 2 = 8 weightage)

Part B

Answer any **four** *questions.*
Weightage 3 for each question.

8. State and prove Neymann Pearson Lemma.
9. Obtain UMPU test for testing $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ in $N(\mu, \sigma^2)$ where σ is known.

Turn over

10. What is the significance of MLR property in testing of hypotheses ? Examine whether discrete uniform distribution possesses this property.

11. State and prove a set of sufficient conditions for a similar test to have Neymann structure.

12. Derive SPRT of strength (α,β) for testing the parameter of Poisson distribution.

13. Compare Kolmogorov-Smirnov and Chi-square for goodness of fittests.

14. Give a suitable non parametric test for testing the equality of medians of two independent populations.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Weightage 5 *for each question.*

15. Derive the expression for sample size required to achieve a test with size α and power $1-\beta$ for testing $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ in $N(\mu, \sigma^2), \sigma^2$ is known. Also give the expression for the sample size when it is one-sided test.

16. Derive the likelihood ratio test statistic for testing the equality of variances of two independent normal populations with unknown parameters.

17. State and prove Wald's fundamental identity.

18. Describe Mann-Whitney U-test.

(2 × 5 = 10 weightage)

C 4771 (Pages : 2) Name.....................................

Reg. No.................................

SECOND SEMESTER M.Sc. DEGREE (REGULAR/SUPPLEMENTARY) EXAMINATION, APRIL 2021

(CBCSS)

Statistics

MST 2C 08—SAMPLING THEORY

(2019 Admissions)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Each question carries weightage 2.

1. Explain the terms : (i) Sampling unit ; and (ii) Sampling frame.
2. Explain stratified sampling.
3. Define circular systematic sampling.
4. Show that the ratio estimators are biased.
5. Distinguish between multistage sampling and multiphase sampling.
6. Explain any one method of selecting a PPS sample without replacement.
7. Distinguish between ordered and unordered estimators.

($4 \times 2 = 8$ weightage)

Part B

Answer any **four** *questions.*
Each question carries weightage 3.

8. Explain the advantages of sampling over census.
9. In case of stratified random sampling, explain Neyman allocation.

Turn over

10. In a linear systematic sampling for $N = nk$, show that the mean of a systematic sample is an unbiased estimator of the population mean.

11. Compare the efficiency of the regression estimator with those based on mean per unit and ratio estimation procedure.

12. In a single stage cluster sampling with clusters of equal size, obtain an unbiased estimator of the population mean and obtain its variance.

13. Explain Hartly - Ross unbiased ratio type estimator.

14. Define Horwitz - Thompson estimator under PPSWOR scheme, for the population total. Is it unbiased ? Obtain the expression for its variance.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each question carries weightage 5.

15. Define simple random sampling. Suggest an unbiased estimator of the population mean in SRS WR. Derive the expression for variance of the estimator.

16. With usual notations, prove that in stratified sampling :

$$V_{opt} \leq V_{prop} \leq V_{ran}.$$

17. Carry out a comparison between systematic sampling, stratified sampling and simple random sampling in the case of a population with linear trend.

18. Define Desraj's ordered estimator for the population total. Using a sample of size two show that it is unbiased. Derive the variance of the estimator.

(2 × 5 = 10 weightage)

C 4770 (Pages : 2) Name..............................

Reg. No..............................

SECOND SEMESTER M.Sc. DEGREE (REGULAR/SUPPLEMENTARY) EXAMINATION, APRIL 2021

(CBCSS)

Statistics

MST 2C 07—ESTIMATION THEORY

(2019 Admissions)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Each question carries weightage 2.

1. Define complete sufficient statistic.
2. State Rao Blackwell theorem.
3. Distinguish between pivot and ancillary statistic.
4. Prove that maximum likelihood estimators need not be unique.
5. Let $X_1, X_2, \ldots, X_n$ be a random sample from $U(0,\theta)$, prove that $X_{(n)}$ is a consistent estimator for θ.
6. Let X_1 and X_2 be a random sample from Poisson (λ). Show that the statistic $T = X_1 + 2X_2$ is not sunicient for λ.
7. Define one parameter exponential family. Examine whether the Poisson family is a member of this class.

(4 × 2 = 8 weightage)

Part B

Answer any **four** *questions.*
Each question carries weightage 3.

8. Obtain minimal sufficient statistic for the parameter of a Poisson distribution based on a sample of size *n* from the Poisson distribution.
9. State and prove Fisher Neyman factorization theorem.

Turn over

10. State and prove the invariance property of consistent estimators.

11. Explain the method of moments for estimation of parameters. Estimate the parameters m and p in the case of Gamma distribution with mean p/m by the method of moments.

12. Let $X_1, X_2,...,X_n$ be a sample from a population with p.d.f.

$f(x,\theta) = \theta(1-\theta)^x, x = 0,1,2,....\ 0<\theta<1,$ find UMVUE of θ.

13. Explain shortest length confidence interval. Let $X_1, X_2,...,X_n$ be a sample from $N(\mu,\sigma^2), \sigma^2$ unknown. Find shortest length confidence interval for μ.

14. Let $X_1, X_2,...,X_n$ be random sample from Binomial distribution with parameters n and θ. If the prior distribution of θ, is uniform (0, 1) Find the posterior distribution of θ.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each question carries weightage 5.

15. a) State and prove Cramer Rao inequality. Show that if the distribution belongs to one parameter exponential family, minimum variance bound estimator exists.

b) Find the lower bound for the variance of unbiased estimator for θ based on a random sample of size n from Cauchy distribution.

16. (a) Describe the pivotal quantity method of constructing confidence interval..

(b) Find $100\ (1-\alpha)\%$ confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$ if random samples of sizes n_1 and n_2 are taken from $N(\mu_1,\sigma_1^2)$ and $N(\mu_2,\sigma_2^2)$ respectively.

17. Let $X \sim N(\mu,\sigma^2)$ and let the prior p.d.f. of μ be $N(\theta,\tau^2)$. Find the Bayes estimate of μ using quadratic error loss.

18. (a) State and prove Lehman Scheffes theorem.

(b) Let $X_1, X_2,...,X_n$ be a random sample from $N(\theta,\theta^2)$. Show that $T = \left(\sum X_i, \sum X_i^2\right)$ is sufficient for θ but not complete.

(2 × 5 = 10 weightage)

C 4769 (Pages : 2) Name.......................................

Reg. No..................................

SECOND SEMESTER M.Sc. DEGREE (REGULAR/SUPPLEMENTARY) EXAMINATION, APRIL 2021

(CBCSS)

Statistics

MST 2C 06—DESIGN AND ANALYSIS OF EXPERIMENTS

(2019 Admissions)

Time : Three Hours Maximum : 30 Weightage

General Instructions

1. *In cases where choices are provided, students can attend* **all** *questions in each section.*
2. *The minimum number of questions to be attended from the Section / Part shall remain the same.*
3. *There will be an overall ceiling for each Section / Part that is equivalent to the maximum weightage of the Section / Part.*

Part A

Answer any **four** *questions.*
Each question carries weight 2.

1. Explain how the local control principle is incorporated in a randomized block design.
2. Explain the features of a fixed-effect model in comparing treatments.
3. Eight treatments are to be compared using 14 blocks each of size 4. Identify the parameters of the balanced incomplete block design.
4. Define a partially balanced incomplete block design with m-associate classes.
5. Give the Yates procedure for getting the treatment totals of a 2^3 factorial experiment.
6. Define total confounding, and give an example.
7. Define rotatability blocking.

(4 × 2 = 8 weightage)

Part B

Answer any **four** *questions.*
Each question carries weight 3.

8. Outline the analysis of a latin square design for comparing treatments.
9. Explain how a missing value is estimated in a randomized block design.

Turn over

10. Show that $\lambda(v-1) = r(k-1)$ for a balanced incomplete block design.

11. Explain the method of computing the adjusted treatment sum of squares in a balanced incomplete block design.

12. Explain the concept of resolution of a block design with an example.

13. Explain the nature of a strip-plot design, and give its use.

14. Explain the method of block orthogonality in response surface designing.

(4 × 3 = 12 weightage)

Part C

Answer any **two** *questions.*
Each question carries weight 5.

15. Give the ANCOVA procedure for comparing treatment effects using completely randomized design.

16. Give the inter-block analysis of a balanced incomplete block design.

17. Construct a 2^3 design in which ABC confounded in two replications and BC in one. Also, give the ANOVA and comment on the differences in the information on the treatments.

18. Explain the methodology of second-order response surface designs.

(2 × 5 = 10 weightage)

C 4007 (Pages : 2) Name.....................................

Reg. No.................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2021

(CCSS)

Statistics

STA 2C 09—REGRESSION METHODS

(2019 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (i) Explain the method of Generalized Least Squares.

(ii) Describe simultaneous interval estimation.

(iii) Explain Likelihood Ratio Test.

(iv) Define R^2 and adjusted R^2. What is the advantage of adjusted R^2 over R^2 ?

(v) What are Dummy Explanatory Variables ? How it is used in Straight Line Regression ?

(vi) Define Orthogonal Polynomials.

(vii) Discuss various Variance stabilizing Transformation.

(viii) Explain regression surface.

(4 × 4 = 16 marks)

Section B

Answer **either** *part-A* **or** *part-B of all questions.*
Each question carries 16 *marks.*

II. A) a) Find the Least Square Estimators of the linear regression model.

b) Prove that the least square estimates of the parameters of linear regression model are BLUE.

(8 + 8 = 16 marks)

Or

B) a) Given $Y_1, Y_2, \ldots, Y_n$ independently distributed as N (0, σ^2). Prove that $\frac{Q}{\sigma^2} \sim \chi^2(n-1)$,

where $Q = \Sigma i\left(Y_i - \bar{Y}\right)2$.

Turn over

b) Given that Y_1, Y_2, Y_3 are random variables with means $\beta_1+\beta_2, \beta_1+\beta_3$ and $\beta_3+\beta_2$ and a common variance . Then, show that $l_1\beta_1+l_2\beta_2+l_3\beta_3$ is estimable if $l_1=l_2+l_3$.

(8 + 8 = 16 marks)

III. A) a) Discuss various Variance stabilizing Transformations.

b) Briefly explain prediction intervals and band for the response.

(8 + 8 = 16 marks)

Or

B) a) Describe the analysis of GLM.

b) Explain the logistic regression model and its uses.

(8 + 8 = 16 marks)

IV. A) a) Explain Weighted Least Square Method. State the Advantages of it.

b) Describe two-phase Linear Regression.

(8 + 8 = 16 marks)

Or

B) a) Describe polynomial regression in one variable.

b) Write a note on piecewise polynomial fitting and spline functions.

(8 + 8 = 16 marks)

V. A) a) Describe the methods of diagnostic checking in regression analysis.

b) Write a short note on residual plots in Regression model. Describe how you use residuals for checking the adequacy of the model fitted.

(8 + 8 = 16 marks)

Or

B) a) Describe various variance stabilizing transformations.

b) Distinguish between ridge-regression and principle component regression.

(8 + 8 = 16 marks)

C 4006 (Pages : 2) Name.....................................

Reg. No.....................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2021

(CCSS)

Statistics

STA 2C 08—DESIGN AND ANALYSIS OF EXPERIMENTS

(2019 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 marks.

1. (i) Distinguish between Fixed effect model and Random effect model in ANOVA.

   (ii) Give a brief note on Regression approach in ANOVA.

   (iii) Introduce the Model for Randomized Block Design. Explain why RBD is much better than CRD.

   (iv) Briefly explain the concept of recovering of intra block information in BIBD.

   (v) Write a note on two - level fractional design and its applications.

   (vi) Explain the analysis of first order factorial design.

   (vii) What are the hierarchical designs ?

   (viii) Explain the use of response surface design.

(4 × 4 = 16 marks)

Section B

Answer either part-A **or** *part-B of* **all** *questions.*
Each question carries 16 marks.

II. A a) Write a short note on optimality criteria for Experimental Design.

b) Explain the Comparison of Individual treatment means. (8 + 8 = 16 marks)

Or

B a) What is Analysis of Variance ? What are its importance ?

b) Write down the model for One way ANOVA and give its analysis. (8 + 8 = 16 marks)

Turn over

III. A a) Evaluate the Efficiency of a Latin Square Design to a completely Randomized Design.

b) What is Greaco-Latin Square Design ? (8 + 8 = 16 marks)

Or

B. a) What are the differences between BIBD and PBIBD ?

b) Describe the intra block and inter block analysis of BIBD.

(8 + 8 = 16 marks)

IV. A a) Explain the analysis of 3^2 factorial Design.

b) Explain the Yate's procedure of obtaining various effects and sum of squares of a 2^k factorial experiments. (8 + 8 = 16 marks)

Or

B. a) Explain Confounding. What are its advantages.

b) Write down the confounded arrangement of a 3^3 design by confounding the Interactions ABC and BC^2 into blocks. Identify other interactions of any which get confounded in our arrangement. Hence analyse the design if there are two replications of the same type of arrangement of the treatment. (8 + 8 = 16 marks)

V A a) What are the Nested Designs ?

b) Explain the analysis of second order surface design and mention its applications.

(8 + 8 = 16 marks)

Or

B. a) List out the advantages and applications of ANCOVA.

b) Develop the analysis of CRD with one concomitant variable. (8 + 8 = 16 marks)

C 4005 (Pages : 2) Name.....................................

Reg. No..................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2021

(CCSS)

Statistics

STA 2C 07—STATISTICAL INFERENCE–I

(2019 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions ; each question carries 4 marks.*

I. (i) Define a sufficient statistics. Let X_1, X_2 follows B (1, *p*. If X_1 and X_2 are independent, examine whether $X_1 + 7X_2$ is sufficient for *p*.

(ii) Define consistency of an estimator. Is a consistent estimator unique, justify your answer.

(iii) Define ancillary statistics. Discuss one example.

(iv) Define one parameter exponential family. Give an example.

(v) Define a CAN estimator. Explain with the help of an example.

(vi) Define Bhattacharya bound.

(vii) Describe percentile method of finding a consistent estimator.

(viii) What is meant by shortest confidence interval ?

(4 × 4 = 16 marks)

Section B

Answer either part A **or** *part B of* **all** *questions ; each question carries 16 marks.*

II. A. a) State and prove Neyman Factorization theorem for discrete case.

b) State and prove Lehman-Scheffe theorem. Find UMVUE of θ based on a sample of size *n* from $U(0, \theta)$.

Or

B. a) State and prove Basu's theorem.

b) Describe a method of finding minimal sufficient statistics. Give an example.

Turn over

III. A. a) Let $X_1, X_2, ... X_n$ be a sample from $U(\theta - 1\backslash 2, \theta + 1\backslash 2), \theta \varepsilon R$. Show that $(\text{Min } X_i, \text{Max } X_i)$ is sufficient for θ but not complete.

b) Find a consistent estimator of the parameter θ based on a sample of size n from Cauchy population with location parameter θ.

Or

B a) Let $X_1, X_2, ... X_n$ be a sample a sample $N(\theta, \theta^2)$. Obtain a sufficient statistics for θ.

b) Let $X_1, X_2, ... X_n$ be a sample a sample Poisson population with parameter, find the UMVUE of $P(X < 1)$.

IV. A. a) Explain conjugate family of prior distribution. If X follows Binomial (n, p) find conjugate family for p. Also find the Bayes estimator of p under squared error loss function.

b) Describe Bayes estimation procedure.

Or

B. a) State and prove invariant property of CAN estimator.

b) Find CAN estimators of μ and σ based on a sample of size n from $N(\mu, \sigma^2)$.

V. A. a) Show that a consistent solution of the likelihood equation is asymptotically normal and is asymptotically efficient.

b) Describe pivotal quantity method of constructing confidence intervals.

Or

B. a) State and prove Cramer-Rao inequality. Give an example where Cramer-Rao lower bound is attained and another where it is not attained.

b) Obtain the shortest confidence interval for σ^2 based on a random sample of size n from $N(0, \sigma^2)$.

(4 × 16 = 64 marks)

C 4004 (Pages : 2) Name..................................

Reg. No..............................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2021

(CCSS)

Statistics

STA 2C 06—PROBABILITY THEORY—II

(2019 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Part A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. 1 Obtain the characteristic function of Cauchy distribution.

2 Show by an example that weak convergence need not imply complete convergence.

3 Define independence of two classes of events. Hence using it define independence of two random variables.

4 State and prove Khintchine's WLLN.

5 Let $\{X_n\}$ be a sequence of independent random variables with $P\{X_n = \pm 2^k\} = 1/2$. Examine Whether SLLN hold.

6 State Lindeberg-Feller form of CLT.

7 Let $\{X_n\}$ be a sequence of independent random variables with $P\{X_k = \pm n^{-a}\} = p$, $P\{X_n = 0\} = 1 - 2p$, $(0 < p < \frac{1}{2})$. For which value of a, $\{X_n\}$ obeys central limit theorem.

8 Define infinite divisibility. Examine the infinite divisibility of Poisson distribution.

(4 × 4 = 16 marks)

Part B

Answer **either** *part A* **or** *part B.*
Each question carries 16 *marks.*

II. (A) (i) State and prove Hellys convergence theorem.

(ii) State and prove Helly Bray lemma.

Or

(B) (i) Show that characteristic function is real iff the distribution function is symmetric about zero.

(ii) State and prove inversion theorem on characteristic functions.

Turn over

III. (A) (i) State and prove Borel Cantelli lemma. Is the converse is true. Justify your answer.

(ii) State and prove Kolmogorov three series theorem.

Or

(B) (i) Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with common law N (0.1). Find the limiting distribution of the random variable

$$W_n = \sqrt{n}\,\frac{X_1 + X_2 + \ldots + X_n}{X_1^2 + X_2^2 + \ldots + X_n^2}.$$

(ii) Define tail event and tail sigma field. Prove that probability of the tail events of a sequence of independent random variable is either 0 or 1.

IV. (A) (i) A Establish Kolmogorov strong law of large numbers for a sequence of independent random variables.

(ii) For the sequence $\{X_n\}$ of independent random variables with $P\{X_k = \pm k\} = \frac{1}{2\sqrt{k}}$ and

$P\{X_k = 0\} = 1 \ldots \frac{1}{\sqrt{k}}$ does the SLLN hold.

Or

(B) (i) State and prove Liapunov CLT.

(ii) Show that Liapunov condition implies Lindberg-Feller condition.

V. (A) (i) Define conditional expectation and state its important properties.

(ii) Define Martingales. Show that if $\{X_n\}$ is a sequence of independent random variables with $E(Y_n) = 0$ then $Y_n = \sum_{k=1}^{n} X_k$ is a martingale sequence.

Or

(B) (i) Define the concept of infinite divisibility of random variables. Show that Poisson (λ) distribution and Negative-binomial (r, p) distribution are infinitely divisible.

(ii) Show that if a sequence (X_m) of infinitely divisible random variables converges in distribution to X, then X is infinitely divisible.

(4 × 16 = 64 marks)

C 4003 **(Pages : 3)** **Name.....................................**

Reg. No.................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2021

(CCSS)

Statistics

STA 2C 09—REGRESSION METHODS

(2013 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 marks.

I. (a) In a Gauss-Markov model show BLUE of a linear parametric function is unique.

(b) Let $Y_1 = \theta + \varepsilon_1, Y_2 = 2\theta - \phi + \varepsilon_2$ and $Y_3 = \theta + 2\phi + \varepsilon_3$, where $E(\varepsilon_i) = 0, i = 1,2,3,$ Find the least squares estimate of θ and ϕ.

(c) For the straight line model $Y = \beta_0 + \beta_1 x_i + \varepsilon_i; i = 1,2,...n,$ obtain the confidence intervals for the slope.

(d) What are Bonferroni t-intervals ? Explain.

(e) Derive a $100(1-\alpha)\%$ simultaneous prediction interval for Y_0, in the model $Y = X\beta + \varepsilon$ at the point $X = X_0$.

(f) What are spline functions ? Explain.

(g) Explain the role of graphical plots of residuals against the fitted values in linear regression models.

(h) What is principal component regression ? Explain.

(4 × 4 = 16 marks)

Section B

Answer **either** *A* **or** *B of all questions.*
Each question carries 16 marks.

II. A. (a) Explain multiple linear regression model $Y = X\beta + \varepsilon$ and estimate the parameters of it using the method of least squares. Show that the maximum likelihood estimate coincides with the least squares estimates.

Turn over

(b) If $E(Y) = X\beta$, where X is $n \times p$ of rank $r(r \le p)$, and $var(Y) = \sigma^2 I_n$ then prove that

$$S^2 = \frac{(Y-\hat{\theta})'(Y-\hat{\theta})}{n-r}, \text{ where } \hat{\theta} = PY \text{ with } P = X(X'X)^{-1}X'.$$

Or

B. (a) Let $Y = X\beta + \varepsilon$, where X is $n \times p$ of rank p. Discuss the method of Lagrange multipliers to find the minimum of $\varepsilon'\varepsilon$ subject to the linear restrictions $A\beta = c$, where A is a known $q \times p$ of rank q, and c is a known $q \times 1$ vector.

(b) Define generalized least squares estimator and show that it has minimum variance in the class of all unbiased estimators whenever the error variance is non-constant.

III. A. (a) For the general linear model $Y = X\beta + \varepsilon$ and to test hypothesis $H : A\beta = C$, prove that

(i) $RSS_H - RSS = (A\hat{\beta} - c)'\left[A(X'X)^{-1}A'\right]^{-1}(A\hat{\beta} - c)$.

(ii) $E(RSS_H - RSS) = \sigma^2 q + (A\beta - c)'\left[A(X'X)^{-1}A'\right]^{-1}(A\beta - c)$.

(b) Define multiple correlation coefficients. Show that the multiple correlation coefficient

$$R^2 = 1 - \frac{RSS}{\sum (Y_1 - \bar{Y})^2}.$$

Or

B. (a) Let $Y_1 = \alpha_1 + \varepsilon_1, Y_2 = 2\alpha_1 - \alpha_2 + \varepsilon_2, Y_3 = \alpha_1 + 2\alpha_2 + \varepsilon_3$, where $\varepsilon \sim N_n(X\beta, \sigma^2 I_n)$. Derive an F-statistic for testing $H : \alpha_1 = \alpha_2$.

(b) Discuss how prediction intervals and bands are constructed for the response.

IV. A. (a) Explain the concept of orthogonal polynomials. How does the assumption of orthogonality simplify the problem of least square estimation in polynomial regression ?

(b) Explain polynomial regression model in several variables.

Or

B. (a) Discuss the estimation of the parameters of the orthogonal polynomial model. Derive the test statistic for testing the significance of the parameters.

(b) Describe the weighted least squares technique for the straight line model with known weights.

V. A. (a) Distinguish between bias due to under fitting and bias due to over-fitting in a multiple regression model, giving an illustrative example.

(b) What are outliers ? How do you detect them ? Describe the treatment of outliers.

Or

B. (a) Explain how to detect non-constant variance and non-normality in the context of regression analysis. Describe the methods of overcome these problems.

(b) What are serial correlations in multiple linear regression models and explain the different sources of it ? How will you detect the presence of it ?

(4 × 4 = 16 marks)

[4 × 16 = 64 marks]

C 4002 (Pages : 2) Name.................................

Reg. No..............................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2021

(CCSS)

Statistics

STA 2C 08—DESIGN AND ANALYSIS OF EXPERIMENTS

(2013 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (a) Distinguish between fixed effect and random effect models.

(b) What do you mean by interaction? How will you determine if any interaction component is necessary in the model ?

(c) What is CRD? Under what situation CRD can be efficiently applied. Give a layout of CRD.

(d) Explain Youden Square design.

(e) What are the advantages of factorial designs ? Give an example.

(f) Distinguish between main effects and interaction effects for a factorial design.

Explain how these effects are measured for a 2^2 design.

(g) What are response surfaces ? Cite some of its applications.

(h) Define a concomitant variable. Explain its relevance in the analysis of covariance.

(4 × 4 = 16 marks)

Section B

Answer **either** *A* **or** *B of all questions.*
Each question carries 16 *marks.*

II. (A) (a) Explain various principles of experimentation and their role in the design of experiments.

(b) Explain one-way classified data. Write down the model, stating clearly the assumptions that are made and prepare its analysis of variance.

Or

(B) (a) State and prove Gauss-Markov theorem.

(b) Explain how Fisher Least Significant Difference (LSD) method and Duncan's multiple range test are used to compare pairs of treatment means.

Turn over

III. (A) (a) Discuss the layout and analysis of RBD.

(b) Distinguish between LSD and Graeco-Latin Square designs. Explain the analysis of variance of a general Graeco-Latin Square design.

Or

(B) (a) Prove the following parametric relationships.

(i) $vr = bk$. (ii) $\lambda(v-1) = r(k-1)$.

(iii) $b \geq v$.

(b) Describe the intra block analysis of BIBD.

IV. (A) (a) For a 2^3 factorial experiment laid as a RBD with r replicates, explain the analysis of variance.

(b) Express the main effects and interactions in terms of treatment means for a 2^4 factorial experiment.

Or

(B) (a) Distinguish between total and partial confounding. Explain total confounding with reference to a 2^3 factorial experiment and write down its analysis.

(b) Explain the important features of one-half fraction of the 2^3 design. How will you construct such designs ?

V. (A) (a) Describe a method of construction of linear response surface designs.

(b) When will you say that a response design is rotatable ? Explain the role of central composite designs to fit the second order response surface designs.

Or

(B) (a) Explain the method of steepest ascent employed in response surface designs.

(b) Consider a RBD with v treatments in r blocks and a covariate x. Writing a suitable linear model derive a test for equality of treatment effects.

($4 \times 16 = 64$ marks)

C 4001 (Pages : 3) Name..................................

Reg. No..................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2021

(CCSS)

Statistics

STA 2C 07—STATISTICAL INFERENCE – I

(2013 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 marks.

I. (i) Obtain Fisher information of θ when X follows $N(\theta,1)$.

(ii) Show that the family $U(0,\theta)$ is complete.

(iii) Prove *or* disprove : Consistent estimators are always unbiased.

(iv) Distinguish between joint and marginal consistent estimators.

(v) Find M.L.E. of the parameters in the following :

(a) $f(x;\theta,\beta)=\frac{1}{\beta}e^{-\frac{1}{\beta}(x-\theta)}$; $x>\theta$.

(b) $f(x;\theta)=\frac{1}{2\theta}$; $-\theta<x<\theta$.

(vi) Define (i) risk function ; and (ii) Bayes risk. Find Bayes estimator θ when $X\sim P(\theta)$ under squared error loss. The prior distribution of θ is $G(\alpha,\beta)$.

(vii) Let X follows exponential with mean θ. Obtain large sample confidence interval of θ using the property of M.L.E.

(viii) Explain Fiducial intervals. Illustrate with an example.

(4 × 4 = 16 marks)

Turn over

Section B

Answer **either** *Part A* **or** *Part B of all questions.*
Each question carries 16 *marks.*

II. A (i) Define minimal sufficiency. Let $X_1, X_2, \ldots, X_n$ be a sample from the density :

$$f(x,\theta) = \begin{cases} \frac{x}{\theta} e^{-x^2/2\theta} & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$$

Show that $\sum_1^n X_i^2$ is minimal sufficient for θ ; but $\sum_1^n X_i$ is not sufficient.

(ii) State and prove Rao-Blackwell theorem. Discuss its merit and demerits.

Or

B (i) Define UMVUE. Let $X_1, X_2, \ldots, X_n$ be a random sample from $p(\lambda)$. Find the UMVUE of

(a) $g(\lambda) = \frac{1}{1-\lambda}$; (b) $g(\lambda) = \lambda^s$; $s > 0$; (c) $g(\lambda) = p(X = 0)$.

(ii) State Cramer-Rao lower bound for the variance. When does the equality hold ? Let $X \sim N(\theta, 1)$. Show that the minimum variance of any estimator of θ^2 is $\frac{4\theta^2}{n}$.

III. A (i) Define consistent estimators. State and prove invariance property of the estimators.

(ii) Let X follows lognormal distribution with parameters μ and σ^2. Obtain consistent estimators of μ and σ^2. Are they unbiased ?

Or

B (i) Let $X \sim P(\theta)$. Obtain CAN estimator for θ^2 and find its asymptotic variance.

(ii) Explain method percentiles for estimation. Show that the estimators so derived are consistent.

IV. A (i) Define M.L.E. show that M.L.E.'s are consistent under certain regularity conditions.

(ii) Let $X \sim b(n, \theta)$ and suppose that the prior density of θ is U (0, 1). Find the Bayes estimator of θ, using loss function :

$$L(\theta, T) = \frac{(\theta - T)^2}{\theta(1-\theta)}.$$

Or

B (i) Find M.L.E. of parameters in the following cases :

(a) $X \sim N(\theta, \theta^2)$.

(b) $X \sim 1G(\mu, \lambda)$. (inverse Gaussian)

(c) $f(x ; \theta, \beta) = \frac{1}{2\beta} e^{-\frac{|x-\theta|}{\beta}} \quad -\infty < x < \infty.$

(ii) Define Bayes estimator. Show that posterior mean is the Bayes estimator under squared error loss function.

V. A (i) Explain pivotal method for finding confidence intervals. Illustrate it with an example.

(ii) Define unbiased confidence intervals. Let X follows $X \sim N(\mu, \sigma^2)$. Obtain $100(1-\alpha)\%$ unbiased confidence interval for σ^2, based on a sample of size.

Or

B (i) Let $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ X and Y are independent. Obtain $100(1-\alpha)\%$ shortest confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$, based on random samples from each population.

(ii) Let $X_1, X_2, \ldots, X_n$ be a sample from a Poisson distribution with parameter λ. The prior distribution of λ is $G(\alpha, \beta)$. Find $100(1-\alpha)\%$ Bayesian confidence interval for λ.

(4 × 16 = 64 marks)

C 4000 (Pages : 4) Name..

Reg. No.....................................

SECOND SEMESTER P.G. DEGREE EXAMINATION, APRIL 2021

(CCSS)

Statistics

STA 2C 06—PROBABILITY THEORY—II

(2013 Admission onwards)

Time : Three Hours Maximum : 80 Marks

Section A

Answer any **four** *questions.*
Each question carries 4 *marks.*

I. (a) Define the characteristic function of a distribution function. Obtain the density function corresponding to the characteristic function : $\phi(u) = e^{-|u|}, -\infty < u < \infty$.

(b) If $\{\phi_n(u)\}$ is a sequence of characteristic function (C.F.)s then show that $\sum_{n=1}^{M} p_n \phi_n(u)$ is also a C.F. if $p_n \geq 0$ and $\sum_{n=1}^{M} p_n = 1$.

(c) Define (i) Weak convergence ; and (ii) Complete convergence. Illustrate with an example that weak convergence need not imply complete convergence.

(d) Define a martingale sequence. Show that a sequence of partial sums of zero mean independent random variable is a martingale.

(e) Show that strong law of large numbers always implies the weak law of large numbers. Is the converse true ?

(f) Check if the weak law of large numbers hold for a sequence of independent random variables with probability mass function :

$$P(X_n = 0) = 1 - \frac{1}{n^3} \text{ and } P(X_n = n) = P(X_n = -n) = \frac{1}{2n^3}, n \geq 1.$$

Turn over

(g) Prove or disprove : Poisson distribution is infinitely divisible.

(h) State Lindeberg-Feller form of central limit theorem. Comment on the supremacy of this theorem over the other central limit theorem.

(4 × 4 = 16 marks)

Section B

Answer **either** *Part A* or *Part B of* **all** *questions.*
Each question carries 16 *marks.*

II. A (a) State Bochner's theorem and show that the characteristic function is always non-negative definite.

(b) Let F be a distribution function and $\{v_n, n \geq 1\}$ be a sequence of absolute moments of F.

Show that if $\sum_{n=1}^{\infty} \frac{v_n u^r}{n!}$ converges for some $u_0 > 0$, then the sequence of row moments $\{\mu'_n\}$ determines the distribution F uniquely.

(6 + 10 = 16 marks)

Or

B (a) Suppose that a sequence of probability density functions (pdf) $\{f_n(x)\}$ corresponding to the distribution functions $\{F_n(x)\}$ converging to a pdf $f(x)$ of the df $F(x)$ for almost all x. Then show that $F_n(x)$ converge completely to $F(x)$.

(b) Let $\{X_n\}$ be a sequence of i.i.d. random varibles with common density function $f(\cdot)$ and the characteristic function : $\phi(u) = \begin{cases} 1 - |u| & \text{if } |u| \leq 1 \\ 0 & \text{if } |u| > 1 \end{cases}$. Check if the sequence of distribution functions of $\left\{Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i, n = 1, 2, \ldots\ldots\right\}$ converges weakly. If it does then find the limit distribution.

(8 + 8 = 16 marks)

III. A (a) If $\{X_n\}$ is a sequence of independent and identicaly distributed random variables with characteristic function $\phi(u)$ then show that $(X_1 + X_2 + \ldots\ldots + X_n)/n$, $n = 1, 2, \ldots.$ converges in *probability* to μ if and only if $\phi'(0) = i\mu$.

(b) State and prove Kolmogorov's strong law of large numbers.

(6 + 10 = 16 marks)

Or

B (a) Let $\{X_n\}$ be a sequence of i.i.d. r.v.s. with common d.f. F and $S_n = X_1 + X_2 + \ldots\ldots + X_n$.

Then prove that there exists a sequence $\{\mu_n\}$ such that $\frac{S_n}{n} - \mu_n \to 0$ in *probability* if $\lim . nP(|X_1| > n) = 0$ as $n \to \infty$.

(b) Let $\{X_n\}$ be a sequence of independent random variables with probability mass functions $P(X_n = n^{3/2}) = P(X_n = -n^{3/2}) = \frac{1}{2n^3}$, $P(X_n = 0) = 1 - \frac{1}{n^3}$, $n = 1, 2, \ldots$ Check if strong law of large numbers hold.

(10 + 6 = 16 marks)

IV. A. (a) State and prove Lyapunov's central limit theorem.

(b) Show that Lyapunov's condition implies Lindberg condition.

(c) Check whether the central limit theorem holds for the sequence of independent r.v.s. with PMF : $P(X_n = n) = P(X_n = -n) = \frac{1}{2\sqrt{n}}$, $P(X_n = 0) = 1 - \frac{1}{\sqrt{n}}$, $n = 1, 2, \ldots$

(8 + 4 + 4 = 16 marks)

Or

B (a) Let $\{X_n\}$ be a sequence of i.i.d. r.v.s. with $E(X_n) = 0$ and $Var(X_n) = \sigma^2 < \infty$. If $\bar{X}_n = \sum_{i=1}^{n} X_i/n$ then show that $\frac{\sqrt{n}\,\bar{X}_n}{\sigma}$ converges in distribution to standard normal random variable as $n \to \infty$.

(b) Check whether the central limit theorem holds for the sequence of independent r.v.s. with

PMF : $(X_n = n) = P(X_n = -n) = \frac{1}{2\sqrt{n}}, P(X_n = 0) = 1 - \frac{1}{\sqrt{n}}, n = 1, 2, \ldots$

(c) Do you agree with the statement "CLT is a generalization of the law of large numbers" ? Justify your arguement.

(8 + 4 + 4 = 16 marks)

V. A (a) Let X and Y be two random variables such that $E(XY)$ and $E(X)$ exist. If X is $\mathscr{D}$ – measurable then prove that $E(XY|\mathscr{D}) = X\,E(Y|\mathscr{D})$ a.s.

(b) Prove that weak limit of a sequence of infinitely divisible distribution functions whenever exists is again infinitely divisible.

(8 + 8 = 16 marks)

Or

B (a) State and prove linearity property of conditional expectations.

(b) Prove that $\{X_n, \mathscr{D}_m\}$ is a martingale if and only of for every $n > 1$ $E(X_n \mid \mathscr{D}_{n-1}) = X_{n-1}$ a.s.

(8 + 8 = 16 marks)

[4 × 16 = 64 marks]