# CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, May 2018

**Programme : M. Sc / M.A**
**Semester : II**
**Course Title : Applied Statistics**
**Course Code : SPMS STAT 01 02 01 GEC 3104**

**Session: 2016-2017**
*Max. Time : **3 Hours***
*Max. Marks : **50***

Question no. 1 has five sub parts and students need to answer any four. Each sub part carries two 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 five marks.

**Question No. 1.**

i) Explain process control and product control

ii) Define an abridge life table.

iii) Explain standardized death rate and Infant mortality rate.

iv) Producer's risk and consumer's risk

v) Explain seasonal variation and cyclical variation.

**Question No. 2**

**a)** What is a time series? What are its main components? Give illustration for each of them.

**b)** Explain clearly what is meant by trend of a time series? What are the methods of measuring the trend? Explain each of them with suitable example.

**c)** Find the trend of annual sales in million rupees of a trading organization by moving average method

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 |
| Sales (Crores Rs.) | 80 | 84 | 80 | 88 | 98 | 92 | 84 | 88 | 80 | 100 |
| Year | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 |
| Sales (Crores Rs.) | 84 | 96 | 92 | 104 | 116 | 112 | 102 | 114 | 108 | 126 |

**Question No. 3**

a) What do you understand by seasonal fluctuations in a time series? Give examples. Explain the link relative of computing the indices of seasonal variation.

b) Compute the seasonal indices by the link relative method for adjoining data relating to the average quarterly prices (Rs. Per kg) of a commodity for five years:

| Year | I Quarter | II Quarter | III Quarter | IV Quarter |
|---|---|---|---|---|
| 1996 | 30 | 26 | 22 | 36 |
| 1997 | 35 | 28 | 22 | 36 |
| 1998 | 31 | 29 | 28 | 32 |
| 1999 | 31 | 31 | 25 | 35 |
| 2000 | 34 | 36 | 26 | 33 |

c) Describe briefly the various steps in the construction of control charts variables and obtain their control limits. Given below the values of sample mean and the range for ten sample of size 5. Draw the appropriate mean chart and comment on the state of control of the process.

| Sample No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Mean | 43 | 49 | 37 | 44 | 45 | 37 | 51 | 46 | 43 | 47 |
| R | 5 | 6 | 5 | 7 | 7 | 4 | 8 | 6 | 4 | 6 |

**Question No. 4**

a) Define the term vital statistics. Describe their nature and the methods of collection of vital statistics. Describe the direct and indirect method of standardizing death rate.

b) Compute the crude and standardized death rate of the two populations A and B, regarding A as standard population, from the data given below:

| Age-group (Years) | A | | B | |
|---|---|---|---|---|
| | Population | Death | Population | Death |
| Under 10 | 20,000 | 600 | 12,000 | 372 |
| 10-20 | 12,000 | 240 | 30,000 | 660 |
| 20-40 | 50,000 | 1250 | 62,000 | 1612 |
| 40-60 | 30,000 | 1050 | 15,000 | 525 |
| Above 60 | 10,000 | 500 | 3,000 | 180 |

c) Explain the main control chart for attributes and obtain their control limits. Discuss the advantages and disadvantages of control chart of variable and control charts of attributes. The following are the figures of defectives in 22 lots each containing 2000 rubber belts:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 425 | 430 | 216 | 341 | 225 | 322 | 280 | 306 | 337 | 305 | 356 |
| 402 | 218 | 264 | 126 | 409 | 193 | 326 | 280 | 389 | 451 | 420 |

Draw control chart for fraction defective and comment on the state of control of the process.

**Question No.**

a) Describe the various components of life table. How is the expectation of life at birth determined from life table?

b) Show that $e_x = \frac{1}{l_x}\left(\sum_{n=1}^{\infty} l_{x+n}\right)$, where symbols have their usual meaning.

c) What do mean by fertility of a population? Explain Crude birth rate, general fertility rate and total fertility rate. Calculate the general fertility rate, total fertility rate and the gross reproduction rate from the following data:

| Age of women | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 |
|---|---|---|---|---|---|---|---|
| No. of women ('000) | 16.0 | 16.4 | 15.8 | 15.2 | 14.8 | 15.0 | 14.5 |
| Total births | 260 | 2244 | 1894 | 1320 | 916 | 280 | 145 |

## CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, May/June 2018

**Programme: M.Sc (Statistics)**
**Semester: II**
**Course Title: Linear Models and Regression Analysis**
**Course Code: SPMS STAT 01 02 03 CC 4004**

**Session: 2017-18**
**Max. Time: 3 Hours**
**Max. Marks: 70**

**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 Gauss-Markov set-up along with basic assumptions.

**b.** What are Normal equations? Write normal equations in estimating the parameters in a two variable linear regression model.

**c.** What are residuals? Write its properties.

**d.** Find the maximum likelihood estimate of error variance $\sigma^2$ in simple linear regression model.

**e.** What are influential observations? Differentiate between an outlier and a leverage point.

**f.** What are Normal Probability Plots? Give some realizations of residual plots for non normal errors in a regression model.

**g.** What is a generalized linear model (GLM)? Write the components of GLM.

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

**a.** Develop tests of significance for testing hypotheses for the intercept and slope parameters in simple linear regression.

**b.** Obtain the least square estimator of the regression coefficient in $y = X\beta + u$ if the errors $u$ in the model follows $N(0, \sigma^2)$.

**c.** Consider the model $E(y_1) = \beta_1 + \beta_2, E(y_2) = \beta_1 , E(y_3) = \beta_1 - \beta_2$ with errors following $N(0, \sigma^2)$. Find the estimates of $\beta_1$ and $\beta_2$ and the Residual Sum of Squares (RSS).

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

a. In multiple linear regression model $= X\beta + u$ , develop a test statistic for testing the general linear hypothesis $H_0 : R\beta = r$ against $H_1 : R\beta \neq r$.

b. What are the exact linear restrictions? Obtain restricted regression estimator of coefficient vector $\beta$ in the multiple linear regression model $y = X\beta + u$, where $u \sim N(0, \sigma^2)$.

c. Explain what is meant by coefficient of determination $R^2$ and adjusted $R^2$ ? Establish the relation between $R^2$ and $F$ statistic in a general linear regression model.

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

a. What is PRESS statistic? Obtain its expression. Prove that PRESS is weighted sum of squares of residuals, with weights related to the leverage of the observations.

b. What is R- Student? Discuss the outlier test based on R-student in detail.

c. Explain the variance stabilizing transformations to correct model inadequacies.

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

a. What do you mean by polynomial in one variable? What are the considerations in fitting polynomial in one variable?

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

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

# CENTRAL UNIVERSITY OF HARYANA

**Term End Examinations, May/June 2018**

**(REAPPEAR)**

**Programme: M.Sc. Statistics** | **Session: 2017-18**

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

**Course Title: Linear Models and Regression Analysis** | **Max. Marks: 70**

**Course Code: SPMS STAT 01 02 03 CC 4004**

---

**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. What is a linear manifold?

b. Explain Gauss-Markov set-up along with basic assumptions.

c. Find the unbiased estimator of $\sigma^2$ in terms of $R_0^2$.

d. Define the ANOVA problem in simple linear regression model.

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

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

g. Discuss the usefulness of logistic regression models.

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

a. Derive the normal equations for Gauss-Markoff model. Also, prove that all models can be reduced to Gauss-MArkoff model.

b. Prove the following result:

(i) $\text{var}(\hat{\beta}_i) = \sigma^2 C_{ii}$; $i = 1,2,\dots,m$, where C is generalized inverse of matrix $X'X$.

(ii) $R_0^2 = \underline{Y}'A\underline{Y}$; where $A = I - X(X'X)^{-}X'$.

c. When a linear parametric function $P'\underline{\beta}$ is estimable? Also prove that $P'\hat{\underline{\beta}}$ is MVUE of

$P'\underline{\beta}$ in class f all linear unbiased estimates when $\hat{\underline{\beta}}$ is solution of normal equations.

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

**a.** State and prove First Fundamental Theorem of Least Squares.

**b.** Discuss the analysis of one-way classified data.

**c.** Explain in detail the two-way classified data for equal but multiple observations per cell.

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

**a.** Discuss confidence interval estimation in multiple linear regression models $y = \beta X + \varepsilon$.

**b.** Discuss the outlier test based on R-student in detail.

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

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

**a.** Discuss the estimation problem in logistic regression using maximum likelihood estimation.

**b.** Explain the problem of homogeneity of variances and hence discuss the Bartlett test in detail.

**c.** Discuss the analysis of covariance for two-way model with one concomitant variable.

# CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, May 2018

**Programme : M.Sc. Statistics**
**Semester : Second**
**Course Title : Distribution Theory**
**Course Code : SPMS STAT 01 02 01 CC 4004**

**Session: 2017-2018**
*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)

**i)** Obtain the skewness and kurtosis of a log-normal distribution.

**ii)** Define central and non central $t$ and $\chi^2$ distributions.

**iii)** Explain the bivariate normal distribution.

**iv)** Define mixture distribution.

**v)** Arithmetic mean and standard deviation of a binomial distribution are respectively 4 and $\sqrt{8/3}$. Find the values of n and p.

**vi)** 2% of the items made by a machine are defective. Find the probability that 3 or more items are defective in a sample of 100 items.

**vii)** Explain compound and truncated distributions.

**Question No. 2.**

**i)** Given $f(x) = \lambda e^{-\lambda x}$ obtain the distribution of $e^{-\lambda x}$.

**ii)** Prove the recurrence relation between the moments of Poisson distribution

$\mu_r = \lambda\left(r\mu_{r-1} + \frac{d\mu_r}{d\lambda}\right)$, where $\mu_r$ is the $r^{th}$ moments about mean $\lambda$. Also show that $\beta_2 - \beta_1 - 3 = 0$.

**iii)** For a trinomial distribution show that the correlation coefficient between $X$ and $Y$ is

$$\rho = -\sqrt{\left(\frac{p_1}{1-p_1}\right)\left(\frac{p_2}{1-p_2}\right)}.$$

**Question No. 3.**

**i)** For a Laplace distribution, show that $E(X-\mu)^{2r} = \sigma^{2r}\sqrt{(2r+1)}$.

**ii)** If $m$ things are distributed among $a$ men and $b$ women, show that the probability that the no. of things received by men is odd is $\frac{1}{2}\left[\frac{(b+a)^m - (b-a)^n}{(b+a)^m}\right]$.

iii) For a normal distribution $N(\mu,\sigma^2)$, obtain $E|X-\mu|$.

**Question No. 4.**

i) For the $t-$ distribution with $k$ degree of freedom, prove that

$$E[X^r]=\frac{k^{r/2}\Gamma\left(\frac{k-r}{2}\right)\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{k}{2}\right)},$$

and hence show that $\beta_2-\beta_1-3=\frac{6}{k-4}$, where symbols have their usual meanings.

ii) Show that for t-distribution with $k$ degrees of freedom

$$P[X\le x]=1-\frac{1}{2}I_{\frac{k}{k+x^2}}\left(\frac{k}{2},\frac{1}{2}\right),$$

where

$$I_x(a,b)=\frac{1}{B(a,b)}\int_0^x y^{a-1}(1-y)^{b-1}dy$$

is the incomplete beta function.

iii) Define the non-central chi square distribution and obtain its mean and variance.

**Question No. 5.**

i) A box contains $N$ identical ball numbered $1,2,\ldots,N$ of these $n$ balls drawn at a time. Let $X_1,X_2,\ldots,X_n$ denote the numbers on the $n$ ball drawn. Find $V(S_n)$, where $S_n=\sum_{i=0}^{n}X_i$.

ii) Obtain the moment generating function of bivariate normal distribution.

iii) Show that if $X_1$ and $X_2$ are standard normal variates with correlation coefficient $\rho$ between them, then the correlation coefficient between $X_1^2$ and $X_2^2$ is given by $\rho^2$.

# CENTRAL UNIVERSITY OF HARYANA

Term End Examinations May/June 2018

Programme: M.Sc. (Statistics) Session: 2017-18
Semester: II Max. Time: 3 Hours
Course Tilte: Distribution Theory Max. Marks: 70
Course Code: SPMS STAT 01 02 01 CC 4004

**Instructions:**

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

Question No. 1. (4 × 3.5 = 14)

(a) Let $X \sim \text{Bin}(n,p)$. Derive the moment generating function of $X$ and hence find $E(X)$ and $\text{Var}(X)$.

(b) If $X$ has a geometric distribution with probability mass function $P(X = x) = p(1-p)^x$, $x = 0,1,2,...$; $0 < p < 1$, then prove that

$$P(X > m + n | X > m) = P(X \geq n),$$

for any two nonnegative integers $m$ and $n$.

(c) Let $X_1$ and $X_2$ be i.i.d. random variables having exponential distribution with mean $\frac{1}{3}$. Find $P(X_1 + X_2 > 1)$.

(d) Let $X \sim N(\mu,\sigma^2)$ and let $Y = \ln X$. Find the value of $E(Y)$.

(e) Discuss, in brief, compound distribution with an example describing its application inreal life.

(f) Discuss, in brief, mixture distribution with an example describing its application in reallife.

(g) Let $(X,Y)$ have joint probability density function

$$f_{X,Y}(x,y) = \frac{1}{6\pi\sqrt{7}} \exp\left[-\frac{8}{7}\left(\frac{x^2}{16} - \frac{31}{32}x + \frac{xy}{8} + \frac{y^2}{9} - \frac{4}{3}y + \frac{71}{16}\right)\right],$$

for $-\infty < x < \infty$, $-\infty < y < \infty$. Find the variances of $X$ and $Y$. Also find the correlation coefficient between $X$ and $Y$.

Question No. 2. (2 × 7 = 14)

(a) Define characteristic function of a random variable. Let $X_1,...,X_n$ be independent random variables and let $Y = \sum_{k=1}^{n} X_k$. Prove that the characteristic function of $Y$ is the product of characteristic functions of $X_k$'s. Hence show that the sum of $n$ independent Poisson random variables with parameters $\lambda_1,...,\lambda_n$ is also a Poisson random variable with parameter $\lambda = \sum_{k=1}^{n} \lambda_k$.

(b) Write down the probability mass function of random variable $X$ following the negative binomial distribution. Derive its moment generating function and probability generating function. Hence, or otherwise, derive its mean and variance.

(c) The probability mass function of hypergeometric random variable $X$ is

$$P(X = x) = \binom{N}{n}^{-1}\binom{M}{x}\binom{N-M}{n-x}, \quad \max(0, M+n-N) \leq x \leq \min(M,n)$$

Show that

$$E(X) = \frac{n}{N}M \quad \text{and} \quad \text{Var}(X) = \frac{nM}{N^2(N-1)}(N-M)(N-n)$$

Question No. 3. (2 × 7 = 14)

(a) Let the random variable $X$ have a uniform distribution on the interval $(a,b)$, $-\infty < a < b < \infty$. Write down its probability density function. Derive its cumulative distribution function, moment generating function. Hence, or otherwise, derive its mean and variance.

(b) Let $X$ be distributed with probability density function

$$f(x) = \begin{cases} \frac{1}{12}x^2(1-x), & \text{if } 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Identify the distribution of $X$ and derive the value of $E(X^n)$ for $n \in \mathbb{N}$. Hence, find $E(X)$ and $Var(X)$.

(c) Let $X \sim N(\mu,\sigma^2)$. Derive the expression for charateristic function of $X$.

Question No. 4. (2 × 7 = 14)

(a) Let $X$ be a random variable with probability mass function

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, \ldots; \lambda > 0$$

Suppose that the value $x = 0$ cannot be observed. Find the probability mass function of the truncated random variable, its mean, and its variance.

(b) Let the random variable $X$ follow noncentral chi-square distribution with probability density function

$$f_X(x) = e^{-\nu^2/2} \sum_{s=0}^{\infty} \frac{(\nu^2/2)^s}{s!} \frac{e^{-x/2}x^{(k/2)+s-1}}{\Gamma(\frac{k}{2}+s)2^{(k/2)+s}}, \quad x > 0$$

Show that the moment generating function of $X$ is given by

$$M_X(t) = (1-2t)^{-k/2}e^{\nu^2 t/(1-2t)}.$$

(c) Let $X$ and $Y$ have the joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} c\,x\,y\,e^{-(x^2+2y^2)}, & \text{if } x > 0, y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Evaluate the constant $c$ and show that $X$ and $Y$ are independent random variables.

Question No. 5. (2 × 7 = 14)

(a) Let $(X,Y)$ be a bivariate normal random vector with parameters $\mu_1$, $\mu_2$, $\sigma_1^2$, $\sigma_2^2$, and $\rho$. Derive the expression for the conditional expectation of $X$, given that $Y = y$.

(b) Let $(X,Y)$ be a bivariate normal random vector with parameters $\mu_1$, $\mu_2$, $\sigma_1^2$, $\sigma_2^2$, and $\rho$. Find a necessary and sufficient condition for $X + Y$ and $X - Y$ to be independent.

(c) Let the $p$-variate random vector **X** be distributed according to multivariate normal distribution with mean vector $\mu$ and variance-covariance matrix $\Sigma$. Prove that **Y** = $C$**X** is distributed according to $N_p(C\mu, C\Sigma C^0)$ for $C$ nonsingular.

# CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, May/June 2018

**Programme: M.Sc (Statistics)** **Session: 2017-18**
**Semester: IV** **Max. Time: 3 Hours**
**Course Title: Bayesian Inference** **Max. Marks: 70**
**Course Code: SPMS STAT 01 04 02 CC4004**

**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) A purse contains 5 quarters and one of them is double-headed. A coin is chosen at random and tossed 5 times. Every time a head turns up. What is the probability that it is the quarter with two heads?
b) State and prove the Bayes' theorem.
c) Explain the following terms with suitable example (i) uniform prior and (ii) noninformative prior.
d) What is 0-1 loss function? Give a situation where it is applicable.
e) Write short notes on Bayes rule and Bayes actions.
f) Differentiate between frequentist and Bayesian inference approach.
g) Describe the credible and highest posterior density (HPD) interval.

Question No. 2. (2X7=14)

a) Derive the asymptotically locally invariant prior for exponential family of distributions.
b) Explain approaches for the subjective determination of prior density with appropriate example.
c) Obtain Bayes estimators of $\mu$ and $\sigma^2$ under symmetric loss function in the lognormal probability density function $f(x|\mu,\sigma)=\frac{x^{-1}}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2\sigma^2}(\log x-\mu)^2\right]\infty<\mu<\infty,\sigma,x>0$ when the joint prior $g(\mu,\sigma)\propto\frac{1}{\sigma^c},\ c>0.$

Question No. 3. (2X7=14)

a) Show that under absolute error loss, the Bayes rule is the median of the posterior distribution.
b) Let $\underline{x}=(x_1,x_2,...,x_n)$ be a random sample from the exponential probability density function $f(x|\mu,\theta)=\frac{1}{\theta}\exp\left(-\frac{x-\mu}{\theta}\right),\ 0<\mu<x,\theta>0$. Obtain $100(1-\alpha)\%$ credible and highest posterior density (HPD) intervals for $\mu$ and $\theta$ under the vague-prior $g(\mu,\theta)\propto\frac{1}{\theta},\ \theta>0.$

c) If a sufficient statistic exists for the parameter $p$ of binomial distribution $B(n, p)$, then show that the family of conjugate priors exists for $p$.

Question No. 4. (2X7=14)

a) For given $\underset{\sim}{x}=(x_1, x_2, ..., x_n)$ and the sufficient statistic $T=\sum_{i=1}^{n} x_i$, construct the natural conjugate prior for the Poisson distribution, $f(x \mid \lambda)=\frac{\exp(-\lambda)\lambda^x}{x!}$, $\lambda > 0$, $x=0,1,2,....$ Hence, obtain Bayes estimator $\lambda^*$. When does the Bayes estimator $\lambda^*$ identify itself with the corresponding maximum likelihood estimator $\hat{\lambda}$?

b) Discuss the location invariant and hierarchical prior. Write objectives of a hierarchical prior.

c) Define the empirical Bayes estimator. Suppose $X$ : $B(n,\theta), 0<\theta<1, x=0,1,2,...,n$ and $\theta$ has a $Be(\alpha,\beta)$ prior distribution. Find marginal distribution $m(x)$ and show that, if $m(x)$ is constant, then it must be the case that $\alpha=\beta=1$.

Question No. 5. (2X7=14)

a) Explain the prior and posterior analysis of exponential distribution. Find the Bayes estimator of the parameter in exponential distribution.

b) Describe the method of obtaining predictive distribution and interval for Rayleigh probability density function $f(x \mid \sigma)=\frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right)$, $x, \sigma > 0$ with the prior $g(\sigma) \propto \frac{1}{\sigma^3}$, $\sigma > 0$.

c) Show that noninformative priors are scale invariant.

**CENTRAL UNIVERSITY OF HARYANA**

Term End Examinations, Nov/Dec 2017

| | | |
|---|---|---|
| **Programme:** | **M.Sc. Statistics** | |
| **Session:** | **2018-19** | |
| **Semester:** | **I** | **Max. Time: 3 Hours** |
| **Course Title:** | **Statistical Methods** | **Max. Marks: 70** |
| **Course Code:** | **SPMS STAT 01 01 02 CC 4004** | |

**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) Find the variance of residual in a simple linear regression model.

b) Explain Chi-square test of goodness of fit.

c) Define Null hypothesis, Critical region and Type I error and Type II error

d) Explain correlation coefficient and linear regression

e) Define mean deviation and coefficient of variation

f) Define empirical distribution function and discuss its important properties

g) Explain Binomial distribution

Question No. 2. (2X7=14)

a) Express first four central moments in terms of moments about origin. What do you mean by skewness and kurtosis of a distribution? How it is measured? Distinguish clearly, by giving figures, between positive and negative skewness. Also show the relative positions of mean, median and mode in the figures, for positively and negatively skewed distribution.

b) What is the problem of association of attributes? Difine the different measures of association in a 2X2 contingency table.

c) Explain with suitable examples the term disperson. State the relative and absolute measures of dispersion and describe the merits and demerits of standard deviation. Also define moments.

Question No. 3. (2X7=14)

a) Define Spearman's rank correlation. Prove that Spearman's rank correlation coefficient is given by $1-\frac{6\sum d_i^2}{n(n^2-1)}$, where $d_i$ denote the difference between the ranks of $i^{th}$ individual.

b) What is a time series? What are its main components? Give illustration for each of them.

c) Define Normal distribution. Show that the probability distribution of the random variables fulfills all the basic properties of the probability density function. Also find their mean and variance.

Question No. 4. (2X7=14)

a) Find the $(1-\alpha)100\%$ confidence intervals for difference of means (variance unknown) and ratio of variance (mean unknown) in case of normal population.

b) Explain chi-square test. A firm manufacturing rivet want to limit variations in their length as much as possible. The lengths (in cm) of 10 rivets manufactured by a new process are 2.15 1.99 2.05 2.12 2.17 2.01 1.98 2.03 2.25 1.93. In the past, the variation in length of rivets manufactured by the firm has been 0.021 cm, examine whether the new process seems to be superior to the old.

c) Define order statistics and derive the probability distribution of the rth order statistic. Also, obtain the joint distribution of two order statistics.

Question No. 5. (2X7=14)

a) Explain the exact and approximate testing procedure for correlation coefficient.

b) Explain the concept of two-way analysis of variance and its computational procedure under fixed effect model.

c) Define $t$-distribution and state its important properties. Suppose that the nicotine contents of two brands of cigarettes are being measured. If in an experiment, 50 cigarettes of the first brand had an average nicotine contents 2.61 mg, while 40 cigarettes of the second brand had a average nicotine contents 2.38 mg, test the claim that brand first, had more nicotine contents than the brand second at 5% level of significance. The population standard deviations are given 0.12 mg and 0.14 mg respectively.

## CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, May 2018

| | | |
|---|---|---|
| **Programme:** | **M.Sc. (Statistics)** | **Session: 2017-18** |
| **Semester:** | **IV** | **Max. Time: 3 Hours** |
| **Course Title:** | **Multivariate Analysis** | **Max. Marks: 70** |
| **Course Code:** | **SPMS STAT 01 04 01 CC 4004** | |

**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) If the variance-covariance matrix of a p-component random vector $\underline{X}$ is $\Sigma$, then obtain the variance-covariance matrix of the transformed vector $\underline{Y} = AX$, where $A$ is a $p \times p$ matrix of constant elements.

b) If $\underline{X}$ follows $N_p(\underline{\mu},\Sigma)$ distribution, then prove that $(\underline{X}-\underline{\mu})'\Sigma^{-1}(\underline{X}-\underline{\mu})$ follows a central chi-square distribution with $p$ degrees of freedom.

c) If $A \sim W_p(n,\Sigma)$, then prove that $DAD' \sim W_p(n, D\Sigma D')$, where $D$ is a $p \times p$ non-singular matrix of constant elements.

d) Define Hotelling's $T^2$-statistic and prove that it is invariant under a non-singular linear transformation.

e) Define canonical correlations and canonical variates. Prove that canonical correlations are invariant under a non-singular linear transformation.

f) Let $\underline{X} \sim N_3(\underline{0},\Sigma)$, where $\Sigma = \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix}$. Find $\rho$ such that $X_1 + X_2 + X_3$ and $X_1 - X_2 - X_3$ are independent.

g) Let $\underline{X}_\alpha$ (α=1,2,…,n) be a random sample of size $n$ from $N_p(\underline{\mu},\Sigma)$, where $\Sigma$ is known. Develop a suitable test for testing $H_0 : \underline{\mu} = \underline{\mu}_0$ where $\underline{\mu}_0$ is a given vector.

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

a) Show that Wishart distribution is a multivariate generalization of the chi-square distribution.

b) If $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_\alpha, \ldots, \underline{X}_n$ is a random sample of size $n(> p)$ from $N_p(\underline{\mu},\Sigma)$, then show that $\overline{\underline{X}} \sim N_p(\underline{\mu}, \Sigma/n)$.

c) If $A_i$ $(i = 1,2)$ are distributed independently according to $W_p(\upsilon_i, \Sigma)$ respectively, then show that $A_1 + A_2 \sim W_p(\upsilon_1 + \upsilon_2, \Sigma)$.

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

a) $X_1$ and $X_2$ are two random variables with respective expectations $\mu_1$ and $\mu_2$ and variances $\sigma_{11}$ and $\sigma_{22}$ and correlation coefficient $\rho$ such that

$$X_1 = \mu_1 + \sqrt{\sigma_{11}}\, Y_1, \text{ and } X_2 = \mu_2 + \sqrt{\sigma_{22}(1-\rho^2)}\, Y_1 + \sqrt{\sigma_{22}}\, Y_2,$$

where $Y_1$ and $Y_2$ are two independent normal $N(0,1)$ variates. Find the joint distribution of $X_1$ and $X_2$.

b) What do you mean by generalized variance? Obtain the distribution of sample generalized variance.

c) Define $p$-variate normal distribution and obtain its probability density function.

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

a) If $\underline{X}$ (with $p$ components) be distributed according to $N_p(\underline{\mu}, \Sigma)$, then prove that $\underline{Y} = C\underline{X}$ is distributed according to $N_p(C\mu, C\Sigma C')$ for $C$ nonsingular.

b) If $\underline{X} \sim N_p(\underline{0}, I)$ and if $A$ and $B$ are real symmetric matrices of order $p \times p$, then show that

i) $E(\underline{X}' A \underline{X}) = tr\, A$

ii) $V(\underline{X}' A \underline{X}) = 2\, tr\, A^2$

iii) $\text{Cov}(\underline{X}' A \underline{X}, \underline{X}' B \underline{X}) = 2\, tr\, AB$.

c) Define Fisher's discriminant function and show that it is invariant under a non-singular linear transformation. Also, obtain a test for testing the goodness of fit of a hypothetical discriminant function.

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

a) Let $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$. Show that $M_{\underline{X}-\underline{\mu}}(t) = \exp\left(\frac{1}{2}\underline{t}'\Sigma \underline{t}\right)$, hence find the moment generating function of $\underline{X}$.

b) Let $\underline{X}_\alpha$ ($\alpha$=1,2,...,n) be a random sample of size $n$ from $N_p(\underline{\mu}_0, \Sigma)$, where $\underline{\mu}_0$ is a given vector. Derive the maximum likelihood estimator of $\Sigma$.

c) A set of p variates following multivariate normal distribution is grouped into q sub-sets having $p_1, p_2, \ldots, p_q$ variates, respectively, such that $p_1 + p_2 + \ldots + p_q = p$. Derive the likelihood ratio criterion to test the hypothesis that these q groups are independent of one another and obtain the test-statistic for large samples.

# CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, May/June 2018

**Programme: M.Sc (Statistics)** **Session: 2017-18**
**Semester: IV** **Max. Time: 3 Hours**
**Course Title: Decision Theory and Sequential Analysis** **Max. Marks: 70**
**Course Code: SPMS STAT 01 04 03 DCEC 4004**

**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) Define risk function and likelihood principle.
b) Explain the need of randomization in decision problem.
c) Define a decision rule and R-better decision rule.
d) Describe Wald's sequential probability ratio test.
e) Let Z be a random variable with finite second moment. Show that $f(b) = E(Z - b)^2$ is minimized when $b = E(Z)$.
f) Show that every complete class is 'essentially complete'.
g) What are the basic elements of game theory? Explain a graphical method for solving a game.

Question No. 2. (2X7=14)

a) Explain the term convex set. Show that the intersection of any number of convex sets is convex, hence that the convex hull is convex.
b) If $p_0, p_1$ denote elements of $\wp^*$; $p_0 < p_1$ and $0 \le \lambda < \mu \le 1$, then prove that $\lambda p_1 + (1-\lambda)p_0 < \mu p_1 + (1-\mu)p_0$.
c) Let X has the pdf $f(x,\theta) = \begin{cases} \theta e^{-\theta x}; x \ge 0, \ \theta > 0 \\ 0, \quad elsewhere \end{cases}$. For testing $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$, construct the SPRT and obtain its ASN and OC functions.

Question No. 3. (2X7=14)

a) Prove that Bayes decisions are admissible and unique up to risk equivalence.
b) Explain the finite action and invariant decision problems.
c) The following is the loss matrix of a particular no-data problem

| | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ |
|---|---|---|---|---|---|
| $\theta_1$ | 1 | 0 | 2 | 1 | −1 |
| $\theta_2$ | 3 | 5 | 4 | 7 | 0 |
| $\theta_3$ | −2 | −3 | −2 | −1 | 1 |

(i) Obtain admissible and inadmissible decisions (actions)?

(ii) If $\theta$ has the following prior distribution $\pi(\theta_1)=\frac{1}{2}, \pi(\theta_2)=\frac{1}{4}, \pi(\theta_3)=\frac{1}{4}$, then find the Bayesian expected loss of each decision.

(iii) Find the Bayes decision of this no-data problem.

Question No. 4. (2X7=14)

a) Define admissible decision, complete class of decisions, essentially complete class of decisions, minimal complete class of decisions and minimal essentially complete class of decisions.

b) Solve the game whose pay-off matrix is $\begin{bmatrix} -1 & -2 & 8 \\ 7 & 5 & -1 \\ 6 & 0 & 12 \end{bmatrix}$.

c) Explain how the problems of point estimation and hypothesis testing fit in with the general theory of decisions?

Question No. 5. (2X7=14)

a) Describe the steps for construction of utility (U) function.

b) Let $X : B(1,\theta), \theta \in \Theta = \left\{\frac{1}{4}, \frac{1}{2}\right\}$ and $A = \{a_1, a_2\}$. The loss function is defined as follows:-

| | $a_1$ | $a_2$ |
|---|---|---|
| $p_1 = \frac{1}{4}$ | 1 | 4 |
| $p_2 = \frac{1}{2}$ | 3 | 2 |

The set of decision rules includes four functions; $d_1, d_2, d_3, d_4$ defined by $d_1(0) = d_1(1) = a_1$; $d_2(0) = a_1, d_2(1) = a_2$; $d_3(0) = a_2, d_3(1) = a_1$; $d_4(0) = d_4(1) = a_2$. Find the minimax rule.

c) Write a short note on the problem of classification as decision problem.