Total No. of Questions: 6 *Total No. of Printed Pages:3*

Enrollment No.....

Faculty of Engineering

End Sem (Even) Examination May-2022

CS3ET01 Statistics & Probability

Programme: B.Tech. Branch/Specialisation: CSE

Duration: 3 Hrs. **Maximum Marks: 60**

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d.

Q.1 i. According to … "Statistics is the science of estimates and probabilities." **1**

(a) Boddington (b) Bowley

(c) King (d) None of these

ii. A condition or characteristic that can take on different values or categories is called____. **1**

(a) Variable (b) Constant

(c) A cause-and-effect relationship (d) None of these

iii. Most stable measure of "measure of central tendency" is- **1**

(a) Mean (b) Mode (c) Median (d) None of these

iv. Which statement is true- **1**

(a) Mean = 3 Median -2 Mode (b) Median =3 Mode -2 Mean

(c) Mode = 3 Median -2 Mean (d) None of these

v. A Type I error can occur when the null hypothesis is- **1**

(a) Correct (b) Incorrect

(c) Either correct or incorrect (d) None of these

vi. If the significance level, α, is increased then the chance of a Type II error will- **1**

(a) Remains the same (b) Decrease

(c) Increase (d) None of these

vii. The value of $E(x - \bar{x})$ is- **1**

(a) 1 (b) 0 (c) $\bar{x}$ (d) None of these

viii. If *x* and *y* are two random variables, then $cov\ (x, y)$defined by- **1**

(a) $E(x - y)$ (b) $E(x + y)$

(c) $E(x)E(y)$ (d) None of these

P.T.O.

ix. The probability of getting an even number when throwing a die is- **1**

(a) 1 (b) 0 (c) $\frac{1}{2}$ (d) None of these

x. If E_1 and E_2 are mutually exclusive events, then- **1**

(a) $E_1 \cap E_2 = \emptyset$ (b) $E_1 \cap E_2 = 0$

(c) $E_1 \cap E_2 = 1$ (d) None of these

Attempt any two:

Q.2 i. Examine the important definitions of statistics, which is in your opinion is the best definition? Explain the scope and limitations of statistical methods. **5**

ii. Define Statistical data. Explain their characteristics. **5**

iii. Write a note on application of statistics in different research areas. **5**

Attempt any two:

Q.3 i. An incomplete frequency distribution is given below: **5**

Interval	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency	12	30	?	65	?	25	18

If median = 46, then find the missing frequencies.

ii. Find standard deviation for the following data: **5**

x	0	4	5	8	9	13
f	5	6	1	4	7	2

iii. Find lower and upper quartiles for the following distribution: **5**

Interval	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
Frequency	5	6	15	10	5	4	2	2

Attempt any two:

Q.4 i. Define Null and alternative hypothesis with example. Explain errors of first and second kinds. **5**

ii. In a population 10 persons are selected at random. Their heights (in inches) are as follows: **5**

63,63,64,65,66,69,69,70,70,71

Consider the proposition that the mean height in the population is 65 inches. Given that at 9 degree of freedom student t static at 5% level of significance is 2.262.

iii. Write a note on one way and two-way ANOVA. **5**

Attempt any two:

Q.5 i. Find the expected value of the number of points obtained by throwing a simple die once. Also find variance. **5**

ii. If moment generating function of a random variable x is- **5**

$\left(\frac{1}{3}+\frac{2}{3}e^t\right)^5$, then find $P(X=2)$.

iii. Show that the mean deviation about the mean of the Poisson distribution is $\frac{2}{e}$ times the standard deviation of the unit mean. **5**

Q.6 Attempt any two:

i. In a survey of 200 students of a school, it was found that 120 study Mathematics, 90 study Physics and 70 study Chemistry, 40 study Mathematics and Physics, 30 study Physics and Chemistry, 50 study Chemistry and Mathematics and 20 none of these subjects. Find the number of students who study all the three subjects **5**

ii. Two bags A and B contain 2 white and 5 black balls and 3 white and 6 black balls respectively. A ball is places from bag A into bag B and one ball is drawn from bag B. Find the probability that the ball is black. **5**

iii. When two dice are thrown into the air, find the probability that the sum of the digits on their upper face is 11, if 5 appear on the top of the first die. **5**

******

Medi-Caps University
Marking Scheme

Course name : Statistics and Probability
Programme : B.tech
Branch : CSE-DS (2nd year, Even)

Q.1 MCQs

i)	Boddington	+1
ii)	Variable	+1
iii)	Mean	+1
(iv)	Mode = 3 Median − 2 Mean	+1
v)	Correct	+1
vi)	decrease	+1
vii)	0	+1
viii)	none of these	+1
ix)	1/2	+1
x)	$E_1 \cap E_2 = \phi$	+1

Q.2(i) Croxton and Cowden defined Statistics more concisely as the science which deals with the collection, analysis and interpretation of numerical data. This definition covers all the aspects of Statistical methods. +1

Scopes of Statistics :—

1) Statistics and Industry — In Industry Statistics is widely used in quality control. In production engineering to find whether the product is conforming to specification or not. +1

Statistics and Business: Business executives are laying more and more on statistical techniques. Before starting the production process he must have an overall idea about the data need to enquired. +1

Limitations of Statistics:

- Study does not study individuals infact it deals with an aggregate of object
- Statistics is not suited to the study of qualitative phenomenon. +1
- Statistical laws are not exact, they are true only on average
- Statistics is liable to be misused. +1

1.2(ii) Statistical data are those numric value which represents the information of given items or object. Eg:- height, weight, marks etc. +1

Characteristics of Statistical data:

(i) Statistical data are aggregates of facts ie. Single and isolated quantity is not Statistics +1

(ii) Statistical data works on quantitative +1

information not on qualitative information

(iii) Data Collection should be done in systematic manner, disorderly arranged data will lead to wrong conclusion. +1

v) Statistical data are always collected when there is a predetermined purpose. +1

2(iii) Statistical analysis is used in Astrostatistics to understand to astronomical data. Statistical data is used to study of populations which is called Demography. +1

Geostatistics is a branch of geography that deals with the analysis of data from disciplines such as petroleum, hydrology, mineralogy etc. +1

A branch of applied mathematics called 'Operations research' use statistics and probability. +1

Statistical physics is one of the fundamental theories of physics, and uses methods of probability theory in solving physical problems. +1

Quantitative psychology is the science of statistically explaining and changing mental processes and behaviour in humans +1

Q.3 (i)

Interval	Frequency	CF
10-20	12	12
20-30	30	42
30-40	x	42+x
40-50	65	107+x
50-60	y	107+x+y
60-70	25	132+x+y
70-80	18	150+x+y

assume total frequency to be 200

Then $150+x+y = 200$

$\Rightarrow x+y = 50 \quad \leftarrow (1)$ +2

as Median = 46 $\Rightarrow$ 40-50 is Median class

$\therefore l = 40 \quad \frac{N}{2} = 100 \quad cf = 42+x \quad f = 65$

Median $= l + \left(\frac{N/2 - cf}{f}\right) \times h$ +1

$46 = 40 + \frac{(100 - 42 - x)}{65} \times 10$

$36 = \frac{(58-x)2}{13}$

$39 = 58 - x$

$x = 19$

$\therefore y = 50 - 19$

$y = 31$ +2

Q-3(i)

x	f	fx	$x-\bar{x}$	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
0	5	0	−6	36	180
4	6	24	−2	4	24
5	1	5	−1	1	1
8	4	32	2	4	16
9	7	63	3	9	63
13	2	26	7	49	98
					382

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{150}{25} = 6$$ +3

$$S.D(\sigma) = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{N}} = \sqrt{\frac{382}{25}} = 3.90$$ +2

Q-3(ii)

C.I	frequency	C.f
5-10	5	5
10-15	6	11
15-20	15	26
20-25	10	36
25-30	5	41
30-35	4	45
35-40	2	47
40-45	2	49

+1

$$Q_r = l_1 + \left(\frac{rN/4 - cf}{f}\right) \times h$$

for Q_1, $r=1$ i.e. $\frac{N}{4} = \frac{49}{4} = 12.25$

i.e. Q_1 will lie in 15 − 20 +1

$$Q_1 = 15 + \frac{12.25 - 11}{15} \times 5$$

$$\boxed{Q_1 = 15.41}$$ +1

for Q_3, $r = 3$ i.e. $\frac{3N}{4} = 3 \times 12.25 = 36.75$

i.e. Q_3 will lie in 25-30 +1

$$Q_3 = 25 + \frac{36.75 - 36}{5} \times 5$$

$$\boxed{Q_3 = 25.75}$$ +1

Q.4 (i) Null hypothesis : In testing of hypothesis, we always begin with an assumption or hypothesis (i.e. assumed value of population parameter). This is called Null hypothesis. It is denoted by H_0 +1

Alternate hypothesis : Any hypothesis which contradicts the null hypothesis, is called Alternate hypothesis. It is denoted by H_1 +1

For eg: if we have to test whether the population mean μ has a specified value μ_0, then

Null hypothesis is $H_0 : \mu = \mu_0$ &

Alternate hypothesis, $H_1 : \mu \neq \mu_0 \ (\mu > \mu_0, \mu < \mu_0)$ +1

Type-I Error : The ~~error~~ error of rejecting null hypothesis (H_0) when it is true is called Type-I error +1

Type-II Error : The error of accepting null hypothesis (H_0), when it is false, is called Type-II Error +1

(Q-4(ii)) Given, population mean = 65

$$\bar{x} = \frac{63+63+64+65+66+69+69+70+71}{10}$$

$$= 67 \quad +1$$

Now, t-statistic is given by

$$t = \frac{\bar{x} - \mu}{\text{S.E of } \mu} \quad +1$$

where S.E of $\mu = \frac{s}{\sqrt{n-1}}$, s = Sample S.D

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}} = \sqrt{\frac{88}{10}} = 2.966 \quad +1$$

$$\therefore \text{S.E} = \frac{2.966}{\sqrt{10-1}} = 0.988$$

$$t = \frac{67-65}{0.988} = 2.02 \quad +1$$

t (calculated) $= 2.02 < t$ (tabulated) $= 2.262$

hence null hypothesis is accepted i.e. population +1

Q.4 (iii) **One Way ANOVA :**

The One Way analysis of variance (ANOVA) is used to determine whether there are any statistically significant differences between the means of three or more independent group.

Specifically, it tests the null hypothesis :

$H_0 : \mu_1 = \mu_2 = \dots\dots = \mu_k$

where μ = group mean, k = no. of groups +2.5

Two Way ANOVA :

The two way analysis of variance compares the mean differences between groups that have been split on two independent variables. The primary purpose of a two way ANOVA is to understand if there is an interaction between the two independent variable on the dependent variable

For eg. to understand whether there is an interaction between gender and educational level on test anxiety among students. +2.5

Q.5 (i)

X	1	2	3	4	5	6
P(X=x)	1/6	1/6	1/6	1/6	1/6	1/6

+1

$E(x) = \sum p_i x_i = \frac{1\times1}{6} + \frac{2\times1}{6} + \frac{3\times1}{6} + \frac{4\times1}{6} + \frac{5\times1}{6} + \frac{6\times1}{6}$ +1

$= 3.5$

Now Variance $= E(x^2) - \{E(x)\}^2$ +1

$= \Sigma p_i x_i^2 - \{E(x)\}^2$

$= \frac{1 \times 1}{6} + \frac{4 \times 1}{6} + \frac{9 \times 1}{6} + \frac{16 \times 1}{6} + \frac{25 \times 1}{6} + \frac{36 \times 1}{6} - (3.5)^2$ +1

$= 2.91$ +1

Q.5(ii) as $M_X(t) = e^{tx} \times P(X=x)$ +1

Given, $M_X(t) = \left(\frac{1}{3} + \frac{2}{3}e^t\right)^5$ +1

On expanding it using Binomial expan.

$= {}^5C_0\left(\frac{1}{3}\right)^5 + {}^5C_1\left(\frac{1}{3}\right)^4\left(\frac{2}{3}e^t\right)^1 + {}^5C_2\left(\frac{1}{3}\right)^3\left(\frac{2}{3}e^t\right)^2$

$+ \cdots + {}^5C_5\left(\frac{2}{3}e^t\right)^5$ +1

Now $P(X=2) = {}^5C_2\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2$ +1

$= \frac{40}{343}$ +1

Q.5(iii) Mean deviation about mean $= \frac{\Sigma|x-\bar{x}|}{n}$

$\Rightarrow \frac{\Sigma|x-1|}{n} = E(|x-1|)$

$= \Sigma |x-1| \cdot P(X)$

$= \Sigma |x-1| \cdot \frac{e^{-1}x^x}{x!}$ +1

as $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$ for poisson distribution

$$\frac{\Sigma|x-\bar{x}|}{n} = e^{-1}\left[\sum_{x=0}^{\infty}\frac{|x-1|}{x!}\right] \quad +1$$

$$= \frac{1}{e}\left[1 + \frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \cdots\right] \quad +1$$

$$= \frac{1}{e}\left[1 + \sum_{x=1}^{\infty}\frac{x-1}{x!}\right]$$

$$= \frac{1}{e}\left[1 + \sum_{x=1}^{\infty}\frac{x}{x!} - \sum_{x=1}^{\infty}\frac{1}{x!}\right] \quad +1$$

$$= \frac{1}{e}\left[1 + \sum_{x=1}^{\infty}\frac{1}{(x-1)!} - \sum_{x=1}^{\infty}\frac{1}{x!}\right]$$

$$= \frac{1}{e}\left[1 + e - (e-1)\right]$$

$$\frac{\Sigma|x-\bar{x}|}{n} = \frac{2}{e} \times 1$$

Hence mean deviation about mean is $\frac{2}{e}$ times of Standard deviation. $+1$

Q.6(i) Given $n(U) = 200$, $n(M) = 120$, $n(P) = 90$, $n(C) = 70$, $n(A \cup B \cup C)' = 20$, $n(M \cap P) = 40$, $n(P \cap C) = 30$, $n(C \cap M) = 50$, $n(A \cap B \cap C) = ?$ $+1$

$$n(A \cup B \cup C)' = n(U) - n(A \cup B \cup C)$$
$$20 = 200 - n(A \cup B \cup C)$$
$$n(A \cup B \cup C) = 180 \quad +1$$

Now $n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(B \cap C) - n(A \cap C) + n(A \cap B \cap C)$ +1

where A, B & C represent, Maths, physics & chemistry respectively

$$\Rightarrow 180 = 120 + 90 + 70 - 40 - 30 - 50 + n(A \cap B \cap C) \quad +1$$

$$\Rightarrow n(A \cap B \cap C) = 20 \quad +1$$

Q.6 (ii) total balls in Bag A = 2W + 5B = 7 Balls
total balls in Bag B = 3W + 6B = 9 Balls

Let E_1 be the event of drawing white ball from Bag A

Then $P(E_1) = \frac{2}{7}$ +1

Let E_2 be the event of drawing Black ball from Bag A

Then $P(E_2) = \frac{5}{7}$ +1

Prob (drawing Black ball from Bag B)

$$= P(E_1 \cap B) + P(E_2 \cap B) \quad +1$$
$$= P(E_1) \times P\left(\frac{B}{E_1}\right) + P(E_2) \times P\left(\frac{B}{E_2}\right) \quad +1$$
$$= \frac{2}{7} \times \frac{6}{10} + \frac{5}{7} \times \frac{7}{10}$$

$= \frac{12}{70} + \frac{35}{70}$

$= \frac{47}{70}$ +1

Q.6 (iii) Sample space of throwing two dice

$=$ (1,1) (1,2) (1,3) (1,4) (1,5) (1,6)
(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)
(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)
(4,1) (4,2) (4,3) (4,4) (4,5) (4,6)
(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)
(6,1) (6,2) (6,3) (6,4) (6,5) (6,6) +1

ie. $n(S) = 36$ +1

Let E_1 be the event of getting sum 11, when 5 appears on first dice

$E_1 = \{(5,6)\}$ +1

$n(E_1) = 1$

∴ required probability $= \frac{n(E_1)}{n(S)}$ +1

$= \frac{1}{36}$ +1