GUJARAT TECHNOLOGICAL UNIVERSITY
BE –SEMESTER 3(NEW SYLLABUS)EXAMINATION-

Subject Code: 3130006 **9-3-21**

Subject Name: Probability and Statistics

Q.1 **(a)** Find the mean, median and mode for the following frequency distribution: **03**

x	1	2	3	4	5	6	7	8	9	10
f	4	7	8	10	6	6	4	2	2	1

Computation:

x	1	2	3	4	5	6	7	8	9	10
f	4	7	8	10	6	6	4	2	2	1
fx	4	14	24	40	30	36	28	16	18	10
cf	4	11	19	29	35	41	45	47	49	50

$$\bar{x} = 4.4, \quad Median = 4, \quad Mode = 4$$

(b) An insurance company insured 2000 bike drivers, 4000 car drivers and 6000 truck drivers. The probability of an accident involving a bike driver, a car driver and a truck driver is 0.10, 0.03 and 0.15 respectively. One of the insured persons meets with an accident. What is the probability that he is a bike driver? **04**

Let A, B and C denote the events that the insured person is a bike driver, car driver and a truck driver. Let D be the event that an insured person meets with an accident. It is given that

$$P(A) = \frac{1}{6}, \quad P(B) = \frac{1}{3}, \quad P(C) = \frac{1}{2}$$

$$P(D|\,A) = 0.10, \quad P(D|\,B) = 0.03, \quad P(D|\,C) = 0.15$$

To find: $P(A|\,D)$

$$P(D) = \frac{1}{6} \times 0.10 + \frac{1}{3} \times 0.03 + \frac{1}{2} \times 0.15 = 0.1017$$

$$P(A|\,D) = \frac{P(D|\,A)}{P(D)} = \frac{\frac{1}{6} \times 0.10}{0.1017} = 0.1639$$

(c) (i) A manufacturer of external hard drives claims that only 10 % of his drives require repairs within the warranty period of 12 months. If 5 of 20 of his drives required repairs within the first year, does this tend to support or refute the claim? **03**

$b(5; 20, 0.10) = 0.0319$. In fact, $B(5; 20, 0.10) = 0.0432$.

This is a very small probability. So, the claim should be refuted.

(ii) The actual amount of instant coffee that a filling machine puts into "4 –ounce" jars may be looked upon as a random variable having a normal distribution with $\sigma = 0.04$ ounce. If only 2 % of the jars are to contain less than 4 ounces, what should be the mean fill of these jars? Out of 10000 jars sold, how many are expected to contain more than 4.2 ounces? **04**

Given that

$$F\left(\frac{4-\mu}{0.04}\right) = 0.02 \qquad \therefore 1 - F\left(\frac{4-\mu}{0.04}\right) = F\left(\frac{\mu - 4}{0.04}\right) = 0.98$$

$$\therefore \frac{\mu - 4}{0.04} = 2.055 \quad or\ \mu = 4.0822$$

Also,

$$1 - F\left(\frac{4.2 - 4.0822}{0.04}\right) = 1 - F(2.945) = 1 - 0.9984 = 0.0016$$

So, number of jars expected to contain more than 4.2 ounces is

$$10000 \times 0.0016 = 16$$

Q.2 (a) If the probability that an individual suffers a bad reaction from a certain injection is 0.001. Find the probabilities that out of 2000 individuals, (i) more than 2 individuals; (ii) exactly 3 individuals will suffer a bad reaction. 03

$$\lambda = np = 2000 \times 0.001 = 2$$

$(i)\ Req.Prob. = 1 - F(2\,;2) = 1 - 0.677 = 0.323$

$(ii)\ Req.Prob. = f(3;2) = 0.857 - 0.677 = 0.18$

(b) A stenographer claims that she can type at the rate of 120 words per minute. She demonstrated, on the basis of 100 trials, an average speed of 116 words with a standard deviation of 15 words. Does this enable us to reject the null hypothesis $\mu = 120$ against the alternative hypothesis $\mu < 120$ at the 0.05 level of significance? 04

Null Hypothesis: $\mu = 120$

Alternative Hypothesis: $\mu < 120$

Level of Significance: $\alpha = 0.05$

Criterion: If $Z < -z_\alpha = -z_{0.05} = -1.645$, then reject the claim.

Calculation:

$$Z = \frac{116 - 120}{15/\sqrt{100}} = \frac{-4}{1.5} = -2.6667 < -1.645$$

Decision: Reject the claim.

(c) (i) The time to check out and process payment information at an office supplies Web site can be modeled as a random variable with mean $\mu = 63$ seconds and variance $\sigma^2 = 81$ seconds. What is the probability that a random sample of size 36 has mean greater than 66.75? 03

$$Req.Prob. = 1 - F\left(\frac{66.75 - 63}{9/6}\right) = 1 - F(2.5) = 1 - 0.9938 = 0.0062$$

(ii) If two random variables X and Y have the joint density 04

$$f(x,y) = \begin{cases} k(x+y^2), & for\ 0 < x < 1, 0 < y < 1 \\ 0, & elsewhere \end{cases},$$

find k and the mean of the conditional density $f_1(x \mid 0.5)$ where $f_1(x)$ is the marginal probability density of X.

Since $f(x,y)$ is a joint probability density,

$$\int_0^1 \int_0^1 k(x+y^2)\,dxdy = 1$$

$$\therefore k\int_0^1 \left[\frac{x^2}{2} + xy^2\right]_0^1 = k\int_0^1 \left[\frac{1}{2} + y^2\right]dy = k\left[\frac{1}{2}y + \frac{y^3}{3}\right]_0^1 = \frac{5}{6}k$$

$$k = \frac{6}{5}$$

$$f_2(y) = \int_0^1 \frac{6}{5}(x+y^2)\,dx = \frac{6}{5}\left[\frac{x^2}{2} + xy^2\right]_0^1 = \frac{6}{5}\left[\frac{1}{2} + y^2\right]$$

$$\therefore f_1(x \mid y) = \frac{f(x,y)}{f_2(y)} = \frac{x+y^2}{0.5+y^2} \qquad \therefore f_1(x \mid 0.5) = \frac{x+0.25}{0.75}$$

. Therefore, the mean of the conditional density $f_1(x \mid 0.5)$ is

$$\int_0^1 x\frac{x+0.25}{0.75}\,dx = \frac{1}{0.75}\left[\frac{x^3}{3} + 0.25\frac{x^2}{2}\right]_0^1 = \frac{11}{18} = 0.6111$$

OR

(c) (i) The following are the times between 6 calls for an ambulance and the patients's arrival at the hospital: 27, 15, 20, 32, 18 and 26 minutes. Use these figures to judge the reasonableness of the ambulance service's claim that it takes on the average of 20 minutes between the call for an ambulance and the patient's arrival at the hospital. 03

$\bar{x} = 23,\ s = 6.3875.$

$$t = \frac{23 - 20}{6.3875/\sqrt{6}} = 1.15 < 1.476 = t_{0.10}\ for\ 5\ degrees\ of\ freedom$$

So, we cannot reject the claim.

(ii) Let X and Y be two random variables having joint probability mass function $f(x, y) = \frac{1}{27}(2x + y)$, where x and y can assume only integer values 0, 1 and 2. Find conditional distribution of Y for $X = x$. **04**

X \ Y	0	1	2	$f_X(x)$
0	0	$1/27$	$2/27$	$3/27$
1	$2/27$	$3/27$	$4/27$	$9/27$
2	$4/27$	$5/27$	$6/27$	$15/27$
$f_Y(y)$	$6/27$	$9/27$	$12/27$	1

The conditional distribution is

X \ Y	0	1	2
0	0	$1/3$	$2/3$
1	$2/9$	$3/9$	$4/9$
2	$4/15$	$5/15$	$6/15$

Q.3 **(a)** If A and B are independent events with $P(A) = 0.26$, and $P(B) = 0.45$, find $(a)\ P(A \cap B)$; $(b)\ P(A \cap \bar{B})$; $(c)\ P(\bar{A} \cap \bar{B})$. **03**

$(a)\ P(A \cap B) = P(A)P(B) = 0.26 \times 0.45 = 0.117$

$(b)\ P(A \cap \bar{B}) = P(A) - P(A \cap B) = 0.26 - 0.117 = 0.143$

$(c)\ P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(A \cap B)]$

$\therefore P(\bar{A} \cap \bar{B}) = 1 - [0.26 + 0.45 - 0.117] = 0.407$

(b) Compute Karl Pearson's coefficient of correlation between X and Y for the following data: **04**

X	100	98	78	85	110	93	80
Y	85	90	70	72	95	81	74

Computation:

X	Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X})(Y - \bar{Y})$	$(X - \bar{X})^2$	$(Y - \bar{Y})^2$
100	85	8	4	32	64	16
98	90	6	9	54	36	81
78	70	−14	−11	154	196	121
85	72	−7	−9	63	49	81
110	95	18	14	252	324	196
93	81	1	0	0	1	0
80	74	−12	−7	84	144	49
				639	814	544

$$\bar{X} = 92,\ \bar{Y} = 81,\ r = \frac{639}{\sqrt{814}\sqrt{544}} = 0.9603$$

(c) (i) The arithmetic means of runs scored by three batsmen A, B and C, in the same series of 10 innings, are 50, 48 and 12 respectively. The standard deviations of their runs are 15, 12 and 2 respectively. Who is the most consistent of the three? **03**

$$C.V._A = \frac{15}{50} \times 100 = 30\ \%, C.V._B = \frac{12}{48} \times 100 = 25\ \%,$$

$$C.V._C = \frac{2}{12} \times 100 = 16.67\ \%.$$

Therefore, C is the most consistent.

(ii) Calculate the first four moments about the mean of the following data: **04**

x	0	1	2	3	4	5	6	7	8
f	1	8	28	56	70	56	28	8	1

Four moments about the mean

x	f	fx	$d = x - \bar{x}$	fd	fd^2	fd^3	fd^4
0	1	0	-4	-4	16	-64	256
1	8	8	-3	-24	72	-216	648
2	28	56	-2	-56	112	-224	448
3	56	168	-1	-56	56	-56	56
4	70	280	0	0	0	0	0
5	56	280	1	56	56	56	56
6	28	168	2	56	112	224	448
7	8	56	3	24	72	216	648
8	1	8	4	4	16	64	256
	256	1024	0	0	512	0	2816

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = 4, \qquad \mu_1 = \mu_3 = 0,\ \mu_2 = \frac{\sum fd^2}{\sum f} = \frac{512}{256} = 2,\ \mu_4 = \frac{2816}{256} = 11$$

Q.4

(a) The following table gives the probabilities that a certain computer will malfunction 0, 1, 2, 3, 4, 5 or 6 times on any one day: **03**

Number of malfunctions	$x:$	0	1	2	3	4	5	6
Probability	$f(x):$	0.17	0.29	0.27	0.16	0.07	0.03	0.01

Find the mean and variance of this probability distribution.

$$\mu = \sum xf(x) = 1.8, \quad \sum x^2 f(x) = 5.04, \quad \sigma^2 = 1.8.$$

(b) The coefficient of rank correlation of marks obtained by 10 students in English and Economics was found to be 0.6. It was later discovered that the difference in ranks in the two subjects obtained by one of the students was wrongly taken as 7 instead of 1. Find the correct coefficient of rank correlation. **04**

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)} \qquad \therefore 0.6 = 1 - \frac{6\sum d^2}{990} \qquad \therefore \sum d^2 = 66$$

Corrected $\sum d^2 = 66 - 49 + 1 = 18$ Corrected $r = 0.8909$.

(c) (i) Find out mean deviation about median for the following series: **03**

Size	4	6	8	10	12	14	16
Freq.	2	1	3	6	4	3	1

Table:

x	f	cf	$d = \lvert x - M \rvert$	fd
4	2	2	6	12
6	1	3	4	4
8	3	6	2	6
10	6	12	0	0
12	4	16	2	8
14	3	19	4	12
16	1	20	6	6

Median= 10

Mean deviation about median

$$= \frac{\sum fd}{\sum f} = \frac{48}{20} = 2.4$$

(ii) Find Karl Pearson's coefficient of skewness for the following data: **04**

x	$0-10$	$10-20$	$20-30$	$30-40$	$40-50$
f	13	20	30	25	12

Karl Pearson's corfficient of skewness

x	f	x_i	f_ix_i	$(x_i-\bar{x})^2$	$f_i(x_i-\bar{x})^2$	
$0-10$	13	5	65	412.09	5357.17	$\bar{x}=25.3$
$10-20$	20	15	300	106.09	2121.80	$\sigma=12.0379$
$20-30$	30	25	750	0.09	2.7	$Mode=26.67$
$30-40$	25	35	875	94.09	2352.25	$S_k=\frac{\bar{x}-Mode}{\sigma}$
$40-50$	12	45	540	388.09	4657.08	$=\frac{25.3-26.67}{12.0379}$
	100		2530		14491	$=-0.1138$

Q.5.

(a) The life in hours of a certain kind of radio tube has the probability density **03**

$$f(x)=\begin{cases}100/x^2, & for\ x\geq 100\\ 0, & elsewhere\end{cases},$$

find the distribution function and use it to determine the probability that the life of tube is more than 150 hrs.

Case 1: If $x<100,\quad F(x)=0.$

Case 2: If $x\geq 100,$

$$F(x)=\int_{100}^{x}\frac{100}{x^2}dx=100\left[-\frac{1}{x}\right]_{100}^{x}=100\left[-\frac{1}{x}+\frac{1}{100}\right]=1-\frac{100}{x}$$

$$F(x)=\begin{cases}1-\frac{100}{x}, & if\ x\geq 100\\ 0, & if\ x<100\end{cases}.$$

$$Req.Prob.=1-F(150)=1-100/_{150}=\frac{1}{3}=0.3333$$

(b) The number of flaws in a fiber optic cable follows a Poisson process with an average of 0.6 per 100 feet. **04**

(i) Find the probability of exactly 2 flaws in a 200 foot cable.

(ii) Find the probability of exactly 1 flaw in the first 100 feet and exactly 1 flaw in the second 100 feet.

Here $\alpha=0.6.$

(i) $\lambda=2*0.6=1.2.$

$$Req.Prob.=f(2;1.2)=0.879-0.663=0.216$$

(ii) Getting 1 flaw in the first 100 feet and getting 1 flaw in the second 100 feet are independent events. Therefore,

$$Req.Prob.=f(1;0.6)\times f(1;0.6)=0.329\times 0.329=0.1082$$

(c) The population (p) of a small community on the outskirts of a city grows rapidly over a 20 −year period: **07**

t	0	5	10	15	20
p	100	200	450	950	2000

As an engineer working for a utility company, you must forecast the population 5 years into the future in order to anticipate the demand for power. Employ an exponential model and linear regression to make this prediction.

$p=ae^{bt}\qquad\therefore \log p=\log a+bt\qquad P=A+bt$

t	p	$P=\log p$	tP	t^2	
0	100	4.6052	0	0	$b=0.15$
5	200	5.2983	26.4916	25	$A=\bar{P}-b\bar{t}=4.594$
10	450	6.1092	61.0925	100	$a=e^A=98.89$
15	950	6.8565	102.8469	225	$\therefore p=98.89e^{0.15\,t}$
20	2000	7.6009	152.0180	400	When $t=25,$
50		30.4701	342.4490	750	$p\approx 4205$

Q.6

(a) The joint probability density of two random variables is given by **03**

$$f(x_1,x_2)=\begin{cases}6e^{-2x_1-3x_2}, & for\ x_1>0,x_2>0\\ 0, & elsewhere\end{cases}.$$

Find the marginal densities of both the random variables and hence show that the two random variables are independent.

$$f_1(x_1) = 6e^{-2x_1}\int_0^{\infty} e^{-3x_2}\,dx_2 = 6e^{-2x_1}\left[\frac{e^{-3x_2}}{-3}\right]_0^{\infty} = 2e^{-2x_1}$$

$$f_2(x_2) = 6e^{-3x_2}\int_0^{\infty} e^{-2x_1}\,dx_1 = 6e^{-3x_2}\left[\frac{e^{-2x_1}}{-2}\right]_0^{\infty} = 3e^{-3x_2}$$

It can be seen that $f(x_1,x_2) = f_1(x_1)\cdot f_2(x_2)$. So, the two random variables are independent.

(b) The probability that an electronic component will fail in less than 1000 hours of continuous use is 0.25. Use the normal approximation to find the probability that among 200 such components fewer than 45 will fail in less than 1000 hours of continuous use. **04**

$$p = 0.25,\quad n = 200,\quad \mu = np = 50,\quad \sigma = \sqrt{np(1-p)} = 6.12$$

$$Req.Prob. = F\left(\frac{44.5-50}{6.12}\right) = F(-0.9) = 0.1841$$

(c) Fit a parabola $y = a + bx + cx^2$ to the following data: **07**

x	1	2	3	5	6
y	1.1	5.8	17.5	55.9	86.7

Computation:

x	y	x^2	x^3	x^4	xy	x^2y
1	1.1	1	1	1	1.1	1.1
2	5.8	4	8	16	11.6	23.2
3	17.5	9	27	81	52.5	157.5
5	55.9	25	125	625	279.5	1397.5
6	86.7	36	216	1296	520.2	3121.2
17	167	75	377	2019	864.9	4700.5

$$\begin{bmatrix} 5 & 17 & 75 \\ 17 & 75 & 377 \\ 75 & 377 & 2019 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 167 \\ 864.9 \\ 4700.5 \end{bmatrix}$$

$$a = 2.7227,\quad b = -4.5528,\quad c = 3.0771$$

$$y = 2.7227 - 4.5528\,x + 3.0771\,x^2$$

Q. 7 ~~08~~ **(a)** In a study of automobile collision insurance costs, a random sample of 80 body repair costs for a particular kind of damage had a mean of 33065 $Rs.$ and a standard deviation of 4364 $Rs.$ If $\bar{x} = 33065\ Rs.$ is used as a point estimate of the true average repair cost of this kind of damage, with what confidence can one assert that the error does not exceed 700 $Rs.$? **03**

$$E = z_{\alpha/2}\cdot\frac{s}{\sqrt{n}}\qquad \therefore z_{\alpha/2} = 700\times\frac{\sqrt{80}}{4364} = 1.4347$$

$$\therefore F(z_{\alpha/2}) = 0.9236\qquad \therefore 1-\alpha/2 = 0.9236\qquad \therefore \alpha = 0.1528$$

$$\therefore confidence = (1-\alpha)\times 100 = 84.72\ \%$$

(b) In a certain city, the daily consumption of electric power (in millions of kilowatt-hours) can be treated as a random variable having a gamma distribution with $\alpha = 2$ and $\beta = 3$. If the power plant of this city has a daily capacity of 12 million kilowatt-hours, what is the probability that this power supply will be inadequate on any given day? Also, find the mean of this probability density. **04**

Gamma distribution with $\alpha = 2$ and $\beta = 3$ is

$$f(x) = \begin{cases} \frac{1}{9}xe^{-x/3}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

$$Req.Prob. = \frac{1}{9}\int_{12}^{\infty} xe^{-x/3}\,dx = \frac{1}{9}\left[x\frac{e^{-x/3}}{-1/3} - \frac{e^{-x/3}}{1/9}\right]_{12}^{\infty}$$
$$= \left[-\frac{x}{3}e^{-x/3} - e^{-x/3}\right]_{12}^{\infty} = 5e^{-4} = 0.0916$$

The mean of this probability density is

$$\frac{1}{9}\int_{0}^{\infty} x^2 e^{-x/3}\,dx = \frac{1}{9}\left[x^2\frac{e^{-x/3}}{-1/3} - 2x\frac{e^{-x/3}}{1/9} + 2\frac{e^{-x/3}}{-1/27}\right]_{0}^{\infty} = 6.$$

(c) (i) Ten bearings made by a certain process have a mean diameter of $0.506\ cm$ and a standard deviation of $0.004\ cm$. Assuming that the data may be looked upon as a random variable from a normal population, construct a 95 % confidence interval for the actual average diameter of bearings made by this process. **03**

$$\alpha = 0.05, \quad t_{0.025}\ for\ 9\ degrees\ of\ freedom = 2.262$$
$$\bar{x} - t_{\alpha/2}\cdot\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\cdot\frac{s}{\sqrt{n}}$$
$$0.506 - 2.262\cdot\frac{0.004}{\sqrt{10}} < \mu < 0.506 + 2.262\cdot\frac{0.004}{\sqrt{10}}$$
$$0.5031 < \mu < 0.5089$$

(ii) A consumer protection agency wants to test a paint manufacturer's claim that the average drying time of his new paint is 20 minutes. It instructs a member of its research staff to paint each of 36 boards using a different 1 –gallon can of the paint, with the intention of rejecting the claim if the mean of the drying times exceeds 20.75 minutes. Otherwise, it will accept the claim. Find the probability of a Type I error. Also, find the probability of a Type II error when $\mu = 21$ minutes. Assume that $\sigma = 2.4$ minutes. **04**

The probability of a Type I error is

$$1 - F\left(\frac{20.75 - 20}{2.4/\sqrt{36}}\right) = 1 - F(1.875) = 1 - 0.9696 = 0.0304$$

The probability of a Type II error when $\mu = 21$ minutes is

$$F\left(\frac{20.75 - 21}{2.4/\sqrt{36}}\right) = F(-0.625) = 0.2660$$

OR

Q.8 **(a)** The dean of a college wants to use the mean of a random sample to estimate the average amount of time students take to get from one class to the next, and she wants to be able to assert with 99 % confidence that the error is at most 0.25 minute. If it can be presumed from experience that $\sigma = 1.40$ minutes, how large a sample will she have to take? **03**

$$\alpha = 0.01, \quad E = z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} \qquad \therefore n = \left[z_{0.005}\cdot\frac{1.4}{0.25}\right]^2 \approx 208$$

(b) How exponential distribution is useful in real applications? Find the mean and variance of the exponential distribution **04**

$$f(x) = \begin{cases} \frac{1}{\beta}e^{-x/\beta}, & for\ x > 0, \beta > 0 \\ 0, & elsewhere \end{cases}$$

In a Poisson process, if the mean arrival rate is α, the time until the first arrival, or the waiting time between successive arrivals, has an exponential distribution with $\beta = 1/\alpha$.

$$\mu = \frac{1}{\beta}\int_0^\infty x\, e^{-x/\beta}\, dx = \frac{1}{\beta}\left[x\frac{e^{-x/\beta}}{-1/\beta} - \frac{e^{-x/\beta}}{1/\beta^2}\right]_0^\infty = \beta$$

$$\int_0^\infty \frac{1}{\beta} x^2\, e^{-x/\beta}\, dx = \frac{1}{\beta}\left[x^2\frac{e^{-x/\beta}}{-1/\beta} - 2x\frac{e^{-x/\beta}}{1/\beta^2} + 2\frac{e^{-x/\beta}}{-1/\beta^3}\right]_0^\infty = 2\beta^2$$

$$\sigma^2 = \mu_2' - \mu^2 = 2\beta^2 - \beta^2 = \beta^2$$

(c) A random sample from a company's very extensive files shows that orders for a certain piece of machinery were filled, respectively, in $10, 12, 19, 14, 15, 18, 11$ and 13 days. Use the level of significance $\alpha = 0.01$ to test the claim that on average such orders are filled in 10.5 days. Choose the alternative hypothesis so that rejection of the null hypothesis $\mu = 10.5$ implies that it takes longer than indicated. Assume normality. **07**

$$\bar{x} = 14, \quad s = 3.2071$$

Null Hypothesis: $\mu = 10.5$
Alternative Hypothesis: $\mu > 10.5$
Level of Significance: $\alpha = 0.01$
Criterion: If $t > t_{0.005} = 3.499$ for 7 degrees of freedom, then reject the claim.
Calculation:

$$t = \frac{14 - 10.5}{3.2071/\sqrt{8}} = 3.0867$$

Decision: Since $t = 3.0867 < 3.499$, we cannot reject the claim.

****