Total No. of Pages: 2 | *Roll No. ......*

B.Tech. (CSE)

(Summer Semester)

MID SEMESTER EXAMINATION | **June-2023**

CO325 Probability and Statistics

Time: 1:30 Hours | *Max. Marks: 25*

Note: Answer *ALL* questions.
Assume suitable missing data, if any.
CO# is course outcome related to question.

Total No. of Pages 2 | *Roll No. ......*

FIFTH SEMESTER | **B.Tech. (CSE)**

MID SEMESTER EXAMINATION | ***September-2022***

CO325 Probability and Statistics

Time: 1:30 Hours | *Max. Marks: 25*

Note: Answer *ALL* questions.
Assume suitable missing data, if any.
CO# is course outcome related to question

Total No. of Pages 3 | *Roll No. ......*

FIFTH SEMESTER | **B.Tech. (CSE)**

END SEMESTER EXAMINATION | ***JULY-2023***

CO325 Probability and Statistics

Time: 3:00 Hours | *Max. Marks: 40*

Note: Answer *ALL* questions.
Assume suitable missing data, if any.
CO# is course outcome related to question

Total No. of Pages 3 | *Roll No.* 2K20/

FIFTH SEMESTER

B.Tech. (CSE)

END SEMESTER EXAMINATION | ***NOVEMBER-2022***

CO325 Probability and Statistics

Time: 3 Hours | *Max. Marks: 50*

Note: Answer *ALL* questions.
Assume suitable missing data, if any.
CO# is course outcome related to question

			Date	Grade/ Marks	Signature
52.					
53.					
54.					
55.					
56.					
57.					
58.					
59.					
60.					
61.					

Date : / /
Page No.

[c] A box contains 20 defective items and 80 non-defective items. If two items are selected at random without replacement, what will be the probability that both items are defective? 2023 SM [2] [CO2]

1[a] A box contains 20 defective items and 80 non-defective items. If two items are selected at random without replacement, what will be the probability that both items are defective? 2022 M [2.5] [CO1]

total no. of items = 100
defective items = 20
non-defective items = 80

$\therefore$ no. of ways picking 2 defective items $= {}^{20}C_2 \times {}^{80}C_0 = x$
no. of ways picking any 2 items $= {}^{100}C_2 = y$

$\therefore$ $P(x = \text{both items are defective}) = x/y = \dfrac{{}^{20}C_2 \times {}^{80}C_0}{{}^{100}C_2} = \dfrac{19}{495}$

1[a] A man is known to speak the truth 3 out of 4 times. He throws a die and reports that the number obtained is a six. Find the probability that the number obtained is actually a six? 2023 SE [2.5] [CO1]

$P(A)$ = probability that man speaks truth = 3/4
$\therefore$ $P(A')$ = " " " lie = 1/4

$P(B)$ = probability of getting 6 = 1/6
$P(B')$ = ―――― not getting 6 = 5/6

~~given~~
probability that we get 6 by man's report $= P(A)\,P(B) + P(A')\,P(B')$
$= 1/6 \cdot 3/4 + 5/6 \cdot 1/4$
$= 8/24 = 1/3$

probability that its actually 6 $= P(A) \cdot P(B) = 1/6 \cdot 3/4 = 3/24$

or by Bayes thm $\therefore$ $P(A|B) = \dfrac{P(B) \cdot P(A)}{P(T)} = \dfrac{P(B)\,P(A)}{P(B)\,P(A) + P(A')\,P(B')}$

$\boxed{= 3/8}$

Date : / /
Page No.

[b] An insurance company insured 2000 scooter drivers, 4000 car drivers, 6000 truck drivers. Probability of an accident involving scooter, car, and truck is 0.01, 0.03, and 0.15 respectively. One of insured persons meet with an accident. What is probability, he is scooter driver?

2023 SE [2.5] [CO2]

Let,

$E_1 \rightarrow$ Scooter driver	$n(E_1) = 2000$
$E_2 \rightarrow$ Car driver	$n(E_2) = 4000$
$E_3 \rightarrow$ truck driver	$n(E_3) = 6000$
$A \rightarrow$ person meets accident.	$n(\text{total}) = 12000$

probality of a person being what kind of driver

$\therefore P(E_1) = 1/6$

$P(E_2) = 1/3$

$P(E_3) = 1/2$

Now, ~~what type of~~ driver meets accident probab for each vehicle.

$P(A|E_1) = 0.01$

$P(A|E_2) = 0.03$

$P(A|E_3) = 0.15$

$\therefore$ Probability that driver is a scooter driver given that he met with an accident

Bayes

$$P(E_1|A) = \frac{P(E_1) \cdot P(A/E_1)}{P(T)} = \frac{P(E_1) \cdot P(A/E_1)}{P(E_1) \cdot P(A|E_1) + P(E_2) \cdot P(A/E_2) + P(E_3) \cdot P(A/E_3)}$$

$$= \frac{1/6 \cdot 1/100}{1/6 \cdot \frac{1}{100} + \frac{1}{3} \cdot \frac{3}{100} + \frac{1}{2} \cdot \frac{15}{100}}$$

$= 1/52$

$\therefore$ Ans $\rightarrow$ 1/52

Date : / /
Page No.

OR

[b] In manufacturing factory, machines A, B, C produce 25%, 35%, 40% bolts. Out of total, 5%, 4%, 2% are defective bolts. A bolt is drawn at random from the product. If bolt drawn is found to be defective, what is the probability, it is manufactured by 'B'?

2023 SE [2.5] [CO2]

let $A \rightarrow$ bolt is produced by A
$B \rightarrow$ ——— B
$C \rightarrow$ ——— C
$X \rightarrow$ bolt is defective.

$P(A) = 1/4 = 25/100$
$P(B) = 35/100$
$P(C) = 40/100$

Now, defective bolt prob given its made by a machine.

$P(X/A) = 5/100$
$P(X|B) = 4/100$
$P(X|C) = 2/100$

by Bayes, $P(B/X) = \dfrac{P(B).P(X/B)}{P(A).P(X|A) + P(B).P(X|B) + P(C).P(X|C)}$

$= \dfrac{35.4}{25.5 + 35.4 + 40.2} = \dfrac{28}{69}$

$\boxed{= \dfrac{28}{69}}$

7. [a] Three coins are tossed. Let X denote the number of heads on the first two and Y denote the number of heads on the last two. Find:

1. The joint distribution of X and Y.
2. Marginal distribution of X and Y.
3. E(X) and E(Y)

2023 SE

[4] [CO1]

1. Joint Distribution

X \ Y	0	1	2	g
0	1/8	1/8	0	2/8
1	1/8	2/8	1/8	4/8
2	0	1/8	1/8	2/8
h	2/8	4/8	2/8	

OUTCOME	X	Y	f(x,y)
H H H	2	2	1/8
H H T	2	1	1/8
H T H	1	1	1/8
H T T	1	0	1/8
T H H	1	2	1/8
T H T	1	1	1/8
T T H	0	1	1/8
T T T	0	0	1/8

2. Marginal Distribution.

$g(x=0) = 2/8$ | $h(y=0) = 2/8$

$g(x=1) = 5/8$ | $h(y=1) = 5/8$

$g(x=2) = 2/8$ | $h(y=2) = 2/8$

3) $E(x) = \sum_{i=0}^{2} p_{x_i} \cdot i = \frac{2}{8} \cdot 0 + \frac{4}{8} \cdot 1 + \frac{2}{8} \cdot 2 = 1$

$E(y) = \sum_{i=0}^{2} i \cdot p_{y_i} = \frac{2}{8} \cdot 0 + \frac{4}{8} \cdot 1 + \frac{2}{8} \cdot 2 = 1$

$\therefore \boxed{E(x) = 1}, \boxed{E(y) = 1}$

Date : / /
Page No.

[b] The likelihood of the 3 teams A, B, C winning a football match are 1/3, 1/5 and 1/9 respectively. Find the probability that neither team A nor team B will win the match. 2023 SM [1] [CO1]

[3] The likelihood of the 3 teams A, B, C winning a football match are 1/3, 1/5 and 1/9 respectively. Find the probability that Neither team A nor team B will win the match. 2022 M [1] [CO1]

let

$A \rightarrow$ team A wins the match	$P(A) = 1/3$
$B \rightarrow$ team B wins the match	$P(B) = 1/5$
$C \rightarrow$ team C wins the match.	$P(C) = 1/9$

$\therefore$ probability that neithe A nor B with win the match

$= P(A') \cdot P(B')$ (independent)

$= (1-1/3) \cdot (1-1/5)$

$= (2/3) \cdot (4/5)$

$\boxed{= 8/15}$

[b] For three events A, B and C,
(i) A and C are independent (ii) B and C are independent (iii) A and B are disjoint (iv) P(AUC) = 2/3, P(BUC) = 3/4, P(AUBUC) = 11/12
Then find P(A), P(B), and P(C). 2022 M [2] [CO1]

given,

$P(A \cup C) = 2/3$	$P(A \cap B) = 0$ (disjoint)	
$P(B \cup C) = 3/4$	$P(A \cap B \cap C) = 0$ (")	
$P(A \cup B \cup C) = 11/12$	$P(A \cap C) = P(A) \cdot P(C)$	$P(B \cap C) = P(B) \cdot P(C)$

$\therefore P(A) + P(C) - P(A \cap C) = 2/3$ — ①

$P(B) + P(C) - P(B \cap C) = 3/4$ — ②

$P(A) + P(B) + P(C) - P(A \cap C) - \cancel{P(A \cap B)}^{0} - P(B \cap C) + \cancel{P(A \cap B \cap C)}^{0} = \frac{11}{12}$ — ③

$\Rightarrow$

by using ①, ②

$P(A) + P(B) + P(C) - (P(A) + P(C) - 2/3) = 11/12$

$\Rightarrow \boxed{P(C) = 1/2}$

using in ①

$\frac{1}{2} + P(A) - \frac{1}{2} P(A) = 2/3 \Rightarrow \boxed{P(A) = 1/3}$

using in ②

$P(B) + \frac{1}{2} - \frac{1}{2} P(B) = 3/4 \Rightarrow \boxed{P(B) = 1/2}$

Date : / /
Page No.

#red = 6
#blue = 5 total = 15
#yellow = 4

2[a] A bag contains 6 red, 5 blue, and 4 yellow balls. 2 balls are drawn, but the first ball is drawn without replacement. Find the P (red, then blue) 2023 SE [1] [CO1]

Probability of drawing red in first turn $= \frac{6}{15} = P(A)$

Prob of drawing blue in second turn $= \frac{\text{no. of blue balls}}{\text{remaining no. of balls}}$

$\therefore$ New no. of balls = 14

$= 5/14 = P(B)$

$\therefore P(\text{red then blue}) = P(A).P(B) = \frac{\cancel{6}^2}{\cancel{15}_3} \times \frac{\cancel{5}}{\cancel{14}_7} = \boxed{\frac{1}{7}}$

1[a] Two players, A and B, play a tennis match. The probability of A winning the match is 0.62. What is the probability that B will win the match? 2023 SM 0. 38 [1] [CO1]

[b] In a neighborhood, 90% children were falling sick due flu and 10% due to measles and no other disease. The probability of observing rashes for measles is 0.95 and for flu is 0.08. If a child develops rashes, find the child's probability of having flu. 2023SM [2] [CO2]

Let,

$A \to$ Doctor finds a rash

$B_1 \to$ Child has measles

$B_2 \to$ child has flu

$\therefore$ Given,

$P(B_1) = 1/10$

$P(B_2) = 9/10$

$P(A/B_1) = 0.95$

$P(A|B_2) = 0.08$

ATQ, flu given rashes

$$\therefore P(B_2/A) = \frac{P(B_2).P(A|B_2)}{P(B_1).P(A|B_1) + P(B_2).P(A|B_2)}$$

$$= \frac{1/10 \cdot 0.95}{1/10 \cdot 0.95 + 9/10 \cdot 0.08}$$

$$= \frac{95}{167}$$

Date : / /
Page No.

2[a] Consider the experiment of rolling a die. Let A be the event 'getting a prime number', B be the event 'getting an odd number'. Write the sets representing the events. [4] [CO1]

(i) A or B
(ii) A and B
(iii) A but not B
(iv) 'not A'

2023 SM

~~Sample~~ Sample Space : $S = \{1, 2, 3, 4, 5, 6\}$

$A = \{2, 3, 5\}$ → prime no. | $B = \{1, 3, 5\}$

~~$P(A) = 1/2$~~ | ~~$P(B) = 1/2$~~

A and B are not indtendent.

i.) A or B = $A \cup B = \{1, 2, 3, 5\}$

ii.) A and B = $A \cap B = \{3, 5\}$

iii.) A but not B = $A - B = \{2\}$

iv.) not A = $A' = S - A = \{1, 4, 6\}$

5. [a] An ordinary 52 card deck is thoroughly shuffled. You are dealt four cards up. What is the probability that all four cards are sevens?

2023SE [2] [CO1]

No of ways to deal 4 cards = ${}^{52}C_4$

No of ways to deal 4 seven cards = ${}^{4}C_4 = 1$

$\therefore$ prob $= \dfrac{1}{{}^{52}C_4} = \dfrac{1}{270725}$

[b] Consider the experiment in which a coin is tossed repeatedly until a head comes up. Describe the sample space. 2023 SE [1] [CO1]

Heads may come in the 1st, 2nd, 3rd ---. toss so till head is obtained, we toss.

$S = \{H, TH, TTH, TTTH ---\}$

[c] Two dice are thrown together. What is the probability that the number obtained on one of the dice is multiple of number obtained on the other dice?

2023 SE [2] [CO1]

total no. of ways = 36

Ways by which its possible = {(1,1), (2,2), (3,3), (4,4), (5,5) (6,6), (2,1), (1,2), (3,1), (1,3), (1,4) (1,5), (4,1) (5,1), (6,1), (1,6), (2,4), (4,2), (3,6) (6,3), (2,6), (6,2)}

$\therefore$ favorable cases = 22

$\therefore$ P(dice has multiple on another) = $22/36 = 11/18$

OR

8. A missile can be accidentally launched if two relays A and B both have failed. The probabilities of A and B failing are known to be 0.01 and 0.03, respectively. It is also known that B is more likely to fail (probability 0.06) if A has failed.
[a] What is the probability of an accidental missile launch?
[b] What is the probability that A will fail if B has failed?
[c] Are the events "A fails" and "B fails" statically independent?
2023 SE [5] [CO2]

$A \rightarrow$ A fails $\quad \therefore P(A) = 0.01$

$B \rightarrow$ B fails $\quad P(B) = 0.03$

given

$P(B/A) = 0.06$

a.) $P(\text{accidental launch}) = P(A) + P(B) - P(A \cap B)$

$= 0.04 - 0.0006$

$= 0.0394$

$P(A \cap B) = P(B/A) \cdot P(A)$

$= 0.06 * 0.01$

$= 0.0006$

b.) $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.0006}{0.03} = 0.02$

c.) No, because $P(A \cap B) \neq 0$.

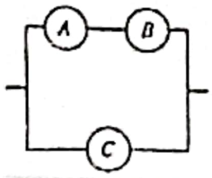
[b] What is the probability that a non-leap year selected at random will contain 53 sundays?
2023 SE [1] [CO1]

A non-leap year has 365 days $\Rightarrow$ 52 full weeks and 1 extra day.

This day can be either of the 7 M, T, W, Th, F, S, Su

$\therefore \boxed{P(\text{sunday}) = 1/7}$

6. [a] The electrical apparatus in the diagram works so long as current can flow from left to right. The three components are independent. The probability that component A works is 0.8; the probability that component B works is 0.9; and the probability that component C works is 0.75. Find the probability that the apparatus works. [2.5] [CO2]



2023 SE

$P(A) = 0.8$

$P(B) = 0.9$

$P(C) = 0.75 \Rightarrow P(C') = 0.25$

$P(A \cap B) = 0.8 * 0.9 = 0.72$

$P((A \cap B)') = 1 - 0.72 = 0.28$

$\therefore P(\text{circuit work}) = 1 - P(\text{ckt fail}) = 1 - P((A \cap B)') \cdot P(C') = 1 - 0.28 * 0.25$

$= 0.93$

Date : / /
Page No.

[b] An examination paper has 150 MCQ of 1 mark each, with each question having 4 choices. Each incorrect answer fetches -0.25 mark. Suppose 1000 students choose all their answers randomly with uniform probability. The sum of the expected marks obtained by all these students is?
2022M [2.5] [CO2]

Let the marks obtained per q be a random variable X

for $x = 1 \Rightarrow P(X) = 1/4$ (4 choices)
for $x = -0.25 \Rightarrow P(x) = 3/4$

$\therefore$ Expected marks per Q $= E(X) = \Sigma x p(x) = \frac{1}{4} \times 1 + \frac{3}{4} \times (-0.25) = \frac{1}{16}$

$\therefore$ Expected marks for 150 Questions (1 student)
$= 150 \times \frac{1}{16} = \frac{75}{8}$

$\therefore$ Expected marks for 1000 students
$= \frac{75}{8} \times 1000 = 9375$ marks

4[a] A coin is weighted so that its probability of landing on head is 20%. Suppose the coin is flipped 20 times. Find a bound for the probability, it lands on head at least 16 times. 2023 SE [2.5] [CO2]

$n = 20$

Binomial

$\therefore P(X \geq 16) = \sum_{k=16}^{20} \left({}^{20}C_k\right) 0.2^k (0.8)^{20-k}$

$= 1.38 \times 10^{-8}$

[b] i. Prove that $Var(ax+b) = a^2 Var(x)$ where a and b are non-zero constants. 2022E [2.5] [CO1]

let $y = ax + b$

$\therefore E(y) = E(ax+b) = E(ax) + E(b) = a E(x) + E(b)$

and $\bar{y} = a\bar{x} + b$

$\therefore Var(ax+b) = Var(y) = E(y - \bar{y})^2$
$= E(ax + \cancel{b} - a\bar{x} - \cancel{b})^2$
$= E(a(x - \bar{x}))^2$
$= E(a^2 (x - \bar{x})^2)$
$= a^2 E(x - \bar{x})^2$

$\boxed{Var(ax+b) = a^2 \; Var(x)}$

M.P.

3[a] Let X be a uniform random variable on the interval (C, D). Write and proof its expression for expectation and variance. 2023SE [2.5] [CO2]

3[a] Let X be a uniform random variable on the interval (C, D). Write and proof its expression for expectation and variance. [2.5] [CO2] 2022 M

given $X \sim U(C, D)$

mean

$$\therefore E[X] = \int_C^D \frac{x}{C-D}\, dx = \frac{1}{C-D}\int_C^D x\, dx = \frac{1}{C-D}\left[\frac{x^2}{2}\right]_C^D$$

$$= \frac{(C^2-D^2)}{2(C-D)}$$

$$\because \left[E[X] = \int_{x\in X} x f(x)\, dx\right.$$

$$\boxed{E[X] = \frac{C+D}{2}} \quad ; \text{ if } X \sim U(C,D)$$

variance

$$Var[X] = E[X^2] - [E[X]]^2$$

$$= \int_{x \in X} x^2 f(x)\, dx - \left[\frac{D+C}{2}\right]^2$$

$$\because \left[E[X^2] = \int_C^D x^2 \frac{1}{(C-D)}\, dx \right.$$

$$= \frac{1}{C-D}\left[\frac{x^3}{3}\right]_C^D = \frac{C^3-D^3}{3(C-D)} = \frac{C^2+CD+D^2}{3} \left.\right]$$

$$\therefore Var[X] = \frac{4(C^2+CD+D^2)}{4(3)} - \frac{(C^2+2CD+D^2).3}{4 \cdot 3}$$

$$= \frac{C^2-2CD+D^2}{12}$$

$$\boxed{Var[X] = \frac{(C-D)^2}{12}}$$

[c] The record of weights of male population follows normal distribution. Its means and standard deviation are 70 kg and 15 kg respectively. If a researcher considers the records of 50 males, then what would be the mean and standard deviation of the chosen sample? 2023 SE

Mean wt. of population, $\mu = 70kg$ (population mean)

Standard deviation of wt. of population, $\sigma = 15kg$ (population s.d.)

Sample size $= 50 = n$

Date : / /
Page No.

Now, for sample,

Sample mean, $\mu_{\bar{x}} = \mu \Rightarrow \boxed{\mu_{\bar{x}} = 70 \text{ kg}}$

sample s.d., $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{50}}$

$\Rightarrow \boxed{\sigma_{\bar{x}} = 2.12 \text{ kg}}$

[b] A random variable X is known to be Gaussian with a_x=1.6 and σ_x=0.4 Find: (a) P {1.4<X<=2.0} and (b) P {-0.6<(X-1.6) <=0.6} 2023 SE [2] [CO1]

given, $X \sim N(1.6, 0.4^2)$

a.) $P(1.4 < X <= 2.0) = P\left(\frac{1.4-1.6}{0.4} < \frac{X-1.6}{0.4} \leq \frac{2-1.6}{0.4}\right)$

$= P(-0.5 < Z < 1)$

$= \phi(1) - \phi(-0.5)$

$= \phi(1) - 1 + \phi(0.5)$

$= 0.8413 - 1 + 0.6915$

$= 0.5328$

b.) $P(-0.6 < (X-1.6) \leq 0.6) = P\left(\frac{-0.6}{0.4} < \frac{X-1.6}{0.4} \leq \frac{0.6}{0.4}\right)$

$= P(-1.5 < Z < 1.5)$

$= \phi(1.5) - \phi(-1.5)$

$= 2\phi(1.5) - 1$

$= 0.8664$

[c] A company sells tractors which fail at a rate of 1 out of 1000. If 500 tractors are purchased from this company, what is the probability of 2 of them failing within the first year? 2022 M [2] [CO1]

Poisson Dist.

$n = 500$

$p(\text{fail}) = 1/1000$.

$\lambda = np = 500 \times (1/1000) = 1/2$

prob of 2 failing,

$\therefore P(x=2) = (\lambda^x e^{-\lambda} / x!)$

$= \{e^{-1/2} \cdot (1/2)^2\} / 2!$

$= e^{-1/2}/8$

8. Assume that the time of arrival of birds at a particular place on a migratory route, as measured in days from the first of the year (January 1 is the first day), is approximated as a Gaussian random variable X with a_x=200 and σ_x=20 days.

Var = 20 × 20 = 400

2023 SE

[a] What is the probability the birds arrive after 160 days but on or before the 210th day?

[b] What is the probability the birds will arrive after the 231st day?

[5] [CO2]

Let X be the arrival day, $X \sim N[200, 400]$,

i) $P(160 < X \leq 210) = P\left(-2 \leq \frac{X-200}{20} \leq \frac{1}{2}\right)$

$= \phi(1/2) - \phi(-2) = \phi(1/2) - 1 + \phi(2)$

$= 0.6688$

ii) $P(X > 231) = 1 - P(X \leq 231)$

$= 1 - F_X(231)$

$= 1 - F\left(\frac{231-200}{20}\right) = 1 - F\left(\frac{31}{20}\right) = 1 - F(1.55)$

$= 0.0606$

[b] Let x be the life in days of a battery with probability density function

$F(x) = \{40000/x^3, \quad x > 100$
$\quad 0, \quad elsewhere\}$

2023 SE
2022 M

Find the expected life of the battery in hours. [2.5] [CO2]

given $F(x) = 4000/x^3, \quad x > 100$
$\quad\quad 0 \quad\quad elsewhere$

$\therefore$ expected battery life $= \int_0^\infty f(x) \cdot x\,dx = \int_0^{100} 0\,dx + \int_{100}^\infty \frac{4000}{x^3} \cdot x\,dx$
E[f(x)]

$= \int_{100}^\infty \frac{4000}{x^2}\,dx = +4000\left[\frac{x^{-1}}{-1}\right]_{100}^\infty$

$= 4000\left(\frac{-1}{\infty} + \frac{1}{100}\right)$

$= 40$ days

$= 40 \times 24$ hours $= 960$ hours

3[a] In a cafe, the customer arrives at a mean rate of 2 per min. Find the probability of arrival of 5 customers in 1 minute using the Poisson distribution formula. 2023 SM [2] [CO2]

Given, mean rate, $\lambda = 2$

hypothesised customer per min $\Rightarrow x = 5$

$$\therefore P(X = x) = \frac{(e^{-\lambda} \lambda^{x})}{x!}$$

$$\Rightarrow P(X=5) = \frac{(e^{-2}(2)^5)}{5!}$$

$$= 0.036$$

The probability of arrival of 5 customers per min is 3.6%.

1. Attempt any two of the following:

[a] A random variable has the following probability function 2022 E

X	1	2	3	4	5	6	7
P(X)	k	2k	2k	3k	k^2	$2k^2$	$7k^2+k$

Determine the value of the following:
i. k
ii. mean
iii. P(0<X<5) [5] [CO1]

i) $\Sigma P(x) = 1 \Rightarrow k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$

$\Rightarrow 9k + 10k^2 - 1 = 0 \Rightarrow 10k^2 + 9k - 1 = 0 \Rightarrow 10k^2 + 10k - k - 1 = 0$

$\Rightarrow 10k(k+1) - 1(k+1) = 0 \Rightarrow k = 1/10, -1 \text{ (not)} \therefore \boxed{k = \frac{1}{10}}$

ii) mean

expected value, $E(x) = 1 \times \frac{1}{10} + 2 \times \frac{2}{10} + 3 \times \frac{2}{10} + 4 \times \frac{3}{10} + 5 \times \frac{1}{100} + 6 \times \frac{2}{100} + 7 \times \left(\frac{7 \times 1}{100} + \frac{1}{10}\right)$

$$= \frac{10 + 40 + 60 + 120 + 50 + 120 + 7 \times 17}{100} =$$

$$= 5.19$$

iii) $P(0<x<5) = \frac{1}{10} + \frac{2}{10} + \frac{2}{10} + \frac{3}{10} = 0.8$

Date : / /
Page No.

2. Attempt any two of the following:

(a) If joint probability distribution of x, y are following:

$$F(x,y) = \begin{cases} 4xy & 0 < x < 1, 0 < y < 1 \\ 0 & otherwise \end{cases}$$

i. Marginal properties [2.5][CO2]

ii. Proof the independent properties of joint and marginal probability distribution. 2022E [2.5] [CO2]

i) Marginal distribution of X

$$f(x) = \int_y f(x,y)\, dy$$

$$= \int_0^1 4xy\, dy = 4x\left[\frac{y^2}{2}\right]_0^1$$

$$\boxed{f(x) = 2x}$$

Marginal distribution of Y

$$f(y) = \int_x f(x,y)\, dx$$

$$= \int_0^1 4xy\, dx = 4y\left[\frac{x^2}{2}\right]_0^1$$

$$\boxed{f(y) = 2y}$$

ii)

Date : / /
Page No.

[b] The average age at first marriage is 25 for women and 27.5 for men in US. If the standard deviation for women is 4 years. What is the probability that a random selection for 32 women have an average age at first marriage is between 26 and 27? 2023SE [2.5] [CO2]

here, $\mu = 25$ | $n = 32$
$\sigma = 4$

$\therefore$ for this sample sd, $s = \sigma/\sqrt{n} = 4/\sqrt{32} = 1/\sqrt{2}$

$\therefore P(26 < x < 27) = P(x < 27) - P(x < 26)$
$= F_x(27) - F_x(26)$
$= F\left(\frac{27-25}{1/\sqrt{2}}\right) - F\left(\frac{26-25}{1/\sqrt{2}}\right)$
$= F(2\sqrt{2}) - F(\sqrt{2})$
$= 0.9772 - 0.6915$
$= 0.2857$

$\therefore$ Probability is 28.57%.

Date : / /
Page No.

[c] Assume that the joint sample space S_j has only three possible elements: (1,1), (2,1), and (3,3) with probabilities of 0.2, 0.3, and 0.5 respectively. find [5] [CO3]

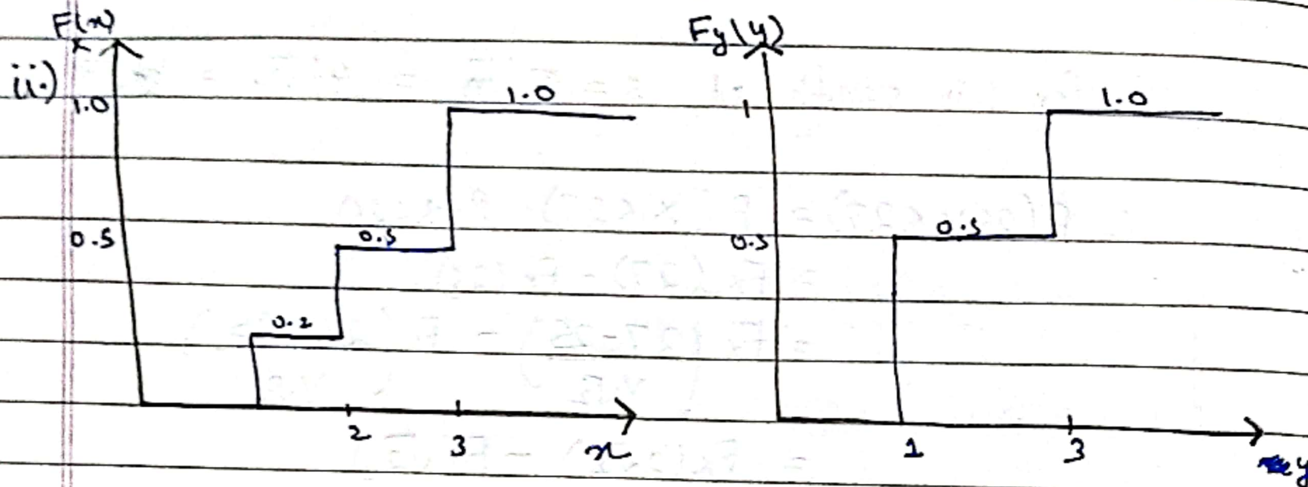
i. $F_{xy}(x, y)$
ii. Draw graph for joint distribution function
iii. Corresponding marginal distributions

2022 E

Unit 4

i.) $f_{x,y}(x,y) = 0.2\,u(x-1)\,u(y-1) + 0.3\,u(x-2)\,u(y-1) + 0.5\,u(x-3)\,u(y-3)$

ii.)



iii.) marginal distribution function

$F_x(x) = 0.2\,u(x-1) + 0.3\,u(x-2) + 0.5\,u(x-3)$

$F_y(x) = 0.2\,u(y-1) + 0.3\,u(y-1) + 0.5\,u(y-3) = 0.5\left(u(y-1) + u(y-3)\right)$

3. Attempt any two of the following:

[A] If X is normally distributed with mean = 12 & S.D. = 4. Find x' such that $P(X > x') = 0.24$ [5] [CO3]

2022E

$\mu = 12$
$\sigma = 4$

here, z-score : $z = \frac{x-\mu}{\sigma} \Rightarrow z.\sigma + \mu = x'$

for, $P(X > x') = 0.24 \Rightarrow P(X \leq x') = 1 - 0.24 = 0.76$

z-score for $0.76 = 0.675$

$\therefore x' = 0.675 \times 4 + 12 \Rightarrow \boxed{x' = 14.82}$

$\because P(x \leq x') = F(z)$

[b] Compute t test for the data given below:
Group A: 10, 4, 3, 2, 4, 2, 5, 10, 5, 5
Group B: 4, 6, 8, 2, 9, 1, 12, 13, 10, 10
Given: Critical value: 2.10 at 0.05 level of significance
2.88 at 0.01 level of significance

2023 SE
[2.5] [CO1]

For GROUP A

$n_1 = 10$

$\bar{x}_1 = \Sigma x_i / n = (10+4+3+2+4+2+5+10+5+5)/10 = 5$

for GROUP B

$n_2 = 10$

$\bar{x}_2 = \Sigma x_i / n = (4+6+8+2+9+1+12+13+10+10)/10 = 7.5$

x_1	$\lvert x_1 - \bar{x}_1 \rvert$	$(x_1 - \bar{x}_1)^2$	x_2	$\lvert x_2 - \bar{x}_2 \rvert$	$(x_2 - \bar{x}_2)^2$
10	5	25	4	3.5	12.25
4	6	36	6	1.5	2.25
3	7	49	8	0.5	0.25
2	8	64	2	5.5	30.25
4	6	36	9	1.5	2.25
2	8	64	1	6.5	42.25
5	5	25	12	4.5	20.25
10	0	0	13	5.5	30.25
5	5	25	10	2.5	6.25
5	5	25	10	2.5	6.25

$\Sigma (x_1 - x_1)^2 = 349$ $\quad\quad$ $\Sigma (x_2 - \bar{x}_{2i})^2 = 152.5$

$\therefore S_1 = 2.87$ $\quad\quad$ $S_2 = 4.12$

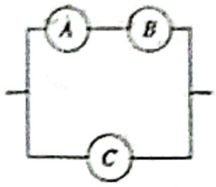
$\therefore$ for t-test

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{n}}} = \frac{2.5 \times 10}{\sqrt{(2.87)^2 + (4.12)^2}} = 1.576$$

$\therefore$ at 1% and 5% the t-test passes are they are of same sample.

4[a] The electrical apparatus in the diagram works so long as current can flow from left to right. The three components are independent. The probability that component A works is 0.8; the probability that component B works is 0.9; and the probability that component C works is 0.75. Find the probability that the apparatus works. [2.5] [CO2]

2022M



Given, $P(A) = 0.8$

$P(B) = 0.9$

$P(C) = 0.75$

$\therefore$ Apparatus works if A, B, C or A, B, or C or A, C or B, C

$\therefore P(\text{work}) = P(A)P(B)P(C) + P(A)P(B)P(\bar{C}) + P(\bar{A})P(\bar{B})P(C) + P(A)P(\bar{B})P(C) + P(\bar{A})P(B)P(C)$

$= 0.8 \times 0.9 \times 0.75 + 0.8 \times 0.9 \times 0.25 + 0.2 \times 0.9 + 0.8$

$+ 0.8 \times 0.1 \times 0.75 + 0.2 \times 0.1 \times 0.75$

$\boxed{P(\text{work}) = 0.93}$

[b] If P(X) =1/4, P(Y) =1/3 and P(XY) =1/12 then find P(Y|X), P(X|Y).

$P(X \cap Y) = 1/12$ 2022M [2.5] [CO1]

Give, $P(X) = 1/4$

$P(Y) = 1/3$

$P(X \cap Y) = 1/12$

$\therefore$ by conditional prob,

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{1/12}{1/3} = \frac{1}{4}$$

$\therefore \boxed{P(X|Y) = 1/4}$

$$P(Y|X) = \frac{P(X \cap Y)}{P(X)} = \frac{1/12}{1/4}$$

$\boxed{P(Y/X) = 1/3}$

Date : / /

Page No.

5. A company sells high-fidelity amplifiers capable of generating 10, 25, and 50 W of audio power. It has on hand 100 of the 10-W units of which 15% are defective, 70 of the 25-W units with 10% defective, and 30 of the 50-W units with 10% defective. 2022M

[a] What is the probability that an amplifier sold from the 10-W units is defective? [2.5] [CO1]

[b] What is the probability that a unit randomly selected for sale is defective? [2.5] [CO1]

Given, total = 100+70+30 = 200

no. of 10W units : 100 → 15% defective

no. of 25W units : 70 → 10% defective

no of 50W units : 30 → 10% defective

a.) P(10W defective) = 0.15 (Given)

b.) $P(\text{random defective}) = \frac{100}{200} \times \frac{15}{100} + \frac{70}{200} \times \frac{10}{100} + \frac{30}{200} \times \frac{10}{100}$

$P(\text{random def}) = 0.125$

$= 1/8$

Date : / /
Page No.

[c] Define central moment. Write and proof the relationship among second central moment and other n^{th} moments. [5] [CO1]

Moments about the mean value of X are called central moments and are given the symbol μ_n. They are defined as expected value of the function:

$$g(X) = (X-\bar{X})^n \qquad n=0,1,2,\ldots$$

which is

$$\mu_n = E[(X-\bar{X})^n] = \int_{-\infty}^{\infty} (x-\bar{X})^n f_X(x)\,dx$$

The moment $\mu_0 = 1$, the area of $f_X(x)$, while $\mu_1 = 0$.

Relation : $\boxed{\mu_2 = \mu_2' - {\mu_1'}^2}$ $\mu_2 \rightarrow$ Variance.

$$\therefore \sigma_X^2 = \mu_2 = E[(X-\bar{X})^2] = E[X^2 - 2\bar{X}X + \bar{X}^2]$$

$$= E[X^2] - 2\bar{X}E[X] + \bar{X}^2 \qquad [E(kX) = kE(X)]$$

$$= E[X^2] - \bar{X}^2 \qquad [\mu_1' = E(X-0) = \bar{X}]$$

$$= \mu_2' - {\mu_1'}^2 \qquad \because [E(X-0)^2 = \mu_2']$$

$$\therefore \boxed{\mu_2 = \mu_2' - {\mu_1'}^2}$$

[d] Write Chebychev's inequality and proof it. [5] [CO2]

For a random variable X with mean value $\bar{X}$ and variance σ_X^2, it states that

$$\boxed{P\{|X-\bar{X}| \geq \epsilon\} \leq \sigma_X^2/\epsilon^2}$$

or

$$\boxed{P\{|X-\mu| \leq \epsilon\} \geq 1 - \frac{\sigma^2}{\epsilon^2}}$$

for any $\epsilon > 0$.

proof:

we have, $P[|X-\mu| \leq \varepsilon] = \int_{\mu-\varepsilon}^{\mu+\varepsilon} f(x)\,dx$

$$\Rightarrow P[\overline{|X-\mu| \leq \epsilon}] = \int_{-\infty}^{\mu-\epsilon} f(x)\,dx + \int_{\bar{X}+\epsilon}^{\infty} f_X(x)\,dx = \int_{|x-\bar{X}|\geq\epsilon}^{\infty} f_X(x)\,dx$$

Date : / /
Page No

but

$$\sigma^2 = E[X - E[X]]^2$$
$$= E[X-\mu]^2$$
$$= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

So, $\sigma^2 = \int_{-\infty}^{\mu-\epsilon} (x-\mu)^2 f(x)\,dx + \int_{\mu-\epsilon}^{\mu+\epsilon} (x-\mu)^2 f(x)\,dx + \int_{\mu+\epsilon}^{\infty} (x-\mu)^2 f(x)\,dx$

non negative term

$$\sigma^2 \geq \int_{-\infty}^{\mu-\epsilon} (x-\mu)^2 f(x)\,dx + \int_{\mu+\epsilon}^{\infty} (x-\mu)^2 f(x)\,dx \quad -(1)$$

Now

$x \leq \mu - \epsilon \Rightarrow (x-\mu)^2 \geq \epsilon^2$

$x \geq \mu + \epsilon \Rightarrow (x-\mu)^2 \geq \epsilon^2$

from (1)

$$\sigma^2 \geq \int_{-\infty}^{\mu-\epsilon} \epsilon^2 f(x)\,dx + \int_{\mu+\epsilon}^{\infty} \epsilon^2 f(x)\,dx \geq \epsilon^2 \left[\int_{-\infty}^{\mu-\epsilon} f(x)\,dx + \int_{\mu+\epsilon}^{\infty} f(x)\,dx\right]$$

$$\geq \epsilon^2 \left[1 - \int_{\mu-\epsilon}^{\mu+\epsilon} f(x)\,dx\right] \qquad [\because P(\bar{X}) = P - P(X)]$$

$$\sigma^2 \geq \epsilon^2 \left[1 - P[|X-\mu| \leq \epsilon]\right] \qquad [\because \text{given}]$$

$$\Rightarrow \frac{\sigma^2}{\epsilon^2} \geq 1 - P[|X-\mu| \leq \epsilon]$$

or

$$\boxed{P[|X-\mu| \geq \epsilon] \leq \frac{\sigma^2}{\epsilon^2}}$$

Hence proved.

Date : / /
Page No

iii. What is central limit theorem and what is the significance of it? E [2.5] [CO1]

CLT is an important result in statistics, which states that the normal distribution is the limiting distribution to the sum of the independent random variables as the no. of RV get indefinitely large.

if X_i $(i=1,2,\ldots n)$ is independently distributed random variable such that $E(X_i) = \mu_i$ and $Var(X_i) = \sigma^2$

then, as $n \to \infty$, the distribution of sum of these

$$S_n = X_1 + X_2 + \ldots + X_n$$

tends to normally distributed $N(\mu, \sigma)$

where $\mu = \sum_{i=1}^{n} \mu_i$, $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$

Significance:

CLT is significant because it allows us to make inferences about populations based on samples.

Ex) Height of all adults in a country can't be measured, so we take a sample of adults and use CLT to estimate population mean based on sample mean.

Its also used in many statistical tests such as t-test, ANOVAs and linear regression.

Date : / /
Page No.

[c] Explain different types of error in hypothesis testing with proper notations by taking a short example. E [5] [CO5]

Different types of error in hypothesis testing are :

TYPE 1 ERROR
In a statistical hypothesis testing experiment, a type-1 error is committed by rejecting the null hypothesis even when its true. It is denoted by 'α'.

Its often called a false positive as an event shows that a given condition is present when its absent.

TYPE 2 ERROR
In a statistical hypothesis testing experiment, a type-2 error is committed by not rejecting the null hypothesis even when its false. It is denoted by 'β'.

Its often called false negative as a real hit was rejected by the test and is observed as a miss.

$\therefore$

	Accept H_0	Reject H_0
H_0 is True	Correct $P = 1-\alpha$	Type 1 error $P = \beta$
H_0 is False	Type 2 error $P = \alpha$	Correct $P = 1-\beta$

Eg.) H_0: A man is guilty of a crime.

type 1 error: He is condemned to crime, though he is not guilty or committed the crime.

type 2 error: He is ~~con~~ condemned not guilty when the court actually does commit the crime by letting the guilty one go free.

5. Attempt any two of the following:

[a] The means of two single large samples of 1000 & 2000 members are 67.5 and 68.0 inches respectively. Can the sample be regarded as drawn from the same population of s.d. 2.5 inches, test at α=.05? [5] [CO5]

Given,

$\bar{x}_1 = 67.5$, $n_1 = 1000$ | $\sigma_1 = 2.5$

$\bar{x}_2 = 68.0$, $n_2 = 2000$ | $\sigma_2 = 2.5$

let null hypothesis H_0 be $\mu_1 = \mu_2$

alternative hypo. H_1 be $\mu_1 \neq \mu_2$

Using studen's z test

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{67.5 - 68.0}{(2.5)\sqrt{\frac{1}{1000} + \frac{1}{2000}}} = \frac{-0.5 \times \cancel{10}^{4} \times \sqrt{20}}{\cancel{2.5}\sqrt{3}}$$

$$= \frac{-2 \times 2\sqrt{5}}{\sqrt{3}} = -5.16$$

$\therefore |z| = 5.16$

at $\alpha = 0.05$, i.e 5% level of significance, table value = 1.96

$\therefore$ Calc Value of z > table value of z

$\therefore$ Hypothesis Reject $\Rightarrow \mu_1 \neq \mu_2 \Rightarrow$ They are from different population

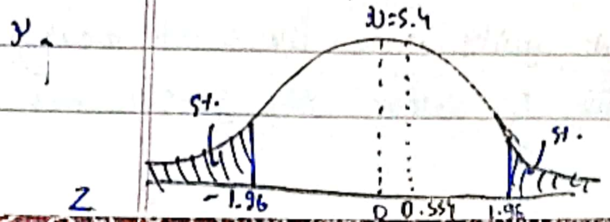
[b] A random sample of 50 items gives the mean 6.2 and variance 10.24. Can it be regarded as drawn from population with mean 5.4 at 5% of level of significance? Show your answer by using curve with proper notations of symbols. [5] [CO5]

Given, $n = 50$

pop mean, $\mu = 5.4$

Sample mean, $\bar{x} = 6.2$

Sample var, $S = 10.24$



null hypo, $H_0 \Rightarrow 5.4 = \mu$ | $H_1 = \mu \neq 5.4$

Using test statistic about mean

$$z = \frac{(\bar{x} - \mu)}{S/\sqrt{n}} = \frac{6.2 - 5.4}{10.24/\sqrt{50}} \approx 0.554$$

at 5% level of significance, $Z_\alpha = 1.96$

$\therefore Z < Z_\alpha \Rightarrow$ Hypo H_0 is accepted

Yes, its drawn from that population.

Date : / /
Page No.

4. Attempt any two of the following:

[a] A manufacturer claimed that at least 98% of the steel pipes which he supplied to a factory conformed to the specifications. An examination of a sample of 500 pieces of pipes revealed that 30 were defective. Test this claim at a significance level of (i) 0·05, (ii) 0·01. [5] [CO4]

Given,

$\boxed{\text{test proportion} = 0.98 = p}$ Sample size, $n = 500$

defective sample pieces = 30

$\therefore$ sample proportion, $\hat{p} = \frac{470}{500} = 0.94$

let H_0 : null test hypothesis $\Rightarrow p \geq 98\%$.

H_1 : alt hypothesis $\Rightarrow p < 98\%$.

now, using Z statistic for normal diff b/w two proportion.

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.94 - 0.98}{\sqrt{\frac{0.98 \times 0.02}{500}}} = -6.389$$

at i) 5% lvl of significance.

$Z_\alpha = 1.960$

$\therefore |Z| > Z_\alpha \rightarrow$ Reject H_0

ii) 1% lvl of significance.

$Z_\alpha = 2.576$

$|Z| > Z_\alpha \rightarrow$ Reject H_0

Claim is false.

[c] The U.S. Bureau of census wishes to estimate the birth rate per 1, 00,000 people in the nation's largest cities. It is known that the standard deviation in the birth rates for these 100 urban centers is 12 births per 1, 00,000 people. Then (a) Calculate the variance and standard error of the sampling distribution of (i) n=8 cities (ii) n=15 cities. [5] [CO4]

a) i) here,

$N = 100$

$n = 8$

$\sigma = 12$

Since $\frac{n}{N} = 0.08 < 0.1$, we use the following formula for sample sd

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

$$\therefore \sigma_{\bar{x}}^2 = \frac{12^2}{8} = 18$$

and standard error is

$\sigma_{\bar{x}} = \sqrt{18} = 4.24.$

ii.) here,

$N = 100$

$n = 15$

$\sigma = 12$

Since $\frac{n}{N} = 0.15 > 0.1$, we use the following formula for sample sd

$$\sigma_{\bar{x}}^2 = \left(\frac{N-n}{N-1}\right)\frac{\sigma^2}{n}$$

$$\therefore \sigma_{\bar{x}}^2 = \left(\frac{100-15}{100-1}\right)\frac{(12)^2}{15}$$

$$= 8.24$$

$\therefore$ std error,

$$\sigma_{\bar{x}} = \sqrt{8.24} = 2.87$$

b.) compare the values obtained in both cases,

on comparing both the values, we observe that, the larger sample has a smaller standard error and will tend to result in less sampling error in estimating the birth rates in the 100 cities.

[b] What is memorylessness property and which distribution contains this property. Prove it for any of two. [5] [CO3]

The memory less property means that a given probability distribution is independent of its history. It refers to the independence of events or event-to-event times.

The following distributions contain this property.

→ Geometric Distribution

→ Exponential Distribution

for GEOMETRIC DISTRIBUTION

If X has a geometric distribution, show that for any two positive integers s and t,

$$\boxed{P(X > s+t \mid X > s) = P(X > t)}$$

PROOF:

Since, X is geometric dist, we have

$$P(x) = q^{x-1}p \quad , x = 1, 2, 3 \ldots$$

Date : / /
Page No.

for any positive integer k, we have

$$P(X>k) = P(X=k+1) + P(X=k+2) + \cdots$$
$$= q^k p + q^{k+1} p + \cdots$$
$$= q^k p\,[1+q+q^2+\cdots]$$
$$= q^k p\left(\frac{1}{1-q}\right) = q^k \cancel{p}\,\frac{1}{\cancel{p}}$$
$$= q^k \quad \because (q+p=1) \quad \text{--- (1)}$$

Hence,

$$P(X>s+t \mid X>s)$$
$$= \frac{P((X>s+t)\cap(X>s))}{P(X>s)} \quad \text{using BT}$$
$$= \frac{P(X>s+t)}{P(X>s)}$$
$$= \frac{q^{\cancel{s}+t}}{q^{\cancel{s}}} \quad \text{using (1)}$$
$$= q^t$$
$$= P(X>t)$$

Hence proved

[b] Explain Z-test and Chi-squared test with proper example. [5] [CO4]

Chi-square Test

The χ^2 test is one of the simplest and most widely used non-parametric tests in statistical work. It makes no assumptions about the population being sampled. The quantity χ^2 describes the magnitude of discrepancy between theory and observation, *i.e.*, with the help of χ^2 test we can know whether a given discrepancy between theory and observation can be attributed to chance or whether it results from the inadequacy of the theory to fit the observed facts. If χ^2 is zero, it means that the observed and expected frequencies completely coincide. The greater the value of χ^2, the greater would be the discrepancy between observed and expected frequencies. The formula for computing chi-square is :

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

where, O = observed frequency

E = expected or theoretical frequency.

The calculated value of χ^2 is compared with the table value of χ^2 for given degrees of freedom at specified level of significance. If the calculated value of χ^2 is greater than the table value, the difference between theory and observation is considered to be significant, *i.e.*, it could not have arisen due to fluctuations of simple sampling. On the other hand, if the calculated value of χ^2 is less than the table value, the difference between theory and observation is not considered significant, *i.e.*, it could have arisen due to fluctuations of sampling.

Z-test

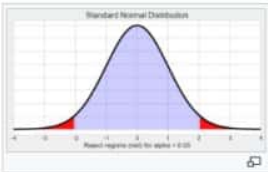
18 languages

Article Talk Read Edit View history Tools

From Wikipedia, the free encyclopedia

"Z test" redirects here. For the "Z-test" procedure in the graphics pipeline, see Z-buffering.

A **Z-test** is any statistical test for which the distribution of the test statistic under the null hypothesis can be approximated by a normal distribution. Z-test tests the mean of a distribution. For each significance level in the confidence interval, the Z-test has a single critical value (for example, 1.96 for 5% two tailed) which makes it more convenient than the Student's *t*-test whose critical values are defined by the sample size (through the corresponding degrees of freedom). Both the Z-test and Student's t-test have similarities in that they both help determine the significance of a set of data. However, the z-test is rarely used in practice because the population deviation is difficult to determine.



Applicability [edit]

Because of the central limit theorem, many test statistics are approximately normally distributed for large samples. Therefore, many statistical tests can be conveniently performed as approximate Z-tests if the sample size is large or the population variance is known. If the population variance is unknown (and therefore has to be estimated from the sample itself) and the sample size is not large ($n < 30$), the Student's *t*-test may be more appropriate (in some cases, $n < 50$, as described below).

Procedure [edit]

How to perform a Z test when T is a statistic that is approximately normally distributed under the null hypothesis is as follows:

First, estimate the expected value μ of T under the null hypothesis, and obtain an estimate s of the standard deviation of T.

Second, determine the properties of T : one tailed or two tailed.

For Null hypothesis H_0: $\mu \geq \mu_0$ vs alternative hypothesis H_1: $\mu < \mu_0$, it is lower/left-tailed (one tailed).

For Null hypothesis H_0: $\mu \leq \mu_0$ vs alternative hypothesis H_1: $\mu > \mu_0$, it is upper/right-tailed (one tailed).

For Null hypothesis H_0: $\mu = \mu_0$ vs alternative hypothesis H_1: $\mu \neq \mu_0$, it is two-tailed.

Third, calculate the standard score:

$$Z = \frac{(\bar{X} - \mu_0)}{\sigma},$$

which one-tailed and two-tailed *p*-values can be calculated as $\Phi(Z)$(for lower/left-tailed tests), $\Phi(-Z)$ (for upper/right-tailed tests) and $2\Phi(-|Z|)$ (for two-tailed tests) where Φ is the standard normal cumulative distribution function

[c] Explain following by taking proper example [5] [CO2]

i. Estimator
ii. Estimate
iii. Point estimation
iv. Criteria for a good estimator

Date : / /
Page No.

E

i) ESTIMATOR

An estimator is a measure/rule for calculating an estimate of a given quantity based on observed data. It is a statistic i.e. a function which takes a sample of data as input and produces an estimate for population as output.

ii) ESTIMATE

It is the value given by the estimator for our sample. It is an approximation of the true value of a population parameter.

iii.) POINT ESTIMATION

It is the process by which we choose an estimator and find the single value of an estimator for estimating an unknown parameter,

Eg.) Sample grades of a population are 80, 85, 90, 95, 95, 90, 90, 90, 85, 80

∴ ESTIMATOR – sample mean

ESTIMATE / POINT ESTIMATE – 88

iv.) CRITERIA FOR GOOD ESTIMATOR

Some of the criteria to judge the quality of an estimator are:

BLUE

1.) UNBIASEDNESS: An unbiased estimator is one that has an expected value that is equal to the true value of the population parameter. , $E(\theta) = \theta$

2.) CONSISTENCY: A consistent estimator is one that gets closer and closer to the true value ~~dsp~~ as sample size increases.

3.) EFFICIENCY: An efficient estimator is one that has a smaller variance than other

4.) SUFFICIENCY: convey as much info as possible

Total No. of Pages 2 | *Roll No. ......*

FIFTH SEMESTER | **B.Tech. [CSE]**

MID SEMESTER EXAMINATION | ***September-2023***

CO325 Probability and Statistics

Time: 1:30 Hours | *Max. Marks: 25*

Note: Answer ***ALL*** questions.
Assume suitable missing data, if any.
CO# is course outcome related to question

1[a] A and B are two events such that $P(A) = 0.54$, $P(B) = 0.69$ and $P(A \cap B) = 0.35$.
Find (i) $P(A' \cap B')$ (ii) $P(A \cap B')$ [2] [CO1]

It is given that $P(A)=0.54, P(B)=0.69, P(A\cap B)=0.35$

(i) We know that $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

$\therefore P(A\cup B)=0.54+0.69-0.35=0.88$

(ii) $A'\cap B'=(A\cup B)'$, [by De Morgan's law]

$\therefore P(A'\cap B')=P(A\cup B)'=1-P(A\cup B)=1-0.88=0.12$

Find (i) $P(A' \cap B')$ (ii) $P(A \cap B')$

[b] Three coins are tossed once. Let A denote the event 'three heads show", B denote the event "two heads and one tail show", C denote the event" three tails show and D denote the event 'a head shows on the first coin". Which events are
(i) mutually exclusive? (ii) simple? (iii) Compound? [3] [CO2]

When three coins are tossed the sample space is given by,
S={HHH,HHT,HTH,HTT,THH,THT,TTH,TTT}
A={HHH}
B={HHT,HTH,THH}
C={TTT}
D={HHH,HHT,HTH,HTT}

$A\cap B=\phi$
$A\cap C=\phi$
$A\cap D=\{HHH\}\neq\phi$
$B\cap C=\phi$
$B\cap D=\{HHT,HTH\}\neq\phi$
$C\cap D=\phi$

(i) Event A and B; event A and C; event B and C; and event C and D are all mutually exclusive.
(ii) If an event has only one sample point of a sample space, it is called a simple event.
Thus A and C are simple events.
(iii) If an event has more than one sample point of a sample space it is called a compound event.
Thus B and D are compound events

2[a] A bag contains 6 red, 5 blue, and 4 yellow balls. 2 balls are drawn, but the first ball is drawn without replacement. Find the following.
P (red, then blue) [2] [CO1]

Let $P(\text{red})$ be the probability of drawing a red item, and $P(\text{blue}|\text{red})$ be the probability of drawing a blue item given that a red item has already been drawn (without replacement).

$$P(\text{red}) = \frac{6}{15}$$
$$P(\text{blue}|\text{red}) = \frac{5}{14}$$

To find the probability of both events occurring (red and then blue), we multiply the probabilities:

$$\begin{aligned} P(\text{red and blue}) &= P(\text{red}) \times P(\text{blue}|\text{red}) \\ &= \frac{6}{15} \times \frac{5}{14} \\ &= \frac{30}{210} \end{aligned}$$

Simplifying the fraction:

$$\frac{30}{210} = \frac{1}{7}$$

Therefore, the probability of drawing a red item followed by a blue item is $\frac{1}{7}$ or a 1 in 7 chance.

[b] Two students Anil and Ashima appeared in an examination. The probability that Anil will qualify the examination is 0.05 and that Ashima will qualify the examination is 0.10. The probability that both will qualify the examination is 0.02. Find the probability that
(i) Both Anil and Ashima will not qualify the examination.
(ii) Atleast one of them will not qualify the examination and
(iii) Only one of them will qualify the examination. [3] [CO1]

Let E be the event that Anil will qualify the examination and F be the event that Ashima will qualify the examination.
Given that:-
Probability that Anil will qualify the exam =P(E)=0.05
Probability that Ashima will qualify the exam =P(F)=0.10
Probability that both will qualify the examination =0.02
∴P(E∩F)=0.02
To find:-
(a) P(both Anil and Ashima will not qualify the examination)=?
(b) P(atleast one of them will not qualify)=?
(c)P(only one of them will qualify)=?
Solution:-
As we know that,
P(E∪F)=P(E)+P(F)−P(E∩F)
⇒P(E∪F)=0.05+0.1−0.02=0.13

- (a) P(both Anil and Ashima will not qualify the examination)

By Demorgan's law,

$P(E'\cap F')=P(E\cup F)'=1-P(E\cup F)$

$\therefore P(E'\cap F')=1-0.13=0.87$

- (b) P(atleast one of them will not qualify)$=1-P(E\cap F)=1-0.02=0.98$
- (c) P(only one of them will qualify)

$P(E\cap F')\cup P(E'\cap F)=P(E\cap F')+P(E'\cap F)-P(E\cap F')\cap P(E\cap F')$

$\Rightarrow P(E\cap F')\cup P(E'\cap F)=P(E\cap F')+P(E'\cap F)[\because P(E\cap F')\cap P(E'\cap F)=0]$

$\Rightarrow P(E\cap F')\cup P(E'\cap F)=P(E)-P(E\cap F)+P(F)-P(E\cap F)$

$\Rightarrow P(E\cap F')\cup P(E'\cap F)=P(E)+P(F)-2P(E\cap F)$

$\Rightarrow P(E\cap F')\cup P(E'\cap F)=0.05+0.1-2(0.02)=0.11$

3[a] Let X be a uniform random variable on the interval (A, B). Write and proof its expression for expectation and variance. [2.5] [CO2]

[b] Let x be the life in days of a battery with probability density function

$$F(x) = \{50000/x^3\,, \quad x > 100 \\ 0\,, \quad elsewhere\}$$

Find the expected life of the battery in hours. [2.5] [CO2]

4[a] Two groups are competing for the position on the Board of directors of a corporation. The probabilities that the first and the second groups will win are 0.6 and 0.4 respectively. Further, if the first group wins, the probability of introducing a new product is 0.7 and the corresponding probability is 0.3 if the second group wins. Find the probability that the new product introduced was by the second group. [3] [CO2]

Let E_1 and E_2 be the respective events that the first group and the second group win the competition. Let A be the event of introducing a new product.

$P(E_1)$ = Probability that the first group wins the competition = 0.6

$P(E_2)$ = Probability that the second group wins the competition = 0.4

$P(A|E_1)$ = Probability of introducing a new product if the first group wins = 0.7

$P(A|E_2)$ = Probability of introducing a new product if the second group wins = 0.3

The probability that the new product is introduced by the second group is given by $P(E_2|A)$.

By using Baye's theorem, we obtain

$$P(E_2|A) = \frac{P(E_2)\cdot P(A|E_2)}{P(E_1)\cdot P(A|E_1) + P(E_2)\cdot P(A|E_2)}$$

$$= \frac{0.4 \times 0.3}{0.6 \times 0.7 + 0.4 \times 0.3}$$

$$= \frac{0.12}{0.42 + 0.12}$$

$$= \frac{0.12}{0.54}$$

$$= \frac{12}{54}$$

$$= \frac{2}{9} = 0.22$$

[b] A lot has 10% defective items. 10 items are randomly selected for this lot. The probability that exactly two of chosen items are defective.

[2] [CO1]

Concept:

The probability that exactly 2 of the chosen items are defective is given as,

By Binomial distribution,

$$P(x = 2) \; = \; {}^{n}C_{x} \; (p)^{x} \; q^{(n-x)}$$

where, p = probability of success, q = probability of failure

Calculation:

Given:

n = 10, x = 2, p = 0.1, q = 0.9

Therefore, P(exactly 2 of the chosen items are defective) = ${}^{10}C_2 \times 0.1^2 \times 0.9^{10-2} \Rightarrow 45 \times 0.01 \times 0.43 = 0.1937$

5[a] A random variable has the following probability function

X	1	2	3	4	5	6	7
P(X)	k	2k	2k	3k	k^2	$2k^2$	$7k^2+k$

Determine:- i) k ii) mean iii) P(0<X<5) [3] [CO1]

[b] A gambler flips a coin 3 times. Draw a sample space S f... this experiment. A random variable X representing his winnin... below. He losses $1 if he gets no heads in three flips. He ... $3 if he obtains 1, 2 or 3 heads respectively. Show how el... to values of X.

2K21/CO/262

Total No. of Pages 3 *Roll No. ......*

FIFTH SEMESTER **B.Tech. [CSE]**

END SEMESTER EXAMINATION ***December-2023***

CO325 Probability and Statistics

Time: 3 Hours *Max. Marks: 50*

Note: Answer *ALL* questions.
Assume suitable missing data, if any.
CO# is course outcome related to question

1. Attempt any two all

[a] The time elapsed, in minutes, between the placement of an order of pizza and its delivery is random with the density function

$$f(x)=\begin{cases}\frac{1}{15} & \text{if } 25 < x < 40\\ 0 & \text{Otherwise}\end{cases},$$

(i) Determine the mean and standard deviation of the time it takes for the pizza shop to deliver pizza.
(ii) Suppose that it takes 12 minutes for the pizza shop to bake pizza. Determine the mean and the standard deviation of the time it takes for the delivery person to deliver pizza. [5] [CO1]

[b] Let the conditional probability density function of X, given that Y=y, be

$$f_{X|Y}(x|y) = \frac{x+y}{1+y}e^{-x}, \quad 0<x<\infty,\ 0<y<\infty$$

Find P(X<1 | Y=2). [5] [CO4]

[c] Define central moment. Write and proof the relationship among second central moment and other n^{th} moments. [5] [CO1]

2. Attempt any two

[a] Let the joint probability mass function of random variables X and Y be given by

$$p(x,y)=\begin{cases}\frac{1}{70}x(x+y) & \text{if } x = 1,2,3, \quad y = 3,4\\ 0 & \text{elsewhere}\end{cases}$$

Find E(X) and E(Y). [5] [CO2]

[b] For the scores on an achievement test given to a certain population of students, the expected value is 500 and the standard deviation is 100. Let X be the mean of the scores of a random sample of 10 students. Find a lower bound for P (460<X<540). Also, proof Chebyshev's inequality. [5] [CO2]

[c] Obtain the equation of the normal curve that may be fitted to the following data; [5] [CO2]

Class	60-65	65-70	70-75	75-80	80-85	85-90	90-95	95-100
frequency	3	21	150	335	326	135	26	4

3. [a] The fill amount of soft drink bottle is normally distributed with a mean of 2.1 liters and a standard deviation of 0.1 L. Bottles that contain less than 95% of the listed net content (2 L in this case) will be rejected by manufacturer quality check. Bottles that contain more than 120 percent of listed net content may cause excess spillage on opening.
Write in detail with formulae and calculation the following [5] [CO3]

[i] What is the proportion of soft drink bottles that will spill out of the total bottles manufactured?
[ii] What proportion of bottles will be rejected by the manufacturer's quality check
[iii] What proportion of bottles will contain exactly 2 L?

[b] It is claimed that the mean of the population is 67 at 5% level of significance. Mean obtained from a random sample of size 100 is 64 with SD 3. Validate the claim. [5] [CO4]

4. Attempt any two
[a] (i) An urn contains 10 white and 12 red chips. Two chips are drawn at random and, without looking at their, are discarded. What is the probability that a third chip drawn is red? [3] [CO5]
(ii) An elevator with two passengers stops at the second, third, and fourth floors. If it is equally likely that a passenger gets off at any of the three floors. What is the probability that the passengers get off at different floors? [2] [CO2]

[b] A poker-dealing machine is supposed to deal cards at random, as if from an infinite deck. In a test, you counted 1600 cards, and observed the following:

Spades 404
Hearts 420
Diamonds 400
Clubs 376

Could it be that the suits are equally likely? Or are these discrepancies too much to be random. Critical value at 0.05 level of significance is 7.815. [5] [CO4]

[c] A cigarette manufacturing firm claims that its brand A out sells brand B by 8%. If it is found that 42 out of 200 smokers prefer brand A, and 18 out of 100 smokers prefers brand B. Test at 5% level of significance, whether the 8% difference is a valid claim. [5] [CO4]

5. Attempt any two

[a] A company wants to improve its sales. The previous sales data indicated that the average sale of 25 salesmen was \$50 per transaction. After training, the recent data showed an average sale of \$80 per transaction. If the standard deviation is \$15, find the t-score. Has the training provided improved the sales? Let the critical value at 0.05 level of significance is 1.711. [5] [CO5]

[b] In a sample of 1000 people in Maharashtra , 540 are rice eaters and rest are wheat eaters. Can we assume that both rice and wheat are equally popular in this state a 1% level of significance? [5][CO5]

[c] (i) What is parametric and non-parametric test? [5][CO4]

(ii) Differentiate between t-test and chi-squared test.

(iii) Write a note on null hypothesis and alternate hypothesis.