PMAT501L - Probability and Statistics

Slot : E2-TE2

Digital Assignment - 2

Q1) Find the equation of the multiple regression plane of Z on X and Y from the following data.

X	30	40	20	50	60	40	20	60
Y	11	10	7	15	19	12	8	14
Z	110	80	70	120	150	90	70	120

Solution

X	Y	Z	X^2	Y^2	XY	ZX	ZY
30	11	110	900	121	330	3300	1210
40	10	80	1600	100	400	3200	800
20	7	70	400	49	140	1400	490
50	15	120	2500	225	750	6000	1800
60	19	150	3600	361	1140	9000	2850
40	12	90	1600	144	480	3600	1080
20	8	70	400	64	160	1400	560
60	14	120	3600	196	840	7200	1680
$\Sigma X =$ 320	$\Sigma Y =$ 96	$\Sigma Z =$ 810	$\Sigma X^2 =$ 14600	$\Sigma Y^2 =$ 1260	$\Sigma XY =$ 4240	$\Sigma ZX =$ 35100	$\Sigma ZY =$ 10470

$n = 8$

For a, b, c

$$\Sigma Z = na + b\Sigma X + c\Sigma Y$$

$$\Sigma ZX = a\Sigma X + b\Sigma X^2 + c\Sigma XY$$

$$\Sigma ZY = a\Sigma Y + b\Sigma XY + c\Sigma Y^2$$

$$810 = 8a + 320b + 96c$$

$$35100 = 320a + 14600b + 4240c$$

$$10470 = 96a + 4240b + 1260c$$

$$a = \frac{2895}{172} \approx 16.83$$

$$b = \frac{-21}{86} \approx -0.244$$

$$c = \frac{675}{86} \approx 7.85$$

$$Z = 16.83 - 0.244X + 7.85Y$$

Q2) It is known that probability of an item produced by a certain machine will be defective is 0.10. If the produced items are sent to the market in packet of 50, then find the no. of packets containing at least, exactly and at most 5 defective items in a consignment of 1000 packets by using (i) Binomial Distribution (ii) Poisson Approximation to the Binomial Distribution.

Solution

$P = 0.10 \qquad n = 50 \qquad N = 1000$

$$P(X=k) = {}^{n}C_{k}\, p^{k} (1-p)^{n-k}$$

Exactly 5, $P(X=5) = {}^{50}C_{5} (0.10)^{5} (0.90)^{45}$

$$= \frac{50!}{45!\,5!} (0.10)^{5} (0.90)^{45}$$

$$= 2118760 \times 0.00001 \times 0.04077 \approx 0.1843$$

Consignment $= 0.1843 \times 1000 = 184.3 \simeq 184$ packets

Atmost 5, $P(X \leq 5)$

$$= P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5)$$

$$= {}^{5}C_{0} p^{0} q^{5} + {}^{5}C_{1} p^{1} q^{4} + {}^{5}C_{2} p^{2} q^{3} + {}^{5}C_{3} p^{3} q^{2} + {}^{5}C_{4} p^{4} q^{1} + {}^{5}C_{5} p^{5} q^{0}$$

$$= 0.0052 + 0.0286 + 0.0779 + 0.1386 + 0.1809 + 0.1849$$

$P(X \leq 5) \Rightarrow 0.6161 \times 1000 = 616.1 \simeq 616$ packets

$$P(X \geq 5) = 1 - P(X < 5) = 1 - P(X \leq 4)$$

$$= 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)]$$

$$= 1 - [0.0045 + 0.0274 + 0.0822 + 0.1645 + 0.2468]$$

$$= 1 - 0.5254$$

$$= 0.4746$$

For $N = 1000$, $\quad 0.4746 \times 1000 = 474.6 \simeq 475$ packets

Poisson Approximation

$\lambda = 5 \qquad N = 1000$

Atleast 5, $\quad P(X \geq 5) = 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)]$

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

$$P(X \geq 5) = 1 - [0.0067 + 0.0337 + 0.0842 + 0.1404 + 0.1755]$$

$$= 0.56$$

for $N = 1000$, $\quad 0.56 \times 1000 = 560$ packets

Exactly 5, $\quad P(X=5) = \dfrac{5^5 e^{-5}}{5!} = 0.1755$

For $N = 1000$, $\quad 0.1755 \times 1000 = 175.5 \simeq 176$ packets

Atmost 5, $\quad P(X \leq 5) = P(X=0) + P(X=1) + P(X=2) = P(X=3) + P(X=4) + P(X=5)$

$$= 0.0067 + 0.0337 + 0.0842 + 0.1404 + 0.1755 = 0.4395$$

For $N = 1000$,

$0.4395 \times 1000 = 439.5 \simeq 440$ packets

Q3) The finish times for marathon runners will a race normally distributed with a mean of 200 minutes and a standard deviation of 50 min

i) What is probability of the runner will complete the marathon within 3 hours?

ii) What is the probability of the runners will complete the marathon between 3 hours and 4 hours?

iii) What is the probability that a runners will complete marathon after 4 hours?

iv) Calculate to the nearest minute, the time by which the first 8% runners have completed the marathon.

$$Z = \frac{X - \mu}{\sigma} \qquad \mu = 200$$

$$\sigma = 50$$

(i) $Z = \frac{3 \times 60 - 200}{50} = \left| \frac{-20}{50} \right| = 0.4$

$P(X \leq 180) = P(Z \leq 0.4) = 0.3446$

(ii) $Z = \frac{240 - 200}{50} = \frac{40}{50} = 0.8$

$P(180 \leq X \leq 240) = P(Z = 0.8) - P(Z = 0.4)$

$= 0.7881 - 0.3446$

$= 0.4435$

Percentage, $0.4435 \times 100 = 44.35\%$

iii) $P(X \leq 240) = 0.7881$ (we know from above)

$P(X \geq 240) = 1 - P(X \leq 240)$

$= 1 - 0.7881 = 0.2119$

Percentage, 21.19%

iv) From Z table

$X = \mu + Z\sigma$

$= 200 + (-1.405) \times 50$

$= 200 - 70.25$

$= 129.75$ min

So, first 8% runners will complete in 130 min approx.

(Q4) The student welfare office of a certain university polled a random sample of 1000 male students and found that 720 were in favor of a new grading system. At the same time, 695 out of a random sample of 900 female students were in favour of the new system. Do who favour the new grading system at 95% level of confidence?

Male Students $n_1 = 1000$ $\quad x_1 = 720$ $\quad P_1 = \frac{720}{1000}$

Female Students $n_2 = 900$ $\quad x_2 = 695$ $\quad P_2 = \frac{695}{900}$

$$P_1 = 0.72 \, , \quad P_2 = 0.7722$$

Null Hypothesis, $H_0 : P_1 = P_2$ (No Difference)

Alternative Hypothesis, $H_1 : P_1 \neq P_2$ (2 tail test)

$$Z = \frac{(p_1 - p_2)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{1415}{1900} \simeq 0.7447$$

$$Z = \frac{0.72 - 0.7722}{\sqrt{0.7447(1 - 0.7447)\left(\frac{1}{1000} + \frac{1}{900}\right)}}$$

$$Z = \frac{0.72 - 0.7722}{0.02003} \approx -2.606$$

For, Two-tailed test at 95% confidence level the critical Z value is ± 1.96

So, -2.606 that falls outside the range

$$-1.96 \leq Z \leq 1.96$$

$\therefore$ Z is less than -1.96

we reject null hypothesis

Q5) Write the detailed report on Applications of Probability Distributions:

Introduction to Probability Distributions

A probability distribution describes how the values of a random variable are distributed. In statistical analysis, it provides the likelihood that a given outcome will occur. Probability distributions are foundational in statistics, with wide applications across various domains, including science, engineering, economics, and everyday decision-making.

There are two main types of probability distributions:

(i) Discrete Probability Distributions: These describe the probability of a discrete random variable (e.g., the number of heads in a coin toss).

(ii) Continuous Probability Distributions: These describe the probability of outcomes of a continuous random variable (e.g., the height of individuals)

Common Types of Probability Distributions

(1) Binomial Distribution

The binomial distribution is used for binary events - situations where there are only two outcomes, often referred to as success and failure.

Application:

- Quality Control: To determine the probability of a defective item in a batch
- Pharmaceutical Trials: To estimate the probability of a patient responding to a new treatment.

(2) Poisson Distribution

The Poisson distribution gives the probability of a given number of events happening in a fixed interval of time or space when the events are independent.

Applications:

- Call centres
- Traffic Engineering
- Natural Events

(3) Normal Distribution

The normal distribution, also called the Gaussian distribution, is symmetric and describes many natural phenomena. It is widely used because of the central limit theorem, which states that the sum of a large number of independent random variables tends to be normally distributed.

Applications

- Finance
- Standardized Testing
- Physics

(4) Exponential Distribution

The distribution is used to model the time between events in a Poisson process, where events occur continuosly and independently.

Applications

- Reliability Engineering
- Queuing Theory
- Healthcare

(5) Uniform Distribution

In a uniform distribution, all outcomes are equally likely. This is used when there is no preference for one outcome over another

Applications

- Gaming
- Random Sampling
- Simulations

(6) Gamma Distribution

It is a two-parameter family of continuous probability distributions. It generalizes the exponential distribution and is used to model the time to complete k independent events.

Applications

- Insurance
- Hydrology
- Queue Mangement