# Probability & Statistics

Model Answer MID SEMESTER

1st APRIL 2025

Q-1) (1) (A) what type of statistical data are used in educational research

Both of them

(B) which of the following is measure of Central Tendency

Median

(C) which measure of central tendency used for categorical data

Mode

(D) which descriptive statistics gives the highest weight to outliers. Mean

(E) Consider the poisson distribution for the tossing of a baised coin. The mean for this distn is $\mu$. The std. deviation for this distribution is given by ____

$\sqrt{\mu}$

(2) True False

(A) let $X \rightarrow N(\mu, \sigma^2)$ If $\mu^2 = \sigma^2$ then the value of $P(X < -\mu / X < \mu)$ in terms of cummulative function $N(0,1)$ is $[1 - P(Z \leq 2)]$

FALSE

Explanation: $X \sim N(\mu, \sigma^2)$

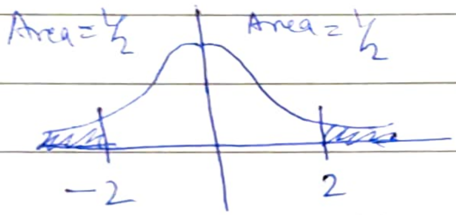
$\mu^2 = \sigma^2 \quad (\mu > 0)$

$P(X < -\mu / X < \mu) = ?$

transforming to $Z \sim N(0,1)$

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

$\sigma^2 = 1 \Rightarrow \sigma = 1$, $\mu > 0$, $\mu^2 = \sigma^2 \Rightarrow \mu = 1$




when $X = -1$

$$Z = \frac{-1-1}{1} = -2$$

when $x = 1$

$$z = \frac{1-1}{1} = 0$$

$\therefore P(X < -\mu / X < \mu)$

$= P(Z < -2 / Z < 0)$

$= \frac{P(Z < -2 \cap Z < 0)}{P(Z < 0)} = \frac{P(Z < -2)}{P(Z < 0)}$

$= \frac{P(Z > 2)}{P(Z < 0)} = \frac{1 - P(Z < 2)}{1/2}$

$= 2[1 - P(Z < 2)]$

(B) For a Binomial variable $x$, If $X_1, X_2, X_3, \dots, X_n$ are independent & identically distributed samples from the distribution of $x$ with sum $Y = \sum^{n} x_i$ then the distribution of $Y$ as $n \to \infty$ can be approximated as Bernoulli distribution.

FALSE

(C) There exists a Binomial distribution with mean 5 and standard deviation 3.

FALSE

(D) For X be a poisson variate, then for Poisson distⁿ with X mean and std. deviations are different.

TRUE

(E) The central Limit Theorem gurantees that the population is normal whenever n is sufficiently large.

FALSE

[3M]

Q.2) (A) Let X be the amount of time a postal clerk spend on his/her customer. The time is known to have an Exponential distribution with the average amount of time equal to four minutes. Find the prob. that a clerk spend four to five minutes with a randomly selected customer.

Ans : Let $X \sim$ Exponential random variable

Given $\mu = 4$ min

$\Rightarrow$ decay parameter $m = \frac{1}{\mu} = \frac{1}{4} = 0.25$

$\sigma = S.D = \mu = 0.25 \Rightarrow \sigma = 0.25$

To find $P(4 < x < 5) = P(x<5) - P(x<4)$

Now $P(x < x) = 1 - e^{-mx}$

$\Rightarrow P(x<5) = 1 - e^{-(0.25)5} = 0.713$

& $P(x<4) = 1 - e^{-(0.25)4} = 0.6321$

$\Rightarrow P(4<x<5) = 0.7135 - 0.6321$

$= 0.0814$

Q.2) (b) In order to identify which 3 cities receive the most sunshine. We simply compare the frequencies bet$^n$ each. First reorder the cities in order from the most to least hrs of sunshine then calculate their relative freq.

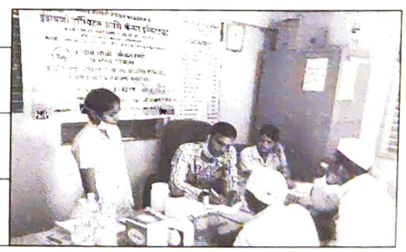
| Country | City | freq. | Relative. |
|---|---|---|---|
| Malta | Valletta | 3054 | 3054/24446 = 12.5 % |
| Greece | Athens | 2773 | 2773/24446 = 11.3 % |
| Italy | Cagliari | 2726 | 11.2 % |
| Spain | | 2696 | 11 % |
| Montenegro | | 2481 | 10.1 % |
| Portugal | | 2468 | 10.1 % |
| Bulgaria | | 2177 | 8.9 % |
| Serbia | | 2112 | 8.6 % |
| Georgia | | 2046 | 8.4 % |
| Croatia | | 1913 | 7.8 % |
| | Total | 24446 | 100 % |

2M

To calculate cummulative freq. we add them & top three cities we get

½M 12.5 % + 11.3 % + 11.2 %
= 35 %

½M ∴ Top 3 cities in list receive more than a quarter of the total sunshine hrs.



Q.2) (C) First in order to calculate row & column frequencies.
will calculate totals of rows & columns.

| Preposition | Purple party | Green Party | Total |
|---|---|---|---|
| A | 13 | 54 | 67 |
| B | 60 | 20 | 80 |
| C | 27 | 26 | 53 |
| Total | 100 | 100 | 200 |

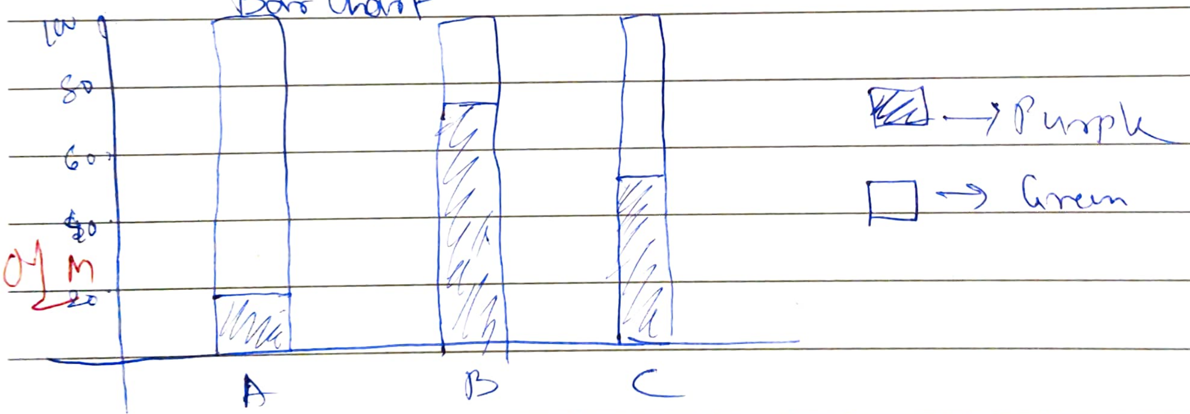
Row frequency & chart

01M

| Preposition | Purple Party | Green Party | Total |
|---|---|---|---|
| A | 19% | 81% | 100% |
| B | 75% | 25% | 100% |
| C | 51% | 49% | 100% |
| | 50% | 5% | 100% |

Bar chart

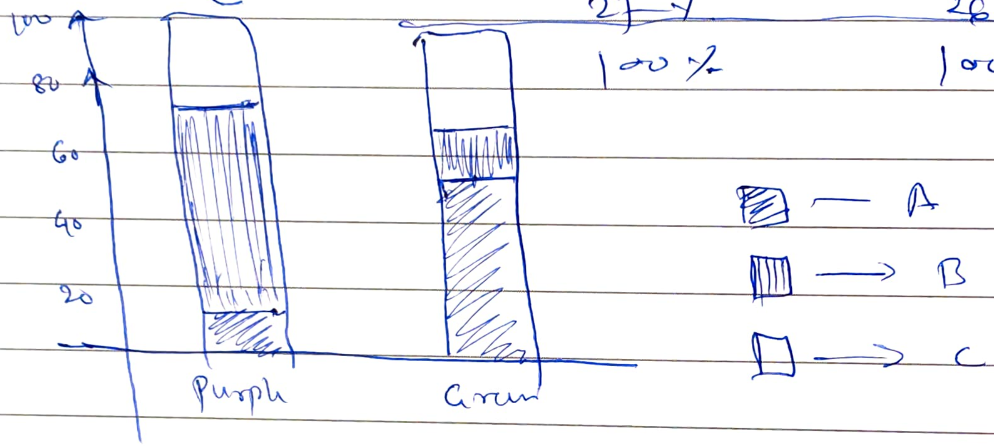
01M



The column freq. & chart is

| Preposition | Purple Party | Green Party | Total |
|---|---|---|---|
| A | 13 % | 54 % | 34 % |
| B | 60 % | 20 % | 40 % |
| C | 27 % | 26 % | 27 % |
| | 100 % | 100 % | 100 % |

01




01

Q.3) (A) $X \sim B(n, p)$, $n = 300$, $p = 0.53$

$np \geq 5$ & $nq \geq 5$ $\therefore$ approximately to normal

$\mu = np = 300 \times 0.53 = 159$

$\sigma = \sqrt{npq} = \sqrt{300 \times 0.53 \times 0.47}$

$= 8.6447$

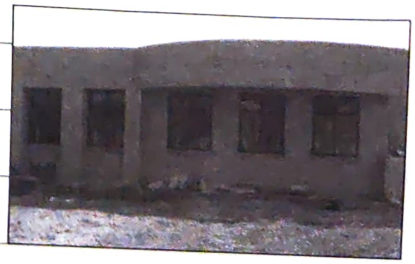
$Y \sim N(159, 8.6447)$

(a) $P(X > 150) \approx P(X \geq 149.5)$

$= 0.8641$

(b) $P(X \leq 160) \approx P(X \leq 160.5)$

$= 0.5689$



OR.

① $P(X > 155) \approx P(X > 155.5)$
$= 0.6572$

② $P(X \le 147) \approx P(X < 146.5)$
$= 0.0741$

Q.3) (B) $X \sim (70, 0.95)$ & $Y \sim (70, 0.05)$

For X, $n > 50$ But $np = 70 \times 0.95 = 66.5$
$np > 5$

Thus X can not be approximated.
However Y, $n > 50$
$np = 70 \times 0.05 = 3.5 < 5$

1M 1) So Y can be approximated by a poisson distribution $Y \sim P_0(3.5)$

1M ② $P(X = 67)$ $P(\cancel{Y=3})$ $= \dfrac{e^{-3.5}(3.5)^3}{3!}$

$= 0.2158$

1M ③ $P(X \ge 67) = P(Y \le 3) = 0.5366$

Q.3) © 1) Let X be poisson variate.

R-command

plot (dpuis (x=1 : 50, lambda = 3))




2) Y <- dbinom(0:25, size 20, prob 0.5)
plot (0:25, Y, type "h")