7. Customer Purchase Analysis:

Dataset

Customer_id	num_purchases	purchase_amount
1	2	35.50
2	1	45.00
3	3	60.00
4	0	0.00
5	4	55.50
6	2	40.00
7	1	30.00
8	3	70.00
9	2	50.00
10	1	25.00
11	0	0.00
12	3	65.00
13	2	45.00
14	1	35.00
15	4	80.00
16	2	50.00
17	1	40.00
18	3	75.00
19	2	55.00
20	1	30.00

Page
Date / /

no. of purchases

∵ values are discrete
⇒ Discrete Distribution

- (μ) Mean $= \frac{38}{20} = 1.90$
- Variance $= \frac{1}{N} \sum_{i=1}^{N} (\text{num\_purchases} - \mu)^2$

$$\sigma^2 = \frac{25.80}{20} = 1.29$$

$$\mu = 1.90, \ \sigma^2 = 1.29 \ (\sigma = 1.136)$$

Poisson Distribution makes the most sense

∵ variance is lower than mean, which means that observations around the mean ~~more~~ are more tightly packed.

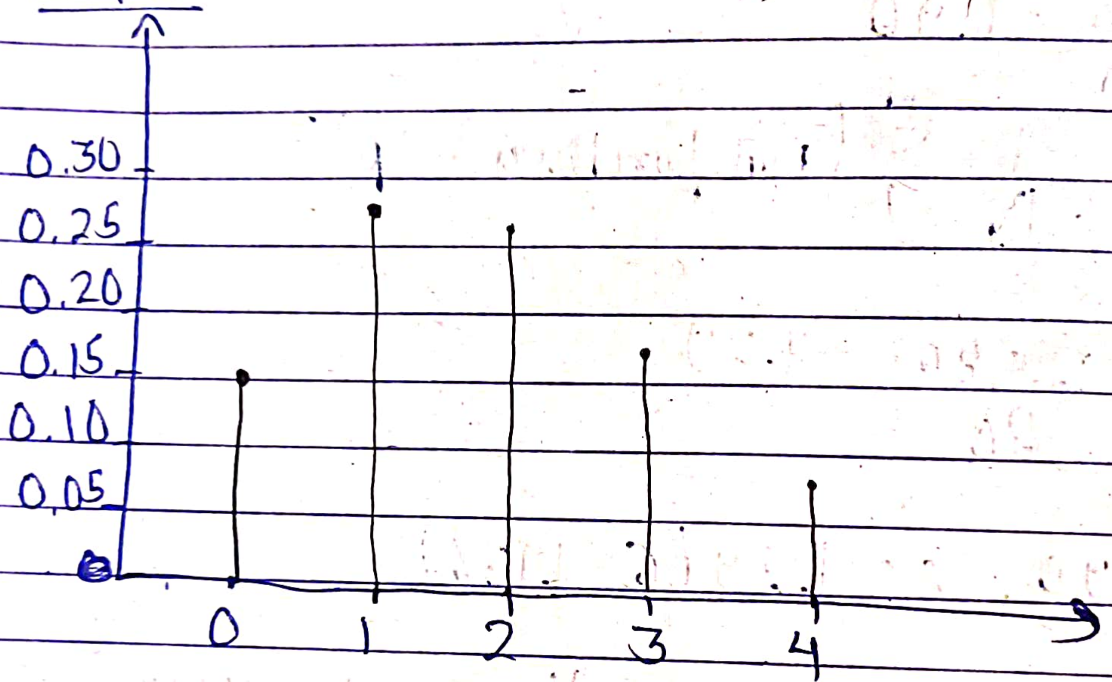
$$\lambda = 1.9 = E(X) = Var(X)$$

Page

Date / /

$$P(X=k) = \frac{\lambda^{k} e^{-\lambda}}{k!}, \quad k = 0, 1, 2, 3, 4$$

Graph (Poisson distribution)



Let us also assume a binomial distribution

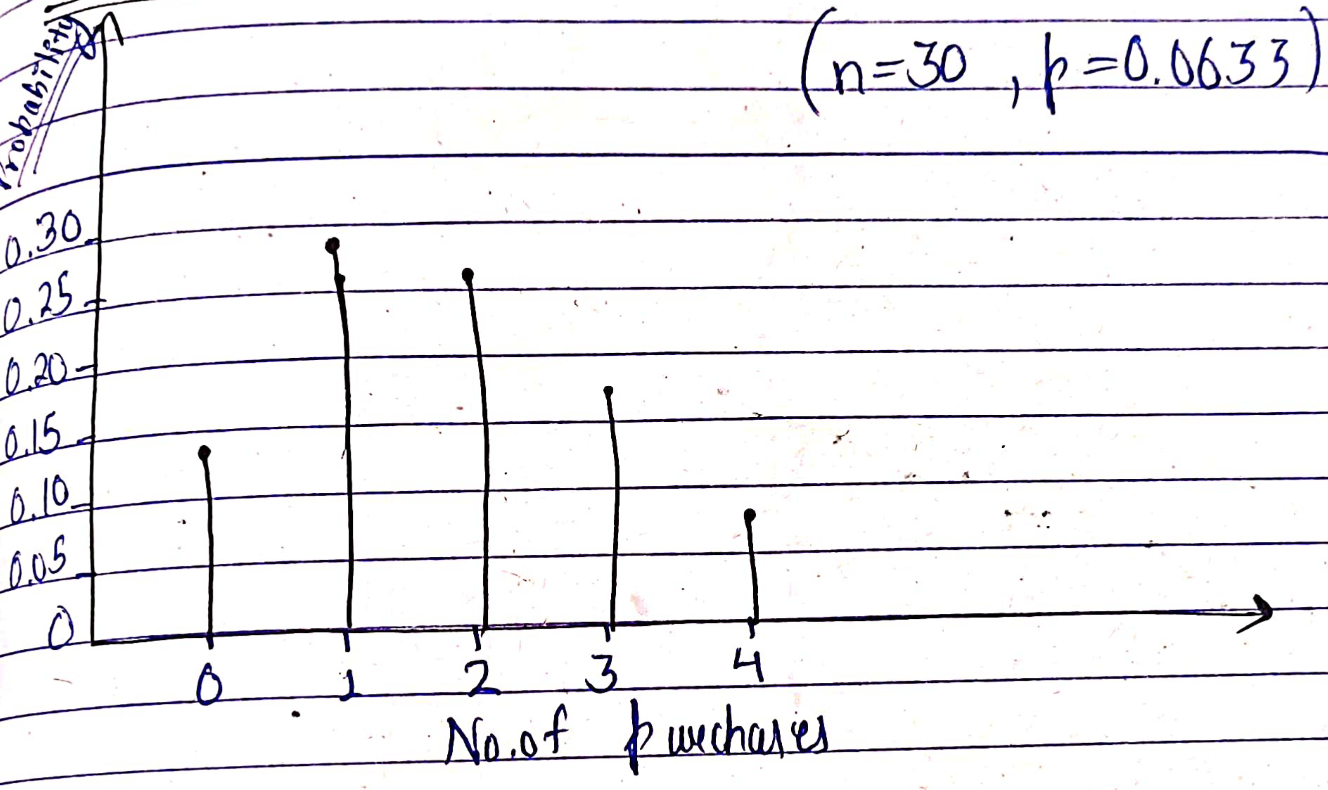
$n = 30$ (assume)

$p = 0.0633$ (chance of a purchase on each day)

mean: $np \approx 1.899$

Variance: $np(1-p) \approx 1.78$

Graph (Binomial)



Purchase Amount:

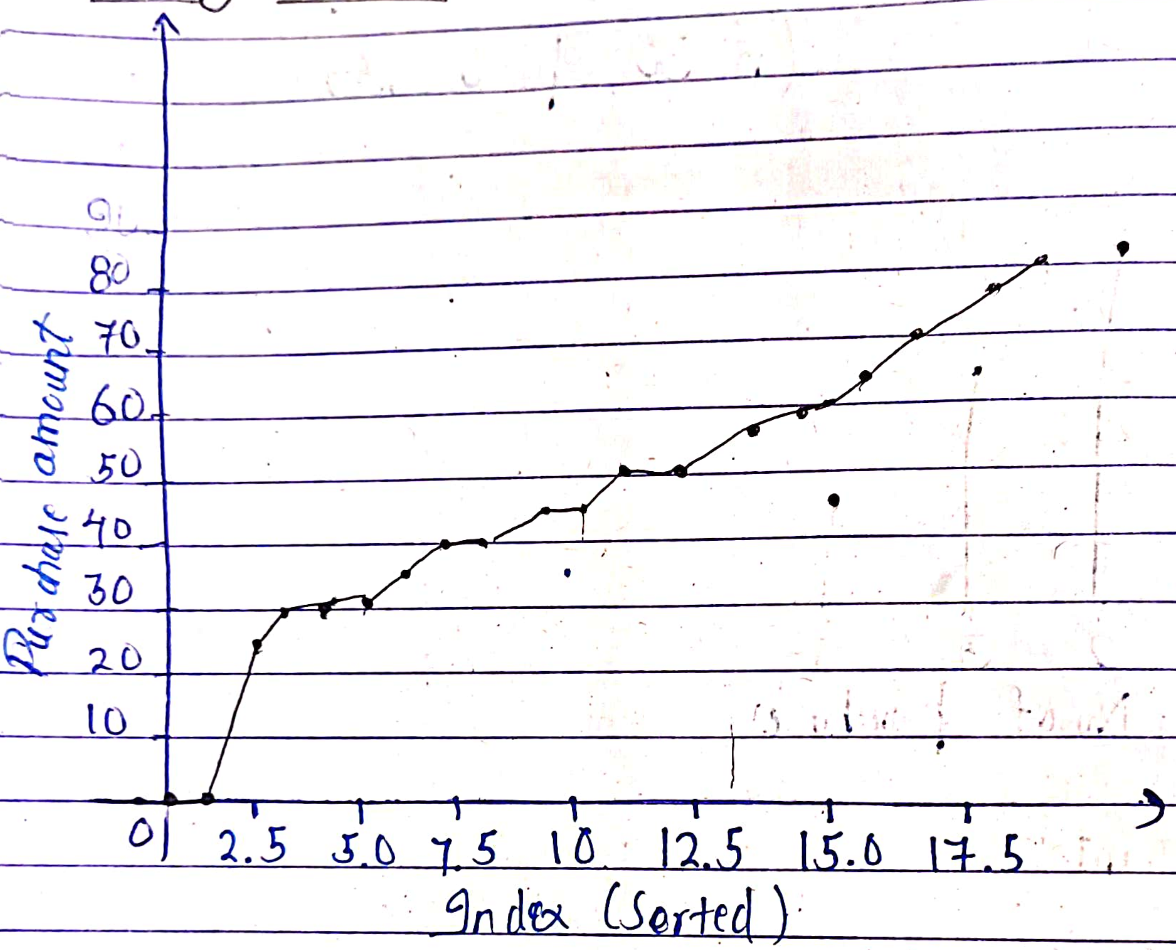
mean = 44.05

Variance = $\sigma^2 = 443.94$

$\sigma = 21.07$

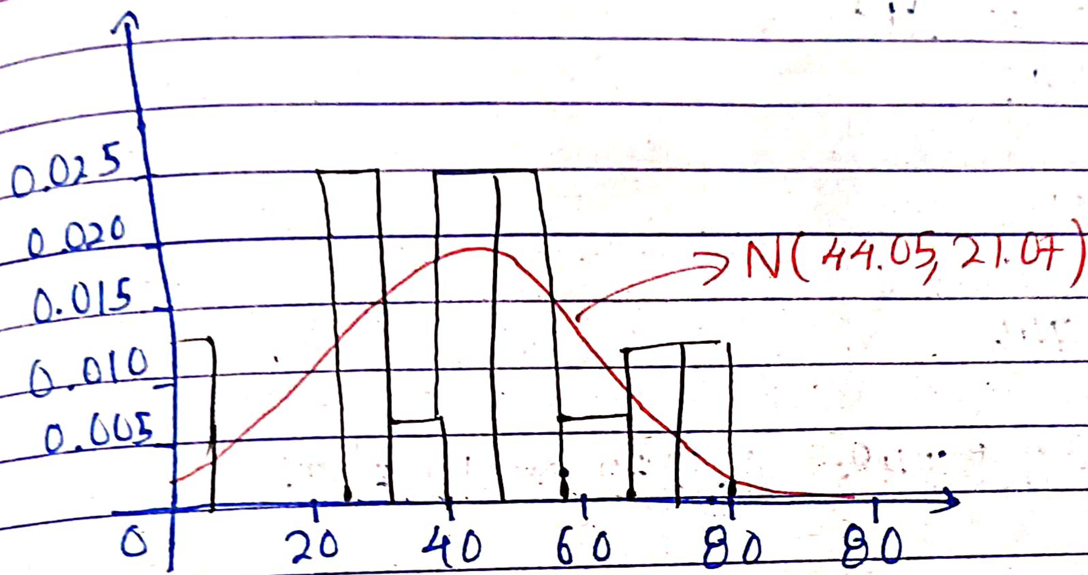
Page
Date / /

Plotting the data :



* The data does not make sense so, we plot a normal distribution.

$$N(\mu = 44.05,\ \sigma^2 = 443.94)$$



- the distribution does not accurately represent the data (but we will proceed with this model only.

- P (customer makes more than 3 purchases in a month)

$$X \sim \text{Poisson}(\lambda = 1.9)$$

$$P(X > 3) = 1 - P(X \leq 3)$$

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, 3, 4$$

$$P(X \leq 3) = \sum_{k=0}^{3} \frac{e^{-1.9} (1.90)^k}{k!}$$

Page
Date / /

- $P(X=0) = e^{-1.9} = 0.1496$
- $P(X=1) = 0.2842$
- $P(X=2) = 0.2699$
- $P(X=3) = 0.1711$

$$P(X \leq 3) = 0.8748$$

$$P(X > 3) = 1 - 0.8748 = 0.1252 \approx 12.52\%$$

The prob. a customer makes more than 3 purchases in a month is about 12.5%.

- P (customer spends more than \$50 in a single purchase)

$$P(Y > 50) = 1 - P(Y \leq 50)$$

$$Z = \frac{Y - \mu}{\sigma} = \frac{50 - 44.05}{21.07} \approx 0.28$$

$$P(Y > 50) = 1 - \text{probability from Z score } (0.28)$$

$$= 1 - 0.61$$

$$= 39$$

$$= 39\%$$

Page
Date / /

The prob. a customer spends more than \$50 (under this Normal assumption) is about 39%.

—x—x—