Lecture-3

Levels of measurement or scale

- **Ratio:**
  - i) Numaric
  - ii) Can be ranked
  - iii) ~~Only difference is meaningful~~. All mathematical calculation can be done
  - → CGPA
  - → Remitence
  - → Pricing
  - → Time duration

- **Interval:**
  - i) Numaric
  - ii) Can be ranked
  - iii) ~~All mathematical calculation can be done~~
  - iii) Only difference is meaningful. (while comparing)
  - → Time point
  - → Temprature

- **Ordinal (order):**
  - (i) Data: Non-numaric, non-Identification
  - (ii) Can be ranked
  - (iii) Non mathematical
  - (iv) Operation can be done

  Ex: Grade, Army rank

- **~~Ordinal~~ Nominal:** Identification
  - i) Data Non-numaric
  - ii) Can't be ranked
  - iii) Non-mathematical
  - iv) Operation can be done
  - → blood group

→ ~~Nominal~~

Continus: measurable & fraction number
Discrite: we got it by counting & no fraction

Nominal

In excel:

Insert → pivot → for frequency.

Frequency table:

Frequency table of the blood group of the STA102 students

Class (group)	Frequency (count)	Tally	Relative (%) frequency
A+	10	𝍸𝍸	31.250%
B+	7	𝍸 ll	21.875%
AB+	5	𝍸	15.625%
O+	9	𝍸 llll	28.125%
O-	1	l	3.125%

Comments: In the table we have considered 32 students. Most of the students are having "A+" Blood group. But only one student have "O-" Blood group.

* Frequency table of marks obtained by the STA102 students:

k = Number of class $= 1 + 3.222 \log(N) = 1 + 3.222 \log(19)$

better → continues

Exclusive method	Inclusive method
50 - 60	51 - 60
60 - 70	61 - 70
70 - 80	71 - 80
80 - 90	81 - 90
90 - 100	91 - 100

15-02-2023

Lecture - 9

⇒ Row is an element in the table. Number of row × num of column = elements

⇒ Column is variable " " " .

⇒ "N" → population of size.

⇒ "n" → Sample size.

L-56   93-56
H-93   = 37
n = 18
k = 6 is the gap in class.

* Frequency table of Marks obtained by the students:

Class	Frequency	Tally	Relative Frequency	Cumulative frequency less than	Cumulative frequency More than
56-63	4	IIII	20%	4	20
63-70	3	III	15%	7	16
70-77	5	~~IIII~~	25%	12	13
77-84	4	IIII	25%	16	8
84-91	2	II	10%	18	4
91-98	2	II	10%	20	2

Comment: From the table we can see that most of The student got the number range between "70-77" and the only few got number between the range "84-91" and "91-98" marks. And in the table we considered data of 20 students.

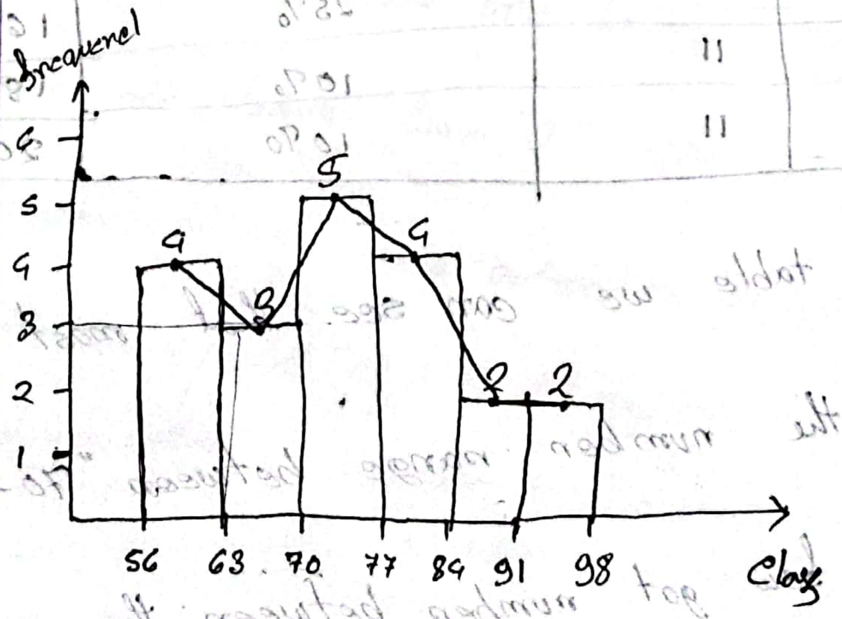
→ New data

Cumulative frequency		Cumulative Rel. frea	
less then type	More then type	less then type	Morethen type

Graphical Presentation

→ Bar diagram : catagory , inclusive

→ Histogram : exclosive



Histogram of the mark of the students.

20-02-2022

Lecture - 5

* Frequency polygon: (line chart)

Class | Frequency | Class Midpoint

~~class~~

* Summery statistics

\# Measures of Central Tendency

tools/ Calculation

↳ Average value
Central value
Representation value

↳ Arithmetic Mean
↳ Median (Position) Doesn't considers all value.
↳ Mode (Most frequent data)
↳ Geometric mean (Grote rate/change rate/ non linear change)
↳ Harmonic mean (Speed / something per something)

Mesures	N	O	I	R	Best use case.
A. Mean	×	×	✓	✓	
Median	×	✓	✓	✓	
Mode	✓	✓	✓	✓	
Geo. Mean	×	×	×	×	
Har. Mean	×	×	×	×	

Propenties:

1) It should consider all the values of a data set.
2) " " not be affected by $\boxed{\text{extreme values/outliers}}$
↳ unusally lowest and height value

Lecture - 6

⇒ Arithmetic mean:

↳ considers all value (Get's biased by extreme value)
↳ if extreme value seen don't use it.

Formula:

Population:

$$\mu = \frac{\Sigma x}{N}$$

Sample:

$$\bar{x} = \frac{\Sigma x}{n}$$

Comment: The estimated average age of the student of the class.

⇒ Median: is a position,

Median is $\left(\frac{n+1}{2}\right)^{th}$ data

$$\frac{5+1}{2} = 3rd \text{ value}$$

⇒ Have to sort the data and then have to identify the 3rd value. (or Median value)

Comment: Hence the avarage...is...

↳ There are 50% responds with income less than BDT 27000 and the

Mode: → most repeated value
→ can be multiple value
→ can't even exist (if no repeted value)

Geometric Mean:

Use: i) Non linear ~~dot~~ groth
ii) to calculate the percentage change over the time period
iii) If the data is in percentage, to calculate the average

Q: After diposi ting 100 taka, a person has received 121 taka after 2 year what is the average annual change per year.

Solution:

$$G.M = \left(\frac{P_n}{P_0}\right)^{1/n} - 1$$

P_n = new value
P_0 = old value
n = number of data

$$= \left(\frac{121}{100}\right)^{1/2} - 1 = 0.1 = 10\%$$

Comment:
So, On average the person has got 10% mean every year.

CamScanner

Q: The profit/large statement of a company for the last 5 years is as follows.

Year	Profit / lose
2017	5% Profit
2018	2% Profit
2019	3% lose
2020	6% lose
2021	10% Profit

⇒ What is the average return (Profit/Loan) of the company in the reported year.

Solution:

$$G.M = (\text{Multiplication of data})^{\frac{1}{\text{Number of data}}}$$

$$= (x_1 \times x_2 \times x_3 \times \ldots \times x_n)^{1/n}$$

Accumulated value	Year
105	2017
102	2018
97	2019
94	2020
110	2021

So, $G.M = (105 \times 102 \times 97 \times 97 \times 110)^{1/5} = 101.44$

Comment: On average the company has made 1.44% profit in those reported years.

27/02/2023

Lecture - 7

Quantiles

Harmonic Mean

Q: What is the average speed?

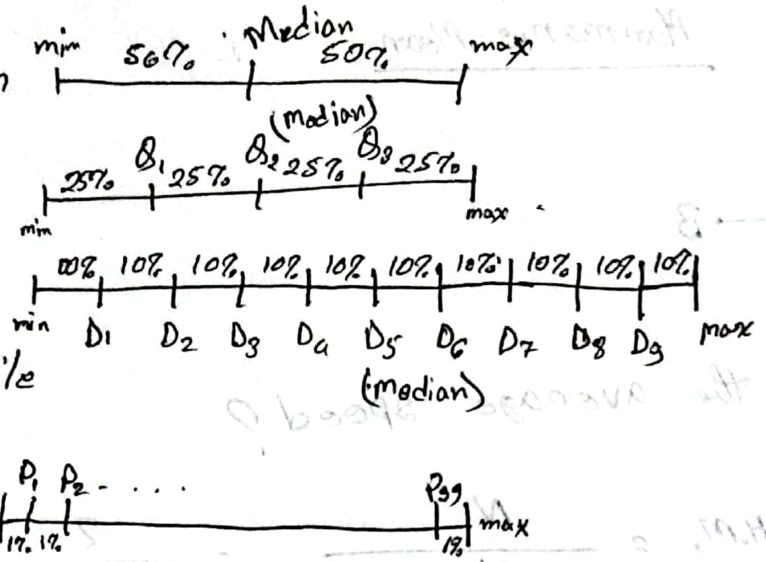
Solution: $$H.M. = \frac{N}{\frac{1}{x_1}+\frac{1}{x_2}+\dots+\frac{1}{x_N}} = \frac{2}{\frac{1}{10}+\frac{1}{15}} = \frac{2}{\frac{1}{6}} = 12$$

Best use: speed related data. (Something per Something)

Comment: On the given data set the average speed is 12 km/h.

Quantiles [Positional Measures]

- Median (half)
- Quartile (Quarter)
- Decile
- Percentile



Q1: Marks

1) Find Q_3

Sol: $Q_3 = \left(\frac{N+1}{100} \times 75\right)$th value $= \left(\frac{21 \times 75}{100}\right) = 15.75$th value

2) Find D_2

$D_2 = \left(\frac{N+1}{100} \times 20\right) = \frac{21 \times 20}{100} = 4.2$th value $= 62 + (0.2 \times 3) = 62.6$

3) Seperate the worst 12% student

commem

→ 25% student got ~~the~~ more than 82.25 and 75% students got less than that

→ 20% students got less ~~more~~ than 62.6, and 80% students got more than that

→ $\left(\frac{21 \times 12}{100}\right) = 2.52$th value $= 59 + (0.52 \times 1) = 59.52$

Comment: The worst 12% students are with marks 56 & 59

~~Assignment-1~~

Lecture-8

tools/ calculation ← Measures of Dispersion (Calculates the realability of Average)

Absolute (single dataset)

→ Range
→ mean Deviation
→ Standard "
→ Quentile "
→ Varience

Relative type (for comparison purpose) two or more data set

Calculated in "%"

→ Co-efficient of range
→ " of mean deviation
→ " of quartile deviation
best → " of variation

CGPA of the stds

1.92, 2.14, 2.26, 2.33, 2.24, 2.11

range = H-L
= 2.33-1.92 = 0.41

→ Gap of height and lowest.

Comment: The gap between the height & lowest CGPA of the data set is 0.41.

অসুবিধা: Can be menupulated by extreme value. does not considers all value

$f(x) = |x|$

Mean Deviation

$$\bar{x} = \frac{1.92 + 2.14 + 2.26 + 2.33 + 2.24 + 2.11}{6}$$

$= 2.167.$

Formula

$$\frac{\sum |x - \bar{x}|}{n} \quad \text{or} \quad \frac{\sum |x - \mu|}{N}$$

$= 0.11$

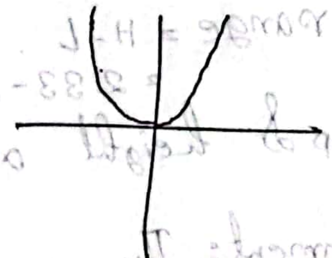
Counts all value

messup by Extrime value

Comment: The average absolute distance from mean is 0.11.

Standard Deviation:

$f(x) = x^2$



Formula,

Population..., $\sigma = \sqrt{\frac{\sum (x-\mu)^2}{N}}$ sample..., $(s) = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}}$

$= 0.148$

comment: The average distance from mean is 0.148.

Variance

→ On an average, the squared distance from mean is 0.021.

13-03-2023

Lecture - 9

Quartile Deviation:

Inter Quartile range (IQR): $Q_3 - Q_1$

Formula = $\frac{IQR}{2} = \frac{Q_3 - Q_1}{2}$

Data: 9 11.5 13.5 14.5 14.5 15.5 17.5 17.5

$Q_1 = \frac{9 \times 25}{100} = 2.25 = ~~11.5~~ \; 12$

$Q_3 = \frac{9 \times 75}{100} = 6.75 = ~~16.5~~ \; 17$

$IQR = 17 - 12 = 5$

Quartile Deviation $= \frac{5}{2} = 2.5$

Comment: The average inter quantile distance is 2.5.

Benefit: Can't be manipulated by Ext value.

Variance:

$$\sigma^2 = \frac{\Sigma (x-\mu)^2}{N} \qquad s^2 = \frac{\Sigma (x-\bar{x})^2}{n-1}$$

Comment: On an average, the squared distance from mean is 0.021.

Co-efficient of variation: (CV)

Formula:

$$\frac{s}{\bar{x}} \times 100\% \quad \text{or} \quad \frac{\sigma}{\mu} \times 100\%$$

Data:

Male: 15.5 18 16.5 17.5 14

Female: 12.5 15 19 17

Solution

Male data:

$$\bar{x}_m = \frac{15.5+18+16.5+17.5+14}{5} = 16.3$$

$$s_m = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{(-0.8)^2+(1.7)^2+(0.2)+(1.2)^2+(-2.3)^2}{4}}$$

$$= \sqrt{\frac{0.64+2.89+0.04+1.44+5.29}{4}} = \sqrt{\frac{10.3}{4}} = \sqrt{2.575}$$

$$= 1.60$$

$$CV_m = 9.89\% = \frac{1.60}{16.3} \times 100$$

$\bar{x}_f = 15.875$

$sd_f = 2.278$

$CV_f = 17.51\%$

Comment: The varition in female data from male classmates.

or → Males data is more consistant than their female classmates.

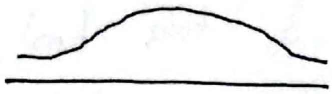
or → On an average, female are more dispersed/scattered than their amale classmates if it is about mid-1 marks.

⇒ Co-efficient of Quintile deviation:

Formula: $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100\%$

Measures of skew ness (Unitless)
or
[Shape characteristics of data]

↳ Right skewed / Positively skewed 

↳ Symmetric 

↳ Left skewed / negatively skewed 

⇒ Karl Pearsons coefficient of skewed $= \dfrac{3 \times (\text{mean} - \text{Median})}{\text{Stdand deviation}}$

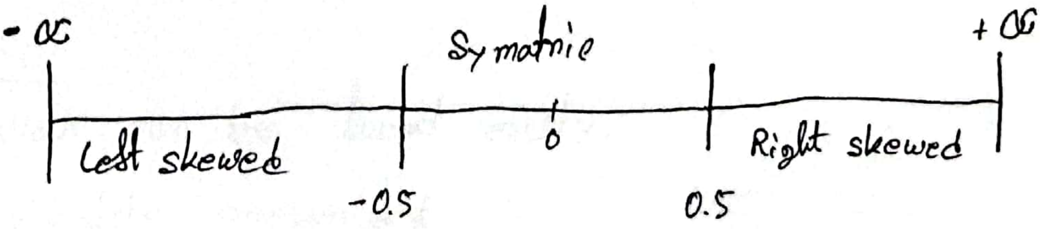
mean = 16.1

median = 16.5

SD = 206

$$sk_p = \frac{3(16.1 - 16.5)}{206}$$

$$= -0.583$$



Comment:

the shape of the data is left skewed.

↳ symmetric meaning : Mean = Median = Mode (or almost)

↳ Right skewed " : Mean > Median > Mode

↳ Left skewed " : Mean < Median < Mode

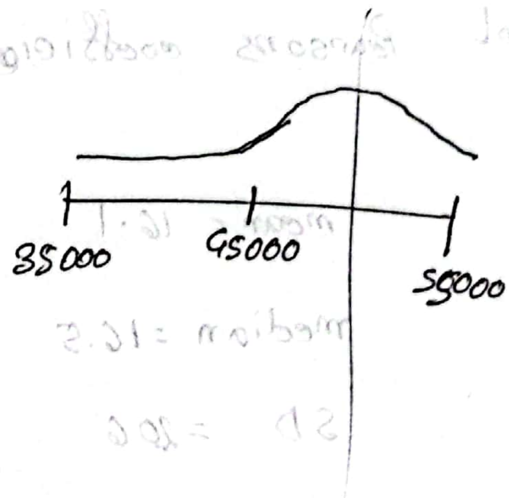
Scenario:

Finding : <=50000

Avg : 45000₹

Std : 10000

Skewed : Left



Stem and Leaf Plot

Marks:

99, 90, 68, 62, 73, 75, 78, 83, 89

Figure: Stem & leaf plot of marks

Stem (10th place)	Leaf ~~Place~~ (unit place)
6	2 8 (2)
7	3 5 8 (5)
8	3 9 (7)
9	0 3 (9)

→ Cumulative frequency in braket / Frequency

Should be done

Conditions:

* Ascending order vertically + Horizontaly
* No comma

Benefits:

→ median can be found easily.

→ Individuality presurived.

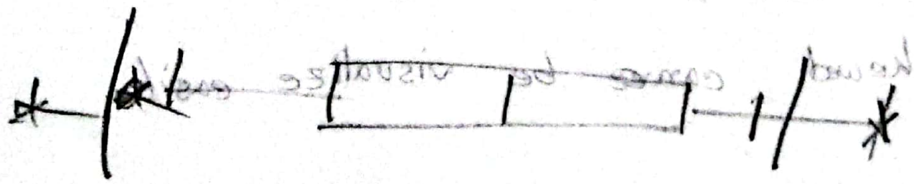
→ Skewed canee be visualize easily.

CGPA:

2.26, 2.33, 3.4, 2.06, 1.67, 4.0

Round up (1 decimal place): 2.3, 2.3, 3.4, 2.1, 1.7, 4.0

Stem (Unit)	Leaf (Decimal)
1	7 (1)
2	1 3 3 (3)
3	4 (1)
4	0 (1)

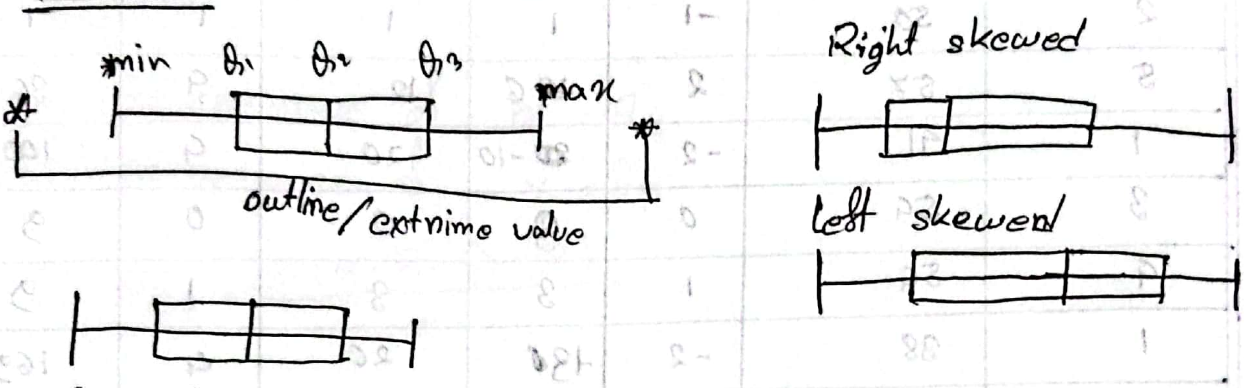


Lecture - 12

Box plot or Box & whiskers plot

Q: Construct a Box & whiskers plot.

Concepts:



Answer:

Process of the calculation is first have to find the Min, Q_1, Q_2, Q_3, Max value. Then have to find the limits.

Min = 62
Q_1 = 70.5
Q_2 = 78
Q_3 = 89.5
Max = 93.

Upper limit = $Q_3 + 1.5\ IQR$ = 118

Lower limit = $Q_1 - 1.5\ IQR$ = 42

Inter Quartile Range = $Q_3 - Q_1$

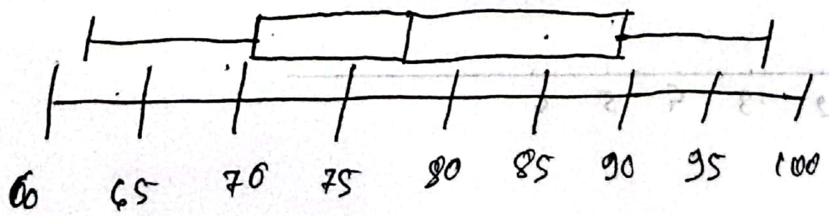


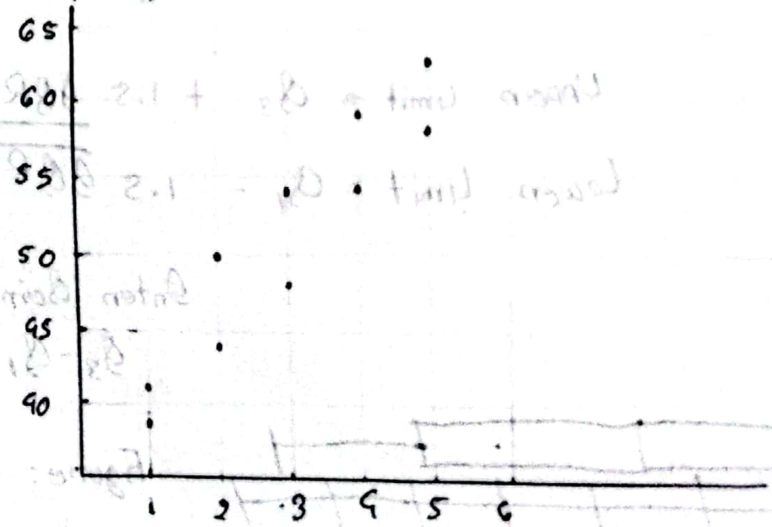
Figure: Box plot of marks obtained by the students.

Correlation

↳ To find linear relationship among related variables.

Number of Commercial (X)	Sales Volume ($100) (Y)	$x-\bar{x}$	$Y-\bar{Y}$	$(x-\bar{x})(Y-\bar{Y})$	$(x-\bar{x})^2$	$(Y-\bar{Y})^2$
2	50	-1	1	1	1	1
5	57	2	6	12	9	36
1	41	-2	-10	20	4	100
3	54	0	3	0	0	9
4	54	1	3	3	1	9
1	38	-2	-13	26	4	169
5	63	2	12	24	4	144
3	48	0	-3	0	0	9
4	59	1	8	8	1	64
2	44	-1	-5	5	1	25
Avg: 3	Avg: 51			Σ=99	Σ=20	Σ=566

Scatter plot:



Comment: The relationship between Number of commercials and sales volume is positive relationship.

Karl pearsons correlation co-efficiant, $r = \frac{S_{xy}}{S_x S_y}$

Here,

co-verience of $x = S_x$

co-verience of $y = S_y$

$$S_{xy} = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{n-1}$$

$$S_{xy} = \frac{99}{9} = 1$$

$$S_x = \sqrt{\frac{20}{9}} = 1.491$$

$$S_y = \sqrt{\frac{560}{9}} = 7.93$$

So,

$$r = \frac{11}{1.491 \times 7.93}$$

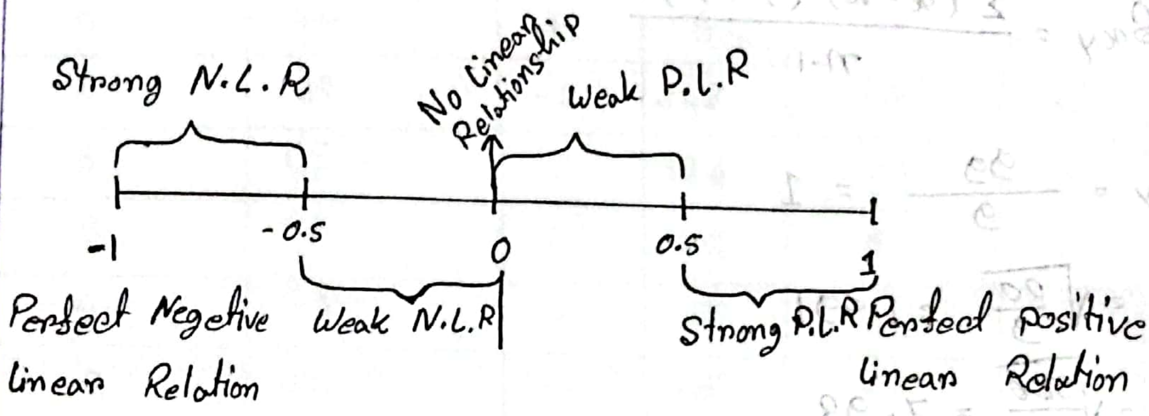
$$= \frac{11}{11.829}$$

$$= 0.93$$

Properties of r:

- $-1 \leq r \leq 1$
- unitless
- $r_{xy} = r_{yx}$

Symmetric property



Comment: There is strong linear relationship between number of commercial and salse volume.

22-03-2023

Lecture -12

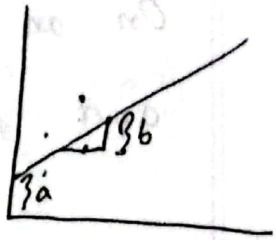
Regression

↳ Explain one variable with the ~~of~~ help of other.

The estimated regression equation is,

estimated → cap means

$\hat{Y} = a + bx$



here,

$b = \cancel{\frac{S_{xy}}{S_x S_y}} \qquad \frac{S_{xy}}{S_x^2}$

$a = \bar{Y} - b\bar{x}$

	a	b
Limit	$-\infty, +\infty$	$-\infty, +\infty$
Unit	Y's unit	Y unit / x unit

So, $b = \frac{11}{(1.491)^2} = \frac{11}{2.223} = 4.95$

and,

$a = 51 - 3 \times 4.95$

$= 51 - 14.849 = 36.16$

Comment: For every commercial the salse volume will increase \$495 dollar.

or,

For every commercial increase, the estimated/average/expected increase of salse volume is \$495.

↳ a is the value of Y is 0".

Standard error of estimate, $S_y = \sqrt{\frac{\Sigma(Y-\hat{Y})^2}{n-2}} = \sqrt{75.95}$

Limit: $[0, \infty)$ $= 3.08$

On an average the gap between the estimated and the actual sales volume is \$308.

Lecture - 13

Coefficiant of determination (R^2)

↳ Formula

$$R^2 = \frac{SSR}{SST} \quad \text{or} \quad = 1 - \frac{SSE}{SST} \quad \text{or} \quad = (r^2)$$

Here,

(SSR) Sum of squere Regration $= \Sigma(\hat{Y} - \bar{Y})^2$

(SSE) " " " Error $= \Sigma(Y - \hat{Y})^2$

(SST) " " " Total $= \Sigma(Y - \bar{Y})^2$

$SSR + SSE = SST$

$(SSE) = 75.95$

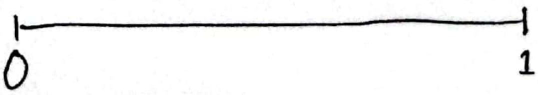
$SST = 566$

$SSR = 490.05$

$$R^2 = \frac{490.05}{\cancel{8}566} \quad \text{or} \quad 1 - \frac{75.95}{566} \quad \text{or} \quad (0.93)^2$$

$$= 0.866 \qquad = 0.866 \qquad = 0.866$$

Limit $[0, 1]$



Comment:

The ~~variable~~ variation in sale volume is 86.6% explained by the variation in number of commercial.

if,

$R^2 = 1$

So, $r = \sqrt{1} = \pm 1$ → sign of b

$S_{y \cdot x} = 0$

Seeing-theory.
brown.edu

Probability

→ uncertainity (No way to know the answer previously)

↳ Random Experiment

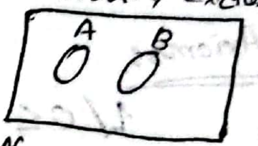
↳ Outcome → Result

↳ Event → collection/set of outcome(s)

↳ Mutually exclusive outcome/event.
Disjoint

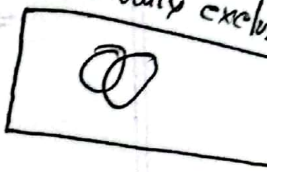
↳ Equally likely outcome/event.
Equal probability

Mitually exclusive



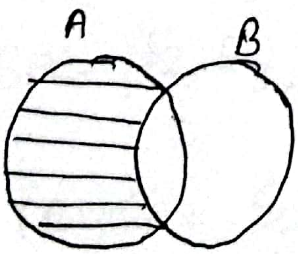
No common

Not Mitually exclu



Univarsal set → Sample set (S)

A-B



Probability:

↳ Subjective / Judgemental prob (No scientific basis)

↳ Classical → $P(A) = \frac{n(A)}{n(S)}$ → Equally likely, → Mutually likely, → Exhaustive

↳ Empirical / Practical → $P(A) = \frac{\text{Number of times A appeared}}{\text{Total number of repetation}}$

Suggest → should be large.

↳ Axiomatic

Axioms:

1/ $0 \leq P(A) \leq 1$

2/ $P(S) = 1$

3/ $P(A \cup B) = P(A) + P(B)$, if A and B are disjoint.

if A & B aren't disjoint

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

L-14

$n = 496$

$a = 11/496 = 0.022.$

$b = \cancel{10}\ 408/496 = 0.82$

$c = 52/496 = 0.11$

d

$A = \{E_1 ; E_2\}$ $\rightarrow$ $U = \{E_1, E_2, E_3, E_4, E_5\}$

$U - B$

$B^0 = \{E_3, E_4\}$ $B^c = \{E_1, E_2, E_5\}$

$A \cup B^c = \{E_1, E_2, E_5\}$

$P(A \cup B^c) = \frac{3}{5} = 0.6$

e

$C = \{E_2, E_3; E_5\}$

$B \cup C = \{E_2, E_3, E_4, E_5\}$

$P(B \cup C) = 4/5 = 0.8$

b

$A = \{E_1, E_4, E_6\}$ $B = \{E_2, E_4, E_7\}$

$A \cup B = \{E_1, E_2, E_4, E_6, E_7\}$

$P(A \cup B) = .05 + 0.2 + .25 + .10 + .05$

$= 0.65$

c

$A \cap B = \{E_4\}$

$P(A \cap B) = 0.25$

Conditional Probability

$P(A|B)$ Given that → have to focused first

Prob. of A given that B.

d)

$P(F \mid \text{not promoted}) = \frac{204}{\cancel{240}\ 876}$ ← not promoted ~~29~~ Q11(57).

e)

$P(M \cap F) = 0$

f)

$P(\text{Prom.} \cup \text{male}) = \frac{\cancel{288}}{324} = \frac{324 + \cancel{672}\ \cancel{300}}{1200}$

03-09-2023

Lecture-15

Conditional probability

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \qquad P(B/A) = \frac{P(A\cap B)}{P(A)}$$

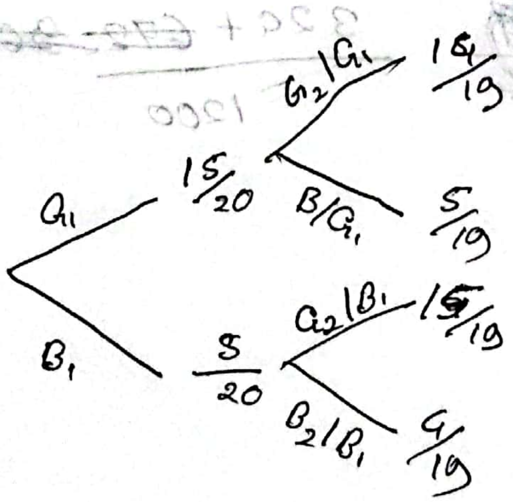
$$P(A\cap B) = P(A|B).P(B) \qquad P(A\cap B) = P(B|A).P(A)$$

Item → 20 defected → 5 sold → 2

both are good = $\cancel{\frac{2\times 15}{20}}$ ~~There is~~ $\frac{2}{20} = \frac{2}{20} \Big/ \frac{15}{20}$

$= \frac{15}{20} \times$ ___

both are bad = $\frac{5}{20} \times$



Both are good = $P(G_1\cap G_2) = P(G_2|G_1) \times P(G_1) = \frac{14}{19} \times \frac{15}{20}$

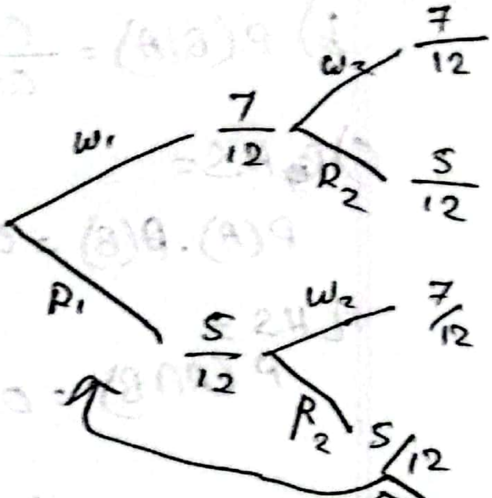
Both " bad = $P(B_1 \cap B_2) = P(B_1|B_2) \times P(B_2) = \frac{4}{19} \times \frac{5}{20}$

12

| 7W 5R | → 2 balls are selected randomly with replacement

i) both red $= \frac{5}{12} \times \frac{5}{12}$

ii) both white $= \frac{7}{12} \times \frac{7}{12}$



Independent scinario

→ If A and B are independent

i) $P(A|B) = P(A)$

ii) $P(B|A) = P(B)$

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

→ checking independence

$L.H.S = P(M \cap Prom) = \frac{288}{1200} = 0.24$

$R.H.S = \frac{960}{1200} \times \frac{324}{1200} = P(M) \times P(Prom) = 0.216$

So, $L.H.S \neq R.H.S$

So, male & promotion are not independent.

Page 217 David Ray Anderson

a) $P(A|B) = \frac{0.40}{0.60}$

b) $P(B|A) = \frac{0.40}{0.50}$

c) R.H.S =

$P(A).P(B) = 0.40 \times 0.50 = 0.20$

L.H.S =

$P(A \cap B) = 0.40$

L.H.S $\neq$ R.HS

So, A and B are not Independent.

Final Lectures: 10-04-2023

Lectures: 16

Probability distribution

Dept	frequency	P(x)
CSE	15	15/33
EEE	10	10/33
Civil	8	8/33

$\Sigma x = 33$

→ Probability distribution

↳ $f(x)$ or $P(x) \geq 0$

$\Sigma f(x)$ or $\Sigma P(x) = 1$

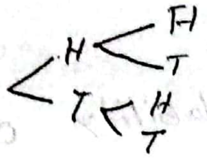
Page 235

a) yes, cause the sum of $\Sigma f(x) = 1$.

b) The prob. of x is 30 is 0.25.

c) 0.35

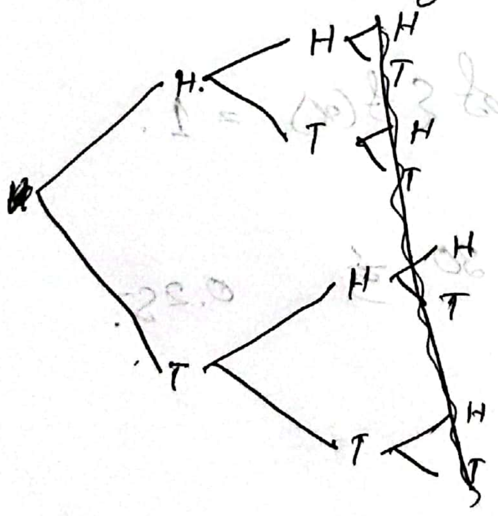
d) 0.40



$S = \{ \underset{2}{HH}, \underset{1}{HT}, \underset{1}{TH}, \underset{0}{TT} \}$

x	$P(x)$
0	1/4
1	2/4
2	1/4
	1

Q: Construct a probability distribution of number of heads for tossing 3 ~~coin~~ fair coins.



$S = \{ \underset{3}{HHH}, \underset{2}{HHT}, \underset{2}{HTH}, \underset{1}{HTT}, \underset{2}{THH}, \underset{1}{THT}, \underset{1}{TTH}, \underset{0}{TTT} \}$

x	$P(x)$

x	$P(x)$
0	1/8
1	3/8
2	3/8
3	1/8
	$= 1$

$E[x] = \text{Mean} = \mu = \Sigma x \cdot P(x) = \Sigma x \cdot f(x)$

$Var[x] = \sigma^2 = \Sigma x^2 \cdot P(x) - (\text{mean})^2$

→ Page 237

12

a) Yes, cause $\Sigma x = 1.2$

b) 0.15

c) 0.10

→ Pag 239

15

a) $E[x] = 0.75 + 3 + 2.25 = 6$

b) $\sigma^2 = (2.25 + 18 + 20.5) - 36 = 4.75$

c) $\sigma = \sqrt{4.75} = 2.18$

$x \sim Bin(\underset{n}{10}, \underset{p}{0.3})$

Binomial Probability Distribution.

→ **Problems:**

1) 2 outcome or can make the outcome intwo.

$$f(x) = \binom{n}{x} p^{x} (1-p)^{(n-x)}$$

where,

x = the number of successes

input { p = the probability of success on one ~~trial~~.

n = the number of trails

$f(x)$ = Probability of x successes in n trials.

→ simulation, to find x in given $f(x)$.

$E(x) = \text{Mean} = np$

$\text{Var}(x) = \sigma^2 = np(1-p)$

is it Binomial ?

what is p ?

Page 251

29

a) $f(3) = \binom{10}{3} 0.3^x (1-0.3)^{(10-3)}$

$= 0.267$

⇒ Formulas exp (page:245 - 257)

b) $0.028 + 0.121 + 0.233 = 0.382$

$P(x \geq 3) = 1 - 0.382 = 0.618.$

35

d) $E[x] = 20 \times 0.2 = 4$

c) $0.042 + 0.576 + 0.137 + 0.205 =$

b) 0.218

a) 0.206

12-04-2023

Lecture - 17

$X \sim Pois(\lambda)$

Poisson Distribution

Limit $(0, \alpha)$

function:

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

μ = Avg.

All rare event distribution.

All event's avg.

can calculate per location.
" " " time. } on both

255 (Page) 40

a) $\mu = 4$ per 5 min $\rightarrow \mu =$

$$f(3) = \frac{4^3 \, e^{-4}}{3!} = 0.195$$

(d)

$$f(0) = \frac{2.4^0 - e^{-2.4}}{0!} = \cancel{0.018}\ 0.091$$

μ

b) $$f(10) = \frac{12^{10} \, e^{-12}}{10!} = 0.105$$

~~d) f~~

c) $$f(0) = \frac{4^0 - e^{-4}}{0!} = 0.018$$

<u>93</u> $\mu = 10$ ~~per~~ passengers per min

<u>b</u>

$$f(0) = \frac{10^0 e^{-10}}{0!} = 4.54\times10^{-5}$$

$$f(1) = \frac{10^1 e^{-10}}{1!} = 4.54\times10^{-4}$$

$$f(2) = \frac{10^2 e^{-10}}{2!} = 2.27\times10^{-3}$$

$$f(3) = \frac{10^3 e^{-10}}{3!} = 7.56\times10^{-3}$$

$$f(0)+f(1)+f(2)+f(3) = 0.0103$$

<u>d</u>

$$\mu = \frac{10\times15}{60} = 2.5 \text{ per}$$

$$f(0) = \frac{25^0 e^{-2.5}}{0!}$$

$$= 0.821$$

So $f(x\geq1) = 1 - f(0) = 0.92$

Mean = Variance = μ =

$X \sim N(\underset{\mu}{300}, \underset{\sigma}{12})$

17-04-2023

Lecture-18

David-Ray-Anderson:- books, 2page needs to be printed out.

Normal distribution

Normal	Binomial	Poission	Exponential
1) Continues	discret	dis	Continus
2) Symmetric	—	—	Right.S
3) $-\infty, \infty$	$[0, n]$	$[0, \infty)$	$[0, \infty]$
4) μ, σ^2	n, p	μ	θ

Normal distribution Function: $\frac{1}{\sqrt{2\pi}\,\sigma^2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

→ have to conv. $\underset{norm}{x} \rightarrow \underset{Std.\ norm}{z}$

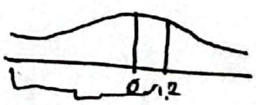
$f(z) = \frac{1}{2\pi} e^{-\frac{z^2}{2}}$

$z = \frac{x-\mu}{\sigma}$

mean, $\mu = 0$
Variance $= 1$
Std dev $= 1$

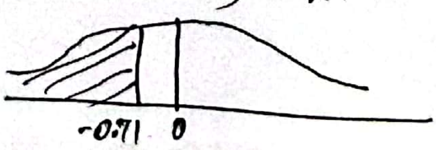
→ values are for less then type in the table.
→ have to round up in 2 decimal places
→ For continues distribution $P(x=a) = 0$

P: 298 (12) (e)



$P(z<1.2) = P(z<1.20) = 0.8849.$

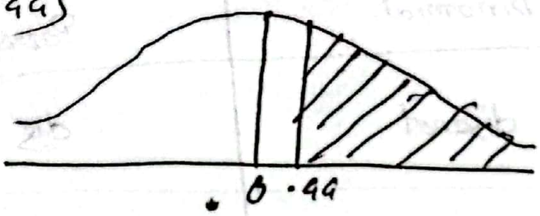
f) $P(z \leq -0.71)$ 0.2198



c)

$P(z > .44)$

$\Rightarrow 1 - P(z > .44)$

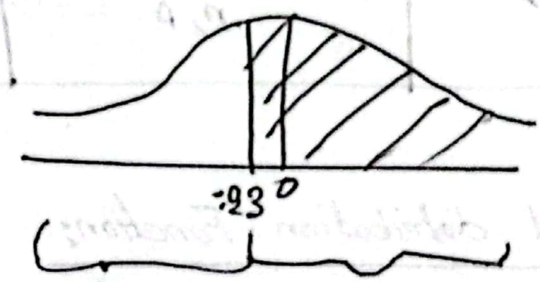


$= 1 - .6700$

$= 0.33$

Ans

d) $P(z \geq -.23)$



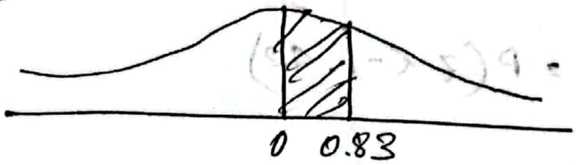
$\Rightarrow 1 - P(z \geq -.23)$

$= 1 - .4090 = \cancel{0.5871}$

4090 58.71%

a) $P(0 \le z \le 0.83)$

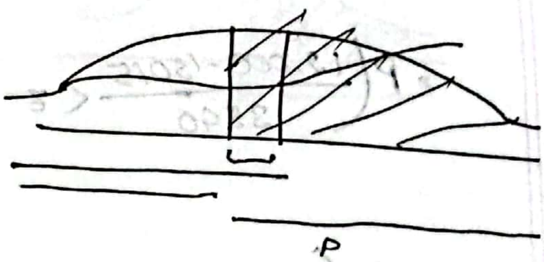
$\Rightarrow P(z < 0.83) - P(z < 0)$



$= .7967 - .5$

$= 0.2967$

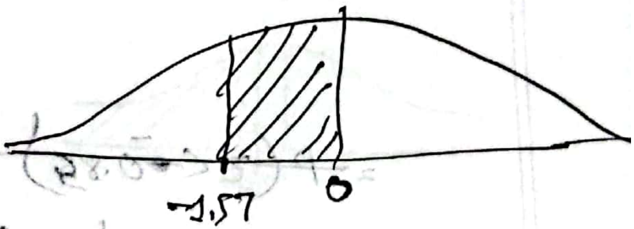
b) $P(-1.57 \le z \le 0)$



$P(z<0) - P(z < -1.57)$

$= .5 - .0582$

$=$



17

a) $P(x > 18000)$

$P\left(\frac{x-\mu}{\sigma} > \frac{18000-15015}{3540}\right)$

$P(z > 0.84) = 1 - P(z > 0.84) = 1 - .7995 = 0.2005$

b $P(x < 10000)$

$= P\left(z < \dfrac{10.000 - 15015}{3540}\right)$

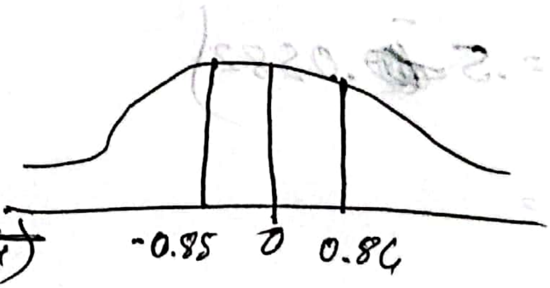
$= P(z < -1.42)$

$= 0.0778$

c) $P(12000 \leq x \leq 18000)$

$= P\left(\dfrac{12,000 - 15015}{3540} < z < \dfrac{18000 - 15015}{3540}\right)$

$= P(-0.85 < z < 0.84)$



$= \cancel{P(z < 0.85) - P(z < 0.84)}$

$= P(z < 0.84) - P(z < -0.85)$

$= \cancel{.1977}\ .7995 - .1977$

$= 0.6018$

03-05-2023

Lecture - 19

Q. 18

a

$\mu = 30$

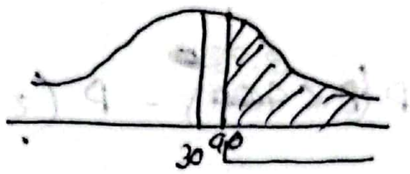
$\sigma = 8.2$

$P(x > 40)$

$= P\left(z > \frac{40-30}{8.2}\right)$

$= P(z > 1.22)$

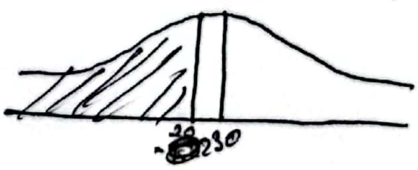
$= 1 - P(z < 1.22)$



$= 1 - 0.8888$

$= 0.1112$

b) $P(x \leq 20)$



$P\left(z < \frac{20-30}{8.2}\right)$

$P(z < -1.22)$

$= 0.1112$

c) $P(x > a) = 0.1$

$P(x < a) = 1 - 0.1 = 0.9$

$P\left(z < \frac{a-30}{8.2}\right) = 0.9$

$P\left(z < \frac{a-30}{8.2}\right) = P(z < 1.28)$

$a - 30/8.2 = 1.28$

$a = 40.496$

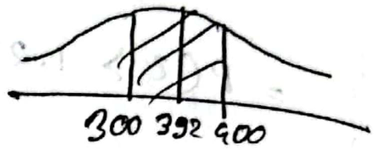
Ans: \$40.496

19

c) $\mu\ \$328 \qquad \sigma = \$92$

$P(300 < x < 400)$

$P\left(\frac{300-328}{92} < z < \frac{400-328}{92}\right)$



$P(z < 0.78) - P(z < -0.30)$

$= 0.7794 - 0.3821$

$= 0.3973$

d) 0.08

$P(x < a) = 0.08$

$P\left(z < \frac{a-328}{92}\right) = 0.92$

$P\left(z < \frac{a-328}{92}\right) = P(z < 1.41)$

$\frac{a-328}{92} = 1.41$

$a = (-1.41 \times 92) + 328$

$a = \$\ \cancel{457.72}\ 198.28$

$X \sim Exp(2)$ Exponential Probability Distribution All about time

warranty, insurance

i) Continuous

ii) $0 \le x < \alpha$

iii) Right skewed

iv) Parameter (θ)

$$f(x) = \frac{1}{\theta} e^{-x/\theta}$$

mean $= \theta$

variance $= \theta^2$

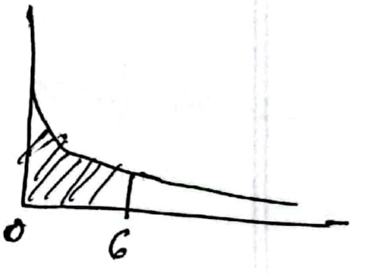
32
291

a

$$P(x \le 6) = \int_0^6 \frac{1}{8} e^{-x/8} \, dx$$

$$= \frac{1}{8} \cdot \frac{1}{-1/8} \left[e^{-x/8} \right]_0^6$$

$$= (-1)\left(e^{-6/8} - e^0\right) = 1 - e^{-6/8} = 0.528$$



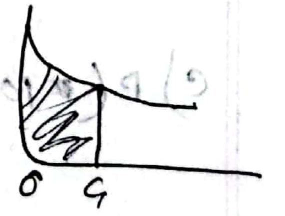
b

$$P(x < 4) = \int_0^4 \frac{1}{8} e^{-x/8} \, dx$$

$$= \frac{1}{8} \cdot \frac{1}{-1/8} \left[e^{-x/8} \right]_0^4 = (-1)\left(e^{-4/8} - e^0\right)$$

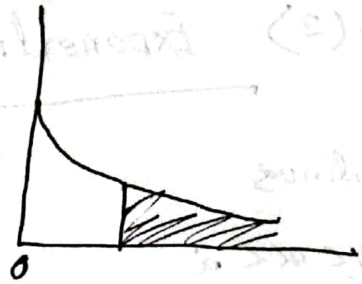
$$= 1 - e^{-4/8}$$

$$= 0.393$$



c) $P(x \geq 6) = 1 - 0.527$

$= 0.473$

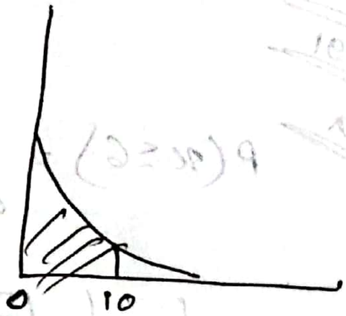


34

$\theta = 12.1$

a) $P(x<10) = \int_0^{10} \frac{1}{12.1} e^{-\frac{x}{12.1}} dx$

$= \frac{1}{12.1} \cdot \frac{1}{-1/12.1} \left[e^{-\frac{x}{12.1}} \right]_0^{10}$



$= (-1) \left(e^{-10/12.1} - e^0 \right)$

$= 1 - e^{-10/12.1}$

$= 0.562$

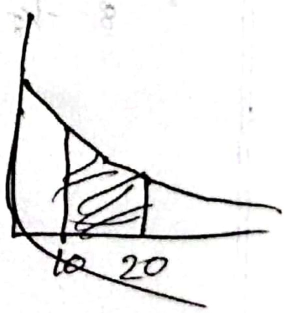
b) $P(x>20) = 1 - P(x<20) = 1 - \int_0^{20} \frac{1}{12.1} e^{-x/12.1} dx$

$= 1 - 0.809 = 0.191$

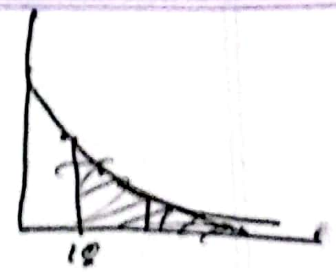
c) $P(10<x<20) = P(x<20) - P(x<10)$

$= 0.809 - 0.562$

$= 0.247$



d. $P(x>18) = 1-P(x<18)$



$$= 1 - \int_0^{18} \frac{1}{12.1} e^{-x/12.1}\, dx$$

$$= 0.226$$

11. Avg life time of a product 3 years.

* If the company can afford to offer only 10% product for warranty. what should be warranty period.

$P(x<a) = 0.1$

$$\Rightarrow \int_0^a \frac{1}{3} e^{-x/3}\, dx = 0.1$$

$$\Rightarrow \frac{1}{3} \cdot \frac{1}{-\frac{1}{3}} \left[e^{-x/3}\right]_0^a = 0.1$$

$$\Rightarrow (-1)\left(e^{-a/3} - e^0\right) = 0.1$$

$$\Rightarrow 1 - e^{-a/3} = 0.1$$

$$\Rightarrow e^{-a/3} = 0.9$$

$$\Rightarrow -a/3 = \log(0.9)$$

$$\Rightarrow a = -3\log(0.9) = 0.3161 \text{ years}$$

Stochastic Process

Syn: Random
Ant: Deterministic

$\{X(t)\}$

Canbe discret/continus
- t = time → Parameter space
- x = variable → State space
- it is a set.

⇒ Stochastic process is the collection of random variables which moves with time.

sec, min → contineus
hour or more → discret

dis	con
count	measure
can't be decimal	can be decim

4 types:

discret	parameter	~~Conti~~ discret	variable.
"	"	continus	"
Continus	"	discret	"
"	"	continus	"

Discret par Discret var

Discret Discret Markove chain.

Short: Markove chain.

$$P(x_3 = G \mid x_2 = B) = P(x_3 = G \mid x_2 = A, x_1 = B, x_{12} = B)$$

Stationary Process:

↳ time gap fixed.

$$P_{BG} = P(x_3 = G \mid x_2 = B)$$

Clear = 0

Rain = 1

$$\begin{array}{cc|cc|} & & \multicolumn{2}{c}{\text{Next day}} \\ & & 0 & 1 \\ \text{Prev day} & 0 & P_{00} & P_{01} \\ & 1 & P_{10} & P_{11} \end{array}$$

One step transition probability matrix

One stem TPM

Condition:

i) $0 \leq \text{Prob} \leq 1$

ii) Row sum = 1 ; $\Sigma P_{ij} = 1$; op. on j

18-2

a

$P_{10} = P_{RC} = 0.3$

$P_{01} = P_{CR} = 0.1$

Matrix of $P =$

$$\begin{array}{cc|cc|} & & \multicolumn{2}{c}{\text{Next day}} \\ & & 0 & 1 \\ \text{Previous day} & 0 & 0.9 & 0.1 \\ & 1 & 0.3 & 0.7 \end{array}$$

18.3

18-7

$$P = \begin{array}{c|cccc|} & \multicolumn{4}{c}{\text{Next}} \\ & 1 & 2 & 3 & 4 \\ \text{Prev } 1 & 0 & P & 0 & 1-P \\ 2 & 1-P & 0 & P & 0 \\ 3 & 0 & 1-P & 0 & P \\ 4 & P & 0 & 1-P & 0 \end{array}$$

here,

prev-next = P

Next - prev = $1-P = q$

18-2

b

Steady state / long-run / recurrent prob.

π_0 or P_0

π_1 or P_1

$$\pi_0 + \pi_1 = 1 \quad \text{—— (I)}$$

$$\pi_0 = 0.9 \times \pi_0 + 0.3\pi_1 \quad \text{—— (II)}$$

$$\pi_1 = 0.1 \times \pi_0 + 0.7\pi_1$$

$$\pi_0 = 0.9(1-\pi_1) + 0.3\pi_1$$

$$1-\pi_1 = 0.9 - 0.9\pi_1 + 0.3\pi_1$$

$$1-\pi_1 = 0.9 - 0.6\pi_1$$

$$0.1 = -0.6\pi_1 + \pi_1$$

$$0.1 = -0.4 \times \pi_1$$

$$\pi_0 = 0.9\pi_0 + 0.3(1-\pi_0)$$

$$= 0.9\pi_0 + 0.3 - 0.3\pi_0$$

$$\pi_0 = 0.6\pi_0 + 0.3$$

$$\pi_0 - 0.6\pi_0 = 0.3$$

$$\pi_0(1-0.6) = 0.3$$

$$\pi_0 = \frac{0.3}{0.4}$$

$$= 0.75$$

Three step TPM

$$P \times P \times P = \begin{vmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{vmatrix} \times \begin{vmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{vmatrix} \times \begin{vmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{vmatrix}$$

$$= \begin{vmatrix} 0.804 & 0.196 \\ 0.588 & 0.412 \end{vmatrix}$$

Lecture - 21

Queueing Process

→ Arrival (Poisson) (constant arrival rate λ)

M/M/1 model: (single server model)

→ service (Exponential) (constant arrival rate θ)

→ both are in rate.

anything / unit time

Ex-9

$\lambda = 6$ / hour

$\theta = 20$/ hour

$W_q = \frac{\lambda}{\theta(\theta-\lambda)} = \frac{6}{20(20-6)} = 0.021$

$\cancel{\pi_0 = 1 - \lambda}$

(3 in file will be busy)

a. probability of busy server $(1-\pi_0) = \frac{\lambda}{\theta} = \frac{6}{20}$

b. $L_q = \frac{\lambda^2}{\theta(\theta-\lambda)} = 0.129 = \frac{6^2}{20(20-6)}$

c. $\cancel{\frac{\lambda}{W_q}}$ W_q

$P(T_{W_q} > 3) = e^{-3/0.021}$

$$W_q = \frac{\lambda}{20(20-\lambda)} = \frac{\lambda}{400-20\lambda} \geq 3$$

$$\Rightarrow \lambda \geq 1200 - 60\lambda$$

$$\Rightarrow \lambda \geq \frac{1200}{61} = 19.67/\text{hour}$$

So, The arrival rate should be increased ~~to~~ by 13.67/h or more.

d

$$P(T_{wq} > t) = e^{-\frac{t}{\alpha w}}$$

e

$$P(T_w > \frac{1}{c}) = 0.097$$

f.

$$1 - \pi_0 = \frac{\lambda}{\theta} = \frac{6}{20} = 0.3 = 30\%$$

So, 7.2 hours the phone is in use ~~hour~~ in a day.

Hypothesis Testing

* In Hypothesis statement we will will use population parameter

Steps:

1: Hypothesis statement

Null Hypothesis H_0 : $\mu = 35$

Alternative " H_a : $\mu \neq 35$

Way 1: claim statement always goes to null H_0.

Way 2: not claim .... null

$S \longrightarrow \sigma$

$S^2 \longrightarrow \sigma^2$

$\bar{X} \longrightarrow \mu$

$P \longrightarrow \pi$ (prob)

Decision

usually $\alpha = 5\%$

	H_0 rejected	H_0 accepted
H_0 true	type-1 error	Right
H_0 false	Right	type-2 error

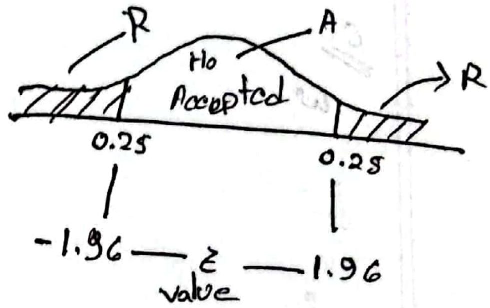
type I error, $\alpha = P$(Rejecting true H_0)

↳ Probability

Step-2:

$\alpha = 5\% = 0.05$

critical value; will come from distribution



Step-3:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

for proporssion;

$$Z = \frac{P - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

Step-4:

decision;