Statistics for Engineers

Sharadindu Adhikari | 19BCE2105

Theory Assignment

Q1. Calculate Median and Arithmetic mean from the given series.

Sol^n:

x	f	cf	d = mid x	fd
0 — 10000	5	5	5000	25000
10000 — 15000	5	10	12500	62500
15000 — 20000	3	13	17500	52500
20000 — 25000	4	17	22500	90000
✓ 25000 — 30000	4	21	27500	110000
30000 — 35000	4	25	32500	130000
35000 — 40000	9	34	37500	337500

$\Sigma f = 34$

$\Sigma fd = 807500$

$\therefore$ Mean $= \dfrac{\Sigma fd}{\Sigma f} = \dfrac{807500}{34} = 23071.42$

Median $= \left(\dfrac{34+1}{2}\right)^{th} = 17.5^{th}$

$\therefore$ Median $= L + \dfrac{h}{f}\left(\dfrac{\Sigma f}{2} - cf\right)$

$= 25000 + \dfrac{5000}{4}\left(\dfrac{34}{2} - 17\right)$

$= 25000$

Q2. Calculate the median and mode (by grouping method) from the given data.

Sol^n:

central size :	5	15	25	35	45	55	65	75
Frequency :	5	9	13	21	20	15	8	3

∴

Central size (x)	Frequency (f_1)	cf	f_2	f_3	f_4	f_5	f_6
5	5	5					
15	9	14	14				
25	13	27		21	27	23	54
35	21	48	34				
45	20	68		41			
55	15	83	35		56	43	
65	8	91		23			26
75	3	94	11				

$\Sigma f_1 = 94$

Median $\Rightarrow$ $\left(\frac{94+1}{2}\right)^{th}$ term $\Rightarrow$ $(97.5)^{th}$ term

$\boxed{\therefore \text{ Median } = 35}$

and Mean, $= \frac{\Sigma x f_i}{\Sigma f_i} = \frac{3690}{94} = 39.25$

$x f_i$
25
135
325
735
900
825
520
225
3690

	35	45	55
f_1	1		
f_2		1	1
f_3	1	1	
f_4	1	1	1
f_5	1	1	1
f_6	1	1	
	$\Sigma 5$	$\Sigma 5$	$\Sigma 3$

because it is bimodal class

$\therefore$ Mode = 3 Median − 2 Mean

$= (3 \times 35) - (2 \times 39.25)$

$= 105 - 78.5$

$\therefore \boxed{\text{Mode} = 26.5}$

Q3. A consignment of 190 articles is classified according to the size of the article. Find the standard deviation and its coefficient.

Sol:- P.t.o. →

19BCE2105

X	f	x = mid	d = x - 45	d^2	fd	fd^2	fx
0–10	6	5	−40	1600	−480	9600	30
10–20	5	15	−30	900	−150	4500	75
20-30	14	25	−20	400	−280	5600	210
30-40	45	35	−10	100	−450	4500	1575
40-50	8	45	0	0	0	0	360
50-60	22	55	10	100	220	2200	1210
60-70	48	65	20	400	960	19200	3120
70-80	22	75	30	900	660	19800	1650
80-90	20	85	40	1600	800	32000	1700
90-100	0	95	50	2500	0	0	0

$\Sigma f = 190$ $\quad \Sigma fd = 1280$ $\quad \Sigma fx = 9930$

$\Sigma fd^2 = 97400$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$

$$\sigma = \sqrt{\frac{97400}{190} - \left(\frac{1280}{190}\right)^2}$$

$$\boxed{\sigma = 21.61}$$ → standard deviation

Mean, $\mu = \frac{\Sigma fx}{\Sigma f}$

$= \frac{9930}{190}$

$$\boxed{\mu = 52.26}$$

∴ coefficient of variation, CV $= \frac{\sigma}{\mu} \times 100$

$= \frac{21.61}{52.26} \times 100$

$$\boxed{CV = 41.85\ \%}$$

19BCE2105

Q4. The scores of 2 batsmen X and Y in ~~a~~ nine innings are given. Find (using coefficient of variation) which of the 2 batsman is more consistent in scoring.

Sol

X	Y	$X-\bar{X}$	$Y-\bar{Y}$
43	56	−3.78	3
45	34	−5.78	−19
23	87	16.22	34
54	24	−14.78	−29
65	98	−25.78	45
34	56	5.22	3
54	10	14.78	−43
23	67	16.22	14
12	45	27.22	−8
$\Sigma X = 353$	$\Sigma Y = 477$		

$$\bar{X} = \frac{353}{9} = 39.22 \quad ; \quad \bar{Y} = \frac{477}{9} = 53$$

$$\sigma_x = \sqrt{\frac{1}{9}\left(14.2884 + 33.4089 + 218.4484 + 664.6084 + 218.4484 + 263.0889 + 27.2484 + 268.0884 + 740.9284\right)}$$

$$\therefore \sigma_x = 18.81$$

$\therefore$ % CV of $x = 100 \times \frac{18.81}{39.22} = 47.96\%$

lly,

$$\sigma_y = \sqrt{\frac{1}{9}\left(9 + 361 + 1136 + 841 + 2025 + 9 + 1849 + 196 + 64\right)}$$

$\therefore \quad \sigma_y = 26.89$

$\therefore$ % CV of $y = 100 \times \frac{26.89}{53} = 50.73\%$

$\because \quad CV(x) < CV(y)$

$\Longrightarrow$ $\boxed{X \text{ is more consistent.}}$

Q5. Let X be a continuous random variable with probability density function $f(x)$:

(i) Determine the constant a.

(ii) Compute $P(X \leq 1.5)$

Sol^n

$$f(x) = \begin{cases} ax & 0 \leq x \leq 1 \\ a & 1 \leq x \leq 2 \\ -ax + 3a & 2 \leq x \leq 3 \end{cases}$$

We know that,

$$\int_0^1 ax\,dx + \int_1^2 a\,dx + \int_2^3 (-ax+3a)\,dx = 1$$

$$\Rightarrow \frac{a}{2} + a - \frac{5a}{2} + 3a = 1$$

$$\therefore \boxed{a = 1/2}$$

$$P(x \le 1.5) = \int_0^1 \frac{x}{2}\,dx + \int_1^{1.5} \frac{dx}{2}$$

$$\boxed{P(x \le 1.5) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}}$$

Q6.

Sol^n: $f(x) = 6x(1-x) \qquad 0 \le x \le 1$

(i) $\int_0^1 6x(1-x)\,dx = 6\int_0^1 (x-x^2)\,dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1$

$= 1$

In the given interval, probability is 1.

$\therefore$ Given function is pdf.

(ii) CDF $\rightarrow$ $\int_0^x 6x(1-x)\,dx = 6\left(\frac{x^2}{2} - \frac{x^3}{3}\right)_0^x$

$= x^2(3-2x)$

(iii) $P\left(X \le \frac{1}{2} \,/\, \frac{1}{3} \le x \le \frac{2}{3}\right) = \dfrac{P\left(\frac{1}{3} \le X \le \frac{1}{2}\right)}{P\left(\frac{1}{3} \le X \le \frac{2}{3}\right)}$

$$P\left(\frac{1}{3} \le X \le \frac{2}{3}\right) = \int_{1/3}^{2/3} 6x(1-x)\,dx = 6\left(\frac{2}{9} - \frac{8}{81} - \frac{1}{18} + \frac{1}{81}\right)$$

$$= \frac{13}{27} = 0.481$$

$$\therefore P\left(X \le \frac{1}{2} \,/\, \frac{1}{3} \le X \le \frac{2}{3}\right) = \frac{0.267}{0.481} = 0.55$$

(iv) $P(X < K) = P(X > K)$

$$\int_0^K 6x(1-x)\,dx - \int_K^1 6x(1-x)\,dx = 0 \quad \text{———} \; (1)$$

But

$$\int_0^K 6x(1-x)\,dx + \int_K^1 6x.(1-x)\,dx = 1 \quad \text{———} \; (2)$$

adding (1) and (2)

$$2\int_0^K (6x(1-x))\,dx = 1$$

$\Rightarrow 6K^2 - 4K^3 = 1$

$\Rightarrow 4K^3 - 6K^2 + 1 = 0$

$K = \underbrace{1.36, -0.36}_{\text{not in range}}, 0.5$ $\quad \therefore \boxed{K = \frac{1}{2}}$

19BCE2105

Q8. Sol^n :

3 coins are tossed

X is tails on first two

Y is heads on first two

	TTT	HTT	TTH	THH	HHT	HHH	THT	HTH
X	2	1	2	1	0	0	1	1
Y	0	0	1	2	1	2	1	1

X \ Y	0	1	2	P(x)
0	0	1/8	1/8	2/8
1	1/8	2/8	1/8	4/8
2	1/8	1/8	0	2/8
P(y)	2/8	4/8	2/8	

$\therefore E(x) = \Sigma x \cdot P(x) = 0 \times P(0) + 1 \times P(1) + 2 \times P(2)$

$= 0 + \frac{4}{8} + 2 \times \frac{2}{8}$

$= 1$

|||ly, $E(y) = \Sigma y \cdot P(y) = 0 \times P(0) + 1 \times P(1) + 2 \times P(2)$

$= 0 + \frac{4}{8} + \frac{2 \times 2}{8}$

$= 1$

19BCE2105

$\therefore \quad cov(x,y) = \cancel{E}\, E(xy) - E(x) \cdot E(y)$

$E(xy) = \sum_y \sum_x x \cdot y \cdot (P_{ij})$

$= \sum_y \left(0 \times y \left(p(0,y) \right) + 1 \times y \left(p(1,y) \right) + 2 \times y \times \left(p(2,y) \right) \right)$

$= \cancel{0} \; 0 + 0 + 0 + 0 + 1 \times \frac{2}{8} + 1 \times 2 \times \frac{1}{8} + 0 + 2 \times 1 \times \frac{1}{8} + 4 \times \frac{1}{8}$

$= 5/4$

$= 1.25$

$\therefore \quad cov(x,y) = 1.25 - 1 = 0.25$

Q9. Sol^n

(i) $f_y(y) = \int_0^1 4x^3 y^2 \, dx$

$= y^2$

$f_x(x) = \int_0^1 4x^3 y^2 \, dy = \frac{4}{3} x^3$

as $f_x(x) \times f_y(y) \neq f(x,y)$

$\therefore$ not independent

19BCE2105

$$f(x/y) = \frac{f(x,y)}{f_y(y)} = \frac{4x^3y^2}{y^2} = 4x^3$$

$$f(y/x) = \frac{f(x,y)}{f_x(x)} = \frac{4x^3y^2}{\frac{4x^3}{3}} = 3y^2$$

(ii) $$P\left(X < \frac{1}{3}\right) = \int_0^{1/3} \left(\int_0^1 4x^3y^2\,dy \right) dx$$

$$= \frac{4}{3} \int_0^{1/3} x^3\,dx$$

$$= \frac{4}{3} \times \frac{1}{4} \times \left(\frac{1}{3}\right)^4$$

$$= \left(\frac{1}{3}\right)^5$$

$$= 0.004115$$

(iii) $$P\left(X < \frac{1}{2},\ Y < \frac{1}{3}\right) = \int_0^{1/2} \int_0^{1/3} 4x^3y^2\,dx\,dy$$

$$= \int_0^{1/2} \frac{4x^3}{3} \times \frac{1}{9}$$

$$= \frac{4}{27} \times \frac{1}{4} \times \left(\frac{1}{2}\right)^4$$

$$= \frac{1}{432}$$

$$= 0.002314$$

—★—

19BCE2105

Q10. Sol^n.

Karl pearson correlation coefficient (r) -

$$= \frac{n\sum xy - \sum x \cdot \sum y}{\sqrt{n\sum x^2 - (\sum x)^2} \cdot \sqrt{n \cdot \sum y^2 - (\sum y)^2}}$$

X	Y	x^2	y^2	XY
17	3	289	9	51
20	11	400	121	220
25	17	625	289	425
34	23	1156	529	782
42	15	1764	225	630
45	51	2025	2601	2295
48	21	2304	441	1008
54	35	2916	1225	1890
$\sum x = 285$	$\sum y = 176$	$\sum x^2 = 11479$	$\sum y^2 = 5440$	$\sum xy = 7301$

$$r = \frac{8 \times 7301 - \cancel{280}\ 285 \times 176}{\sqrt{\left(8 \times 11479 - (285)^2\right)\left(8 \times 5440 - (176)^2\right)}}$$

$$= \frac{8248}{\sqrt{10607 \times 12544}}$$

$$= \frac{8248}{11534.9125}$$

$$= \frac{8248}{11534.9125}$$

$$= 0.71504$$

END