Date: / /

Name Ghazala
IDNO 14093
Subject Biostatistics

(a) Calculate the correlation between X and Y

X	3	4	5	6	7	8	9	10	11	13
Y	25	24	20	20	19	17	~~18~~ 16	13	10	8

Solution:

Let us change the origin of X and Y.
Hence
$U = X - 7$ and $V = Y - 19$, Then $r_{xy} = r_{uv}$.
The calculation needed to find r are give below

X	Y	u	v	u^2	v^2	uv
3	25	-4	6	16	36	-24
4	24	-3	5	9	25	-15
5	20	-2	1	4	1	-2
6	20	-1	1	1	1	-1
7	19	0	0	0	0	0
8	17	1	-2	1	4	-2
9	16	2	-3	4	9	-6
10	13	3	-6	9	36	-18
11	10	4	-9	16	81	-36
13	8	6	-11	36	121	-66
76	172	6	-18	96	314	-170

Now

As we know that

$$r = \frac{\sum uv - (\sum u)(\sum v)/n}{\sqrt{\left[\sum u^2 - \frac{(\sum u)^2}{n}\right]\left[\sum v^2 - \frac{(\sum v)^2}{n}\right]}}$$

Putting the values.

$$r = \frac{-170 - \frac{(6)(-18)}{10}}{\sqrt{\left[96 - \frac{(6)^2}{10}\right]\left[314 - \frac{(-18)^2}{10}\right]}}$$

$$r = \frac{-170 + \frac{108}{10}}{\sqrt{\left[96 - \frac{36}{10}\right]\left[314 - \frac{(324)}{10}\right]}}$$

$$r = \frac{-170 + 10.8}{\sqrt{(96 - 3.6)(314 - 32.4)}}$$

$$r = \frac{-159.2}{\sqrt{26019.84}}$$

$$r = \frac{-159.2}{161.30}$$

Date: / /

$\boxed{r = -0.98}$ Ans.

(b) Given the following set of values.

X	20	11	15	10	17	18	21	25	28
Y	5	15	14	17	8	9	12	16	18

(a) Determine the equation of least square regression

Solution:

(a) The necessary calculation for determining the equation of least square regression line are given below.

X	Y	X^2	Y^2	XY
20	5	400	25	100
11	15	121	225	165
15	14	225	196	210
10	17	100	289	170
17	8	289	64	136
18	9	324	81	162
21	12	441	144	252
25	16	625	256	400
28	18	784	324	504
165	114	3309	1604	2099

The estimate linear regression line Y on X is

$$\hat{Y} = a + bx$$

Date: / /

Where a and b are the least square estimate of the parameter α and β respectively and are given by.

As we know that

$$b = \frac{n \sum XY - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2} \quad \text{and}$$

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

Substituting the sum, we get

Putting the values.

$$b = \frac{(9)(2099) - (165)(114)}{(9)(3309) - (27225)}$$

$$b = \frac{18891 - 18810}{29781 - 27225}$$

$$b = \frac{81}{2556}$$

$$b = 0.031$$

Hence

As we know that

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

Putting the values.

$$a = \frac{114}{9} - 0.031\left(\frac{165}{9}\right)$$

Date: / /

$a = 12.66 - 0.568$

$a = 12.09$

$\boxed{\bar{Y} = 12.09 + 0.031}$ **Ans**

The estimate linear regression line

X on Y

$\bar{X} = a + b\bar{y}$

As we know that

$$b = \frac{n\sum XY - (\sum X)(\sum Y)}{\cancel{n}\ \cancel{(\sum XY}\ n(\sum Y^2) - (\sum Y)^2}$$

Putting the values.

$$b = \frac{(9)(2099) - (165)(114)}{(9)(1604) - (114)^2}$$

$$b = \frac{18891 - 18810}{14436 - 12996}$$

$$b = \frac{81}{1440}$$

$\boxed{b = 0.056}$

Hence

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

$$a = \frac{165}{9} - (0.056)\left(\frac{114}{9}\right)$$

Date: / /

$a = 18.33 - 0.709$

$\boxed{a = 17.62}$ **Ans**

(b) Find the predicated values of X = 20, 11, 15, ~~X~~ 25, 28 and X for Y = 5, 15, ~~19~~ 9, 12, 16, 18.

Solution:

Predicated values of Y ~~for~~ ~~X~~ 20, 11, 15, 25, 28

The predicated values of Y are found by substituting the values in the estimated equation. Thus X = ...

As we know that.

$\bar{Y} = 12.09 + 0.031(20)$

$\bar{Y} = 12.09 + 0.62$

$\bar{Y} = 12.71$

$\bar{Y} = 12.09 + 0.031(11)$

$\bar{Y} = 12.09 + 0.341$

$\bar{Y} = 12.43$

$\bar{Y} = 12.09 + 0.031(15)$

$\bar{Y} = 12.09 + 0.465$

$\bar{Y} = 12.55$

$\bar{Y} = 12.09 + 0.031(25)$

$\bar{Y} = 12.09 + 0.775$

$\bar{Y} = 12.86$

Date: / /

$\bar{Y} = 12.09 + 0.031(28)$
$\bar{Y} = 12.09 + 0.86$
$\bar{Y} = 12.95$

Predicated values of $X = 5, 15, 9, 12, 16, 18$
As we know that.

$\bar{X} = 17.62 + 0.056(5)$
$\bar{X} = 17.62 + 0.28$
$\bar{X} = 17.9$

$\bar{X} = 17.62 + 0.056(15)$
$\bar{X} = 17.62 + 0.84$
$\bar{X} = 18.46$

$\bar{X} = 17.62 + 0.056(9)$
$\bar{X} = 17.62 + 0.504$
$\bar{X} = 18.1$

$\bar{X} = 17.62 + 0.056(12)$
$\bar{X} = 17.62 + 0.672$
$\bar{X} = 18.2$

*
$\bar{X} = 17.62 + 0.056(16)$
$\bar{X} = 17.62 + 0.89$
$\bar{X} = 18.5$

$\bar{X} = 17.62 + 0.056(18)$
$\bar{X} = 17.62 + 1.008$
$\bar{X} = 18.6$

Date: / /

Q3 (a) Construct the ungroup frequency distribution of these data.

Solution:

No of children	No of women	CF	Tally
0	1	1	|
1	4	5	|||||
2	8	13	~~||||~~ |||
3	11	24	~~||||~~ ~~||||~~ |
4	8	32	~~||||~~ |||
5	5	37	~~||||~~
6	4	41	||||
7	3	44	|||
8	2	46	||
9	1	47	|
10	3	50	|||
Total	50		

(b) Construct the group frequency distribution of these data.

Solution:

In this group we find out

1. frequency
2. C. frequency
3. Boundaries
4. Class Mark.

Finding class interval $n = 50$

x_A mini $= 0$

x max $= 10$.

Scanned with CamScanner

$$x = \frac{10}{5} = 2$$

Class interval = 2.

Classes	Entairs	F	C.F	Class Boundries	Mid Marks
0 - 2	0, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2	13	13	0.5 - 2.5	1
2 - 5	3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5	24	37	2.5 - 5.5	4
6 - 8	4, 4, 4, 4, 7, 7, 7, 8, 8	9	46	5.5 - 8.5	7
9 - 11	9, 10, 10, 10	4	50	8.5 - 11.5	10
Total		50			

Q2 (a)

A fair coin is tossed 5 times. Find the ~~prop~~ ~~pa~~ ~~b~~ probability of obtaining various numbers of head.

Solution:

Let us regard the tossing of coin as an experiment. Then observe that.

(i) Each toss of coin (i.e each trial) has two possible outcomes, head (success) and tails (failure)

(ii) the probability of a head (success) is $p = \frac{1}{2}$ and remains the same for successive tosses;

Date: / /

(iii) the successive tosses of the coin are independent; and

(iv) the coin is tossed 5 times.

Therefore the r.v. X which denotes the number of heads (Successes) has a binomial probability distribution with p = 1/2 and n=5. The Possible value of X are 0, 1, 2, 3, 4 and 5. Hence

$$P(\text{no head}) = P(X=0) = \binom{5}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^5 = 1\times\left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

$$P(1\text{ head}) = P(X=1) = \binom{5}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^{5-1} = 5\times\left(\frac{1}{2}\right)^5 = \frac{5}{32}$$

$$P(2\text{ head}) = P(X=2) = \binom{5}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{5-2} = 10\times\left(\frac{1}{2}\right)^5 = \frac{10}{32}$$

$$P(3\text{ head}) = P(X=3) = \binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{5-3} = 10\times\left(\frac{1}{2}\right)^5 = \frac{10}{32}$$

$$P(4\text{ head}) = P(X=4) = \binom{5}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^{5-4} = 5\times\left(\frac{1}{2}\right)^5 = \frac{5}{32}, \text{ and}$$

$$P(5\text{ head}) = P(X=5) = \binom{5}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0 = 1\times\left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

These probabilities can also be obtained by expanding the binomial $\left(\frac{1}{2}+\frac{1}{2}\right)^5$. The binomial

Date: / /

Probability distribution for the number of heads obtained in 5 tosses of a fair coin is

X	0	1	2	3	4	5
$f(x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Date: / /

(b) There fore the bionomial probability ~~dit~~ distribution

$n = 10$

$p = 2/3$

$q = 1 - p$

$q = 1 - 2/3$

X denoted the number of won by A then

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

$$1 - \sum_{x=0}^{3} \binom{10}{x} \left(\frac{2}{3}\right)^{x} \left(\frac{1}{3}\right)^{9} + 45 \left(\frac{2}{3}\right)^{2} \left(\frac{1}{3}\right)^{8} + 120 \left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^{7} \Big]$$

$$1 - \frac{1}{59049} [1 + 20 + 180 + 960]$$

$$1 - 0.0197$$

$$\boxed{P(x \geq 4) = 0.9803}$$

(ii) $P(x=4) = \binom{10}{4} \left(\frac{2}{3}\right)^{4} \left(\frac{1}{3}\right)^{6}$

$$= 210 \left(\frac{16}{81}\right) \frac{1}{729}$$

$$= \frac{3360}{59049}$$

$$\boxed{P(x=4) = 0.056}$$

(iii) $P(x=11) = f(0) =$ because X can take only value.

$0, 1, 2, 3 \ldots\ldots 10$

(iv) 6 or more games

$$P(x \geq 6) = \sum_{x=6}^{10} \binom{10}{x} \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{10-x}$$

$$= \binom{10}{6}\left(\frac{2}{3}\right)^6\left(\frac{1}{3}\right)^4 + \binom{10}{9}\left(\frac{2}{3}\right)^9\left(\frac{1}{3}\right)^1 + \binom{10}{10}\left(\frac{2}{3}\right)^{10}\left(\frac{1}{3}\right)^0$$

$$= 0.228 + 0.261 + 0.196 + 0.087 + 0.018$$

$$\boxed{P(x \geq 6) = 0.79}$$