FINAL TERM PAPER

SUBJECT: BIO STATISTIC

SUBMITTED TO: DR.ANWAR SHAMIM

SUBMITTED BY: RABIA ZAMAN

DEPARTMENT: DT

IDNO# 14180

Date: 22/06/2020

Question N0.1 Calculate the correlation co-efficient between X and Y.

n	Price(X)	Demand(Y)
1	3	25
2	4	24
3	5	20
4	6	20
5	7	19
6	8	17
7	9	16
8	10	13
9	11	10
10	13	8

Solution :

As we know that Correlation between X and Y is given below

$$r = \frac{n\sum xy - \sum x.\sum y}{\sqrt{n\sum x^2 - (\sum x)^2(n\sum y^2 - (\sum y)^2}} \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_}(1)$$

n	Price(X)	Demand(Y)	x^2	y^2	X*Y
1	3	25	9	625	75
2	4	24	16	576	96
3	5	20	25	400	100
4	6	20	36	400	120
5	7	19	49	361	133
6	8	17	64	289	136
7	9	16	81	256	144
8	10	13	100	169	130
9	11	10	121	100	110
10	13	8	169	64	104
Total	76	172	670	3240	1148

Now put all the values from table in equation no .1

$$r = \frac{10 * 1148 - 76 * 172}{\sqrt{10 * 670 - (76)^2 * (10 * 3240 - (172)^2}}$$

$$r = -\frac{1592}{1613.067} = -0.988694$$

The above value of r shows strongly negative correlation in price and demand of a particular commodity.

Question no.1(b)

The regression equation y on x is given below

$$Y = a + bx$$

Where

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

and

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

N	X	y	XY	x^2
1	20	5	100	400
2	11	15	165	121
3	15	14	210	225
4	10	17	170	100
5	17	8	136	289
6	18	9	162	324
7	21	12	252	441
8	25	16	400	625
9	28	18	504	784
Summation	165	114	2099	3309

First we fine value of b

$$b = \frac{9 * 2099 - 165 * 114}{9 * 3309 - (165)^2}$$

$$b = \frac{81}{2556} = 0.0316$$

Now we find value of a

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

$$\overline{y} = 12.6667$$

$$\overline{x} = 18.333$$

$$a = 12.6667 - 0.0316 * 18.333 = 12.40$$

$$Y = 12.40 + 0.0316X$$

Now find the regression of X on Y

$$X = a + bY$$

Where

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

and

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

n	x	y	XY	Y^2
1	20	5	100	25
2	11	15	165	225
3	15	14	210	196
4	10	17	170	289
5	17	8	136	64
6	18	9	162	81
7	21	12	252	144
8	25	16	400	256
9	28	18	504	324
Summation	165	114	2099	1604

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

$$b = \frac{81}{1440} = 0.05625$$

Now we find value of a

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

$$\overline{y} = 12.6667$$

$$\overline{x} = 18.333$$

$$a = 18.333 - 0.05625 * 12.6667 = 17.62049$$

$$X = 17.62049 + 0.05625Y$$

Now to find different values of Y put the value of X in regression equation Y on X.

$$Y = 12.40 + 0.0316X$$

For X=20,11,15,25,28

For X=20

$$Y = 12.40 + 0.0316 * 20 = 13.032$$

For X=11

$$Y = 12.40 + 0.0316 * 11 = 12.747$$

For X=15

$$Y = 12.40 + 0.0316 * 15 = 12.874$$

For X=25

$$Y = 12.40 + 0.0316 * 25 = 13.19$$

For X=28

$$Y = 12.40 + 0.0316 * 28 = 13.2848$$

Now to find different values of X put the value of X in regression equation X on Y.

$$X = 17.62049 + 0.05625Y$$

For Y=5,15,9,12,16,18

For Y =5

$$X = 17.62049 + 0.05625 * 5 = 17.901$$

For Y =15

$$X = 17.62049 + 0.05625 * 15 = 18.4624$$

For Y =9

$$X = 17.62049 + 0.05625 * 9 = 18.070$$

For Y =12

$$X = 17.62049 + 0.05625 * 12 = 18.295$$

For Y =16

$$X = 17.62049 + 0.05625 * 16 = 18.520$$

For Y =18

$$X = 17.62049 + 0.05625 * 18 = 18.632$$

Question no .2

Find the following

a) A fair of coin is tossed 5 times. Find the probability of obtaining various number of heads
b) A and B play a game in which A's Probability of winning is 2/3 . In series of 10 games,what is the probability that A will win (i) at least four games.(ii) Exactly equal to 4/10 games.(iii) Exactly equals to 11 games.(iv) 6 or more games

Solution :

As we know that when a coin is tossed it has two possible outcomes, Head or tail.

So, the probability of a head or tail is $p = \frac{1}{2}$. So here we are interested only in head so when a coin is tossed the probability of head is $p = \frac{1}{2}$

This probability remains the same for successive toss.

The successive tosses are independent

The coin is tossed 5 times.

Therefore the random value of X which denotes the number of head has a binomial distribution with $p = \frac{1}{2}$ and n=5. The possible value of X are 0,1,2,3,4,5

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

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

P(2 heads)= $P(X = 2) \binom{5}{2} (1/2)^2(1/2)^{5-2} = 10 \times (1/2)^5 = \frac{10}{32}$

P(3 heads)= $P(X = 3) \binom{5}{3} (1/2)^3(1/2)^{5-3} = 10 \times (1/2)^5 = \frac{10}{32}$

P(4 heads)= $P(X = 4) \binom{5}{4} (1/2)^4(1/2)^{5-4} = 5 \times (1/2)^5 = \frac{5}{32}$

P(5 heads)= $P(X = 5) \binom{5}{4} (1/2)^5(1/2)^{5-5} = 1 \times (1/2)^5 = \frac{1}{32}$

These probability can also be obtained by using binomial probability distribution by expanding the binomial probability distribuiton $(\frac{1}{2} + \frac{1}{2})^5$

x	0	1	2	3	4	5
f(x)	1/32	5/32	10/32	10/32	5/32	1/32

Part (b)

Given data

Total number of games =10

Probability of A's win = P(A)=2/3=p

Probability of B's win = P(A)=1/3=q

The successive games are independent won of lost

Let X denote the number of games won by A then

(i) A will win at least 4 games.

$$P(x \geq 4) = 1 - P(x \leq 3)$$

$$P(x \geq 4) = 1 - (P(x \leq 3 + P(x \leq 2) + P(x \leq 1) + P(x \leq 0))$$

$$P(x \geq 4) = 1 - \binom{10}{3}(2/3)^3\left(\frac{1}{3}\right)^7 + \binom{10}{2}(2/3)^2\left(\frac{1}{3}\right)^8 + \binom{10}{1}(2/3)^1\left(\frac{1}{3}\right)^9 + \binom{10}{0}(2/3)^0\left(\frac{1}{3}\right)^{10}$$

$$P(x \geq 4) = 1 - (0.01625 + 0.000304 + 0.000338 + 0.0000169))$$

$$P(x \geq 4) = 0.98309$$

(ii) $Exactly\ 10\ by\ 4\ games$

$$P(x=4)=\binom{10}{4}(2/3)^4\left(\frac{1}{3}\right)^4=\frac{1120}{2187}=0.512117$$

(iii) *We cannot find the probability of exactly equal to 11 games because total number of games is equal to 10.*

(iv) 6 or more games

$$P(x\geq 6)=\binom{10}{6}(2/3)^6\left(\frac{1}{3}\right)^4+\binom{10}{7}(2/3)^7\left(\frac{1}{3}\right)^3+\binom{10}{8}(2/3)^8\left(\frac{1}{3}\right)^2$$
$$+\binom{10}{9}(2/3)^9\left(\frac{1}{3}\right)^1+\binom{10}{10}(2/3)^{10}\left(\frac{1}{3}\right)^0$$

$$P(x\geq 6)=0.2276+0.2601+0.1950+0.0867+0.01734=0.7867$$

Quesiton no 3

The ungrouped frequency distribution is given below

n	Tally	Frequency
0	|	01
1	||||	04
2	||||||||	08
3	|||||||||||	11
4	||||||||	08
5	|||||	05
6	||||	04
7	|||	03
8	||	02
9	|	01
10	|||	03
Total		50

The group frequency distribution is given below

Maximum value =10

Minimum value =0

Range =10-0=10

Number of classes $k = 1 + 3.3\ log\ N$

$$k = 1 + 3.3\ log\ (50)$$

$$k = 1 + 3.3 * 1.6989 = 6.6066 = 6$$

Class interval = Range /Number of classes= 10/6=1.66=2

The group distribution table is given below

No.of classes	Classes	Frequency
1	0-1	5
2	2-3	19
3	4-5	13
4	6-7	7
5	8-9	3
6	10-11	3
	Total	50