PATAN MULTIPLE CAMPUS

Patandhoka, Lalitpur

BCA 3rd semester

Probability and Statistics(CACS-202)

Presented by :- Biraj Acharya

INDEX PAGE AND DETAILS

S.N	Lab Works	Date	Signature
1.	Individual Series	2082/01/11	
2.	Individual Series & Consistency	2082/01/11	
3.	Discrete Series	2082/01/11	
4.	Correleation Coefficent	2082/01/16	
5.	Regression Equation	2082/01/16	
6.	One-Way ANOVA	2082/01/16	
7.	Two-Way ANOVA	2082/01/22	
8.	Binomial Distribution	2082/01/25	
9.	Poisson Distribution	2082/01/30	
10.	Normal Distribution	2082/01/31	

Patan Multiple Campus

Practical No. 1: Individual Series

Level:- TU BCA 3rd Semester

Date;- 2081/01/11

Subject:- Probability and Statistics (CACS-202)

Roll No:- 13/080

Name: Bimal Acharya

Question:-

Find no. of observations, minimum value, maximum value, mean, median, mode, quartiles, range, standard deviation, variance and coefficient of variation of the following observations:

100, 80, 79, 50, 52, 80, 79, 80, 105, 79, 82, 80, 84, 64, 70, 71, 80, 45, 54, 80, 95, 80

Working Expression:-

n = no. of observation

S = minimum value

L = maximum value

Mean $(\bar{x}) = \frac{\Sigma x}{n}$

Median (Md) = $\left(\frac{n+1}{2}\right)^{th}$ item value

Mode (M_o) = Maximum repetation number

First Quartile $(Q_1) = \left(\frac{n+1}{4}\right)^{th}$ item value

Third Quartile $(Q_3) = 3\left(\frac{n+1}{4}\right)^{th}$ item value

Range $(R) = L - S$

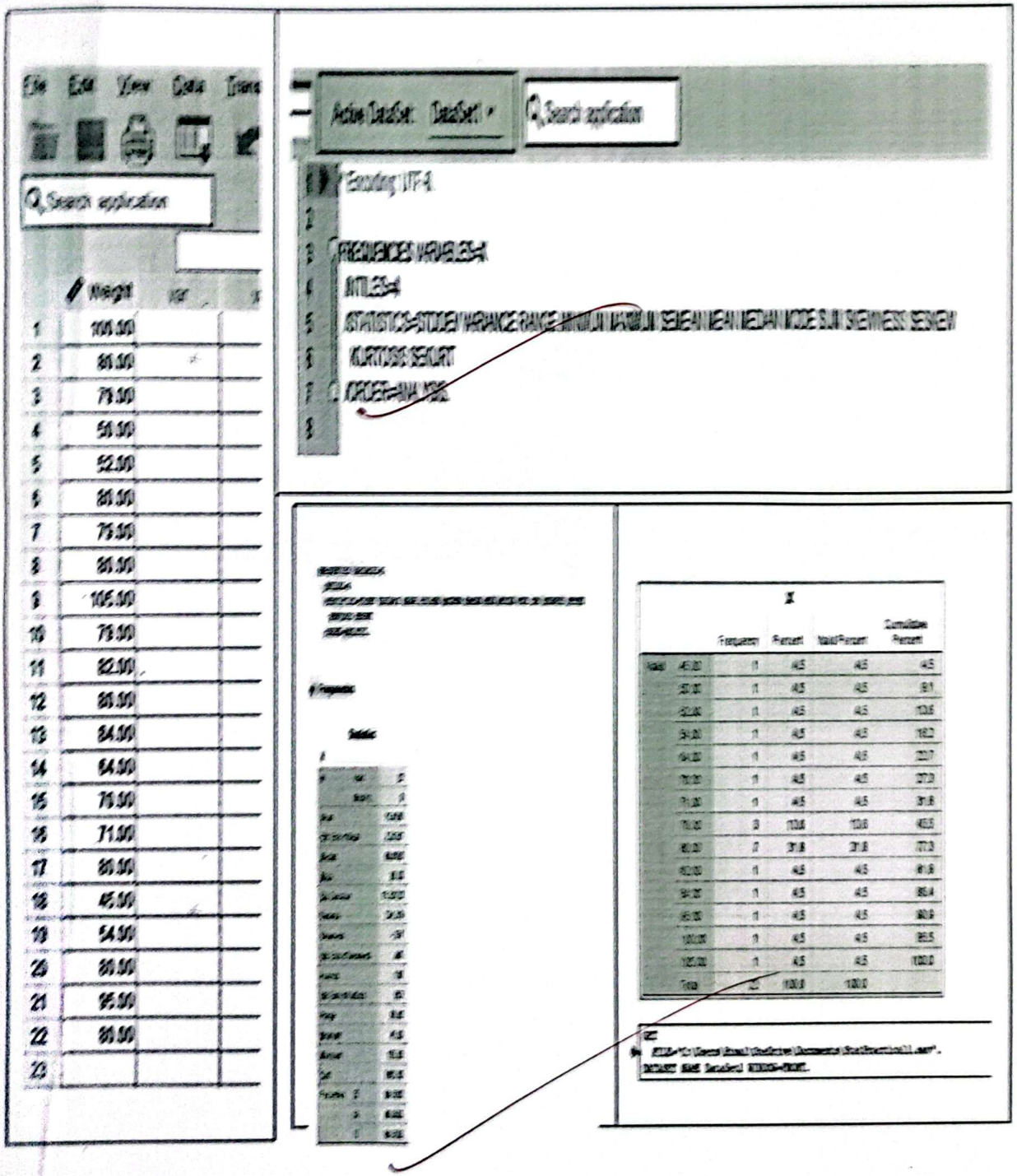
Coefficient of Range $= \frac{L-S}{L+S}$

Standard Deviation (S.D) $= \frac{\Sigma(x-\bar{x})^2}{n}$

Coefficient of variation $= \frac{\sigma}{\bar{x}} \times 100\%$

Results:-

Label:-



Results :-

$n = 22$

$S = 45$

$L = 105$

$\bar{X} = 75.8636$

$Md = 80$

$Mo = 80$

$Q_1 = 68.5$

$Q_3 = 80.5$

$R = 60$

Coefficient of Range $= \frac{105-45}{105+45} = \frac{60}{150} = \frac{2}{5}$

S.d(σ) $= 15.30420$

Variance(σ^2) $= 234.219$

Coefficient of variation $= \left(\frac{\sigma}{\bar{X}}\right) \times 100\% = \frac{15.30420}{75.8636} \times 100 = 20.17\%$

Ans

08/1/12

Patan Multiple Campus

Practical No. 2: Individual Series & Consistency

Level:- TU BCA 3rd Semester

Date:- 2081/01/11

Subject:- Probability and Statistics (CACS-202)

Roll No:- 13/080

Name: Bimal Acharya

Question:-

Find no. of observations, minimum value, maximum value, mean, median, mode, quartiles, range, standard deviation, variance and coefficient of variation of the following observations:

Height	Weight	Age
3.00	12.00	2.00
3.50	15.00	5.00
5.80	50.00	15.00
4.20	55.00	20.00
5.00	60.00	30.00
5.10	59.00	25.00
5.00	55.00	25.00
5.10	48.00	30.00
4.20	55.00	25.00
5.00	50.00	40.00
5.40	54.00	45.00

Working Expression:-

n = No. of observation

S = minimum value

L = maximum value

mean $(\bar{x}) = \frac{\Sigma x}{n}$

median $(M_d) = \left(\frac{n+1}{2}\right)^{th}$ item value

Mode (M_o) = maximum repeated number

First Quartile $(Q_1) = \left(\frac{n+1}{4}\right)^{th}$ item value.

Third Quartile $(Q_3) = 3\left(\frac{n+1}{4}\right)^{th}$ item value.

Range = $L - S$

Standard Deviation $(\sigma) = \frac{\Sigma(x-\bar{x})}{n}$

Coefficient of variation $= \frac{\sigma}{\bar{x}} \times 100\%$

LAB2:-

	Weight	Age	Height	var
1	100.00	25.00	4.11	
2	80.00	31.00	6.00	
3	79.00	32.00	6.11	
4	50.00	56.00	5.00	
5	52.00	45.00	4.10	
6	80.00	11.00	6.10	
7	79.00	56.00	4.50	
8	80.00	98.00	5.00	
9	105.00	65.00	5.11	
10	79.00	45.00	5.80	
11	82.00	32.00	6.00	
12	80.00	.	5.50	
13	84.00	.	5.60	
14	64.00	.	4.90	
15	70.00	.	.	
16	71.00	.	.	
17	80.00	.	.	
18	45.00	.	.	
19	54.00	.	.	
20	80.00	.	.	
21	95.00	.	.	
22	80.00	.	.	
23				

```
1  Encoding: UTF-8.
2
3  DATASET ACTIVATE DataSet2.
4  FREQUENCIES VARIABLES=Weight Age Height
5    /NTILES=4
6    /STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM SEMEAN MEAN MEDIAN MODE SUM SKEWNESS SESKEW
7     KURTOSIS SEKURT
8    /ORDER=ANALYSIS.
9
```

Double-click to activate

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	11.00	1	4.5	9.1	9.1
	25.00	1	4.5	9.1	18.2
	31.00	1	4.5	9.1	27.3
	32.00	2	9.1	18.2	45.5
	45.00	2	9.1	18.2	63.6
	56.00	2	9.1	18.2	81.8
	65.00	1	4.5	9.1	90.9
	98.00	1	4.5	9.1	100.0
	Total	11	50.0	100.0	
Missing	System	11	50.0		
Total		22	100.0		

Height

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4.10	1	4.5	7.1	7.1
	4.11	1	4.5	7.1	14.3
	4.50	1	4.5	7.1	21.4
	4.90	1	4.5	7.1	28.6
	5.00	2	9.1	14.3	42.9
	5.11	1	4.5	7.1	50.0
	5.50	1	4.5	7.1	57.1
	5.60	1	4.5	7.1	64.3
	5.80	1	4.5	7.1	71.4
	6.00	2	9.1	14.3	85.7
	6.10	1	4.5	7.1	92.9
	6.11	1	4.5	7.1	100.0
	Total	14	63.6	100.0	
Missing	System	8	36.4		
Total		22	100.0		

```
DATASET ACTIVATE DataSet2.
FREQUENCIES VARIABLES=Weight Age Height
  /STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM SEMEAN MEAN MEDIAN MODE SUM SKEWNESS SESKEW
    KURTOSIS SEKURT
  /ORDER=ANALYSIS.
```

➧ **Frequencies**

Statistics

		Weight	Age	Height
N	Valid	22	11	11
	Missing	0	11	0
Mean		75.9436	45.0909	5.2736
Std. Error of Mean		3.26767	7.09732	.18625
Median		80.0000	46.0000	5.3050
Mode		80.00	32.00[a]	5.04[a]
Std. Deviation		15.90420	23.53614	.78611
Variance		234.319	554.091	.581
Skewness		-.374	.971	-.415
Std. Error of Skewness		.491	.661	.597
Kurtosis		.136	1.586	-1.055
Std. Error of Kurtosis		.953	1.279	1.154
Range		60.00	87.00	2.01
Minimum		45.00	11.00	4.10
Maximum		105.00	98.00	6.11
Sum		1669.00	496.00	73.63
Percentiles	25	68.5000	31.0000	4.8050
	50	80.0000	46.0000	5.3050
	75	90.5000	58.0000	6.0000

a. Multiple modes exist. The smallest value is shown

Frequency Table

Weight

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	45.00	1	4.5	4.5	4.5
	50.00	1	4.5	4.5	9.1
	52.00	1	4.5	4.5	13.6
	54.00	1	4.5	4.5	18.2
	64.00	1	4.5	4.5	22.7
	70.00	1	4.5	4.5	27.3
	71.00	1	4.5	4.5	31.8
	79.00	3	13.6	13.6	45.5
	80.00	7	31.8	31.8	77.3
	82.00	1	4.5	4.5	81.8
	84.00	1	4.5	4.5	86.4
	95.00	1	4.5	4.5	90.9
	100.00	1	4.5	4.5	95.5
	105.00	1	4.5	4.5	100.0
	Total	22	100.0	100.0	

Results:-
n=11

	Height	Weight	Age
N	11	11	11
S	3	11	2
L	5.8	60	45
$\bar{x}$	4.6636	46.6364	23.3182
Md	5	54	25
Mo	5	55	25
Q_1	4.2	48	15
Q_3	5.1	55	30
R	2.8	4.8	43
Coefficient of Range	0.318	0.67	0.914
σ	0.84294	16.79448	13.05359
σ^2	0.711	282.055	171.364
Coefficient of Variation	18.074%	36.011%	54.96%

Since C.V of height (C.V.H) is less the data of height is more consistent.

082/11/14

Patan Multiple Campus

Practical No. 3: Descriptive Statistics

Level:- TU BCA 3rd Semester

Date:- 2081/01/11

Subject:- Probability and Statistics (CACS-202)

Roll No:- 13/080

Name: Bimal Acharya

Question:-

Find no. of observations, minimum value, maximum value, mean, median, mode, quartiles, range, standard deviation, variance and coefficient of variation of the following observations:

Marks	10	20	30	40	50	60	70	80	90	100
No.of students	2	3	5	15	20	15	4	2	1	0

Working Expression:-

$n = \Sigma f$

S = Smallest value

L = Largest value

mean $(\bar{x}) = \frac{\Sigma fx}{N}$

median $(Md) = \left(\frac{N+1}{2}\right)^{th}$ item value

mode (m_0) = maximum frequency value

first Quartile $(Q_1) = \left(\frac{N+1}{4}\right)^{th}$ item value

Third Quartile $(Q_3) = 3\left(\frac{N+1}{4}\right)^{th}$ item value

Range = L - S

Standard Deviation $(\sigma) = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{N}}$

Variance $= \frac{\Sigma f(x-\bar{x})^2}{N}$

Coefficient of Variation $= \frac{\sigma}{\bar{x}} \times 100\%$

Lab3:-

	X	f	var
1	5.00	1.00	
2	10.00	2.00	
3	15.00	3.00	
4	20.00	5.00	
5	25.00	6.00	
6	30.00	9.00	
7	35.00	20.00	
8	40.00	15.00	
9	45.00	4.00	
10	50.00	5.00	
11	55.00	65.00	
12	60.00	4.00	
13	65.00	32.00	
14	70.00	15.00	
15	75.00	45.00	
16	80.00	20.00	
17	85.00	15.00	
18	90.00	12.00	
19	95.00	30.00	
20	100.00	9.00	
21			

```
* Encoding: UTF-8.

DATASET ACTIVATE DataSet6.
WEIGHT BY f.

FREQUENCIES VARIABLES=X
  /NTILES=4
  /STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM SEMEAN MEAN MEDIAN MODE SUM SKEWNESS SESKEW
    KURTOSIS SEKURT
  /ORDER=ANALYSIS.
```

```
DATASET ACTIVATE DataSet4.
WEIGHT BY f.
FREQUENCIES VARIABLES=x
  /NTILES=4
  /STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM SEMEAN MEAN MEDIAN MODE SUM SKEWNESS SESKEW
    KURTOSIS SEKURT
  /ORDER=ANALYSIS.
```

Frequencies

Statistics

x

N	Valid	317
	Missing	0
Mean		64.1798
Std. Error of Mean		1.20431
Median		65.0000
Mode		55.00
Std. Deviation		21.44221
Variance		459.768
Skewness		-.343
Std. Error of Skewness		.137
Kurtosis		-.507
Std. Error of Kurtosis		.273
Range		95.00
Minimum		5.00
Maximum		100.00
Sum		20345.00
Percentiles	25	55.0000
	50	65.0000
	75	80.0000

x

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5.00	1	.3	.3	.3
	10.00	2	.6	.6	.9
	15.00	3	.9	.9	1.9
	20.00	5	1.6	1.6	3.5
	25.00	6	1.9	1.9	5.4
	30.00	9	2.8	2.8	8.2
	35.00	20	6.3	6.3	14.5
	40.00	15	4.7	4.7	19.2
	45.00	4	1.3	1.3	20.5
	50.00	5	1.6	1.6	22.1
	55.00	65	20.5	20.5	42.6
	60.00	4	1.3	1.3	43.8
	65.00	32	10.1	10.1	53.9
	70.00	15	4.7	4.7	58.7
	75.00	45	14.2	14.2	72.9
	80.00	20	6.3	6.3	79.2
	85.00	15	4.7	4.7	83.9
	90.00	12	3.8	3.8	87.7
	95.00	30	9.5	9.5	97.2
	100.00	9	2.8	2.8	100.0
	Total	317	100.0	100.0	

Result:-

$n = 10$

$S = 10$

$L = 100$

$\bar{x} = 55$

$M_d = 55$

$M_o = 10$

$Q_1 = 27.5$

$Q_3 = 82.5$

Range $= 90$

Coefficient of Range $= \frac{100-10}{100+10} = \frac{90}{110} = 0.818$

S.d $(\sigma) = 30.27655$

Variance $(\sigma^2) = 916.667$

Coefficient of Variation $= \frac{\sigma}{\bar{x}} \times 100\%$

$= \frac{30.2765}{55} \times 100\%$

$= 55.04\%$

Ans
082/1114

Patan Multiple Campus

Practical No. 4: Correlation and Regression

Level:- TU BCA 3rd Semester

Date:- 2081/01/16

Subject:- Probability and Statistics (CACS-202)

Roll No:- 13/080

Name: Bimal Acharya

Question:-

The students got the following percentage of marks in statistics and mathematics:

Statistics	78	36	98	25	75	82	90	62	65	39	80
Mathematics	84	51	91	60	68	62	86	58	63	25	47

i) Find Karl-Pearson's correlation coefficient between marks in statistics and mathematics and interpret it.
ii) Find probable error and test for significance.
iii) Find units of population correlation coefficient.
iv) Find coefficient of determination and interpret it.

Working Expression:-

i) Karl-Pearson's correlation coefficient (r) is :-

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

Interpretation of r :-

a) If $r = 0$, there is no correlation between x & y.
b) If $r > 0$, there is positive correlation between x and y.
c) If $r < 0$, there is negative correlation between x and y.

ii) Probable Error (PE) $= \frac{0.6745(1-r^2)}{\sqrt{n}}$

For significance test,

a) If $|r| < PE$, it is not significant
b) If $|r| > 6PE$, it is significant
c) Other than this, nothing is there to give conclusion to this correlation.

Ans. Write valid XML and CSS to display following details.

iii) Limits of population correlation coefficient (ρ) is :-
$r \pm PE$

$\therefore$ Lower limit of $\rho = r - PE = 0.670 - 0.112 = 0.558$

$\therefore$ Upper limit of $\rho = r + PE = 0.670 + 0.112 = 0.782$

iv) Coefficient of determination $= r^2 = (0.670)^2 = 0.4489$

Interpretation of r^2

If $r^2 \times 100\%$ of variation is explained by total variation.

Since, $r^2 \times 100\% = 0.4489 \times 100\% = 44.89\%$ of variation explained by total variation, it means only 44.89% of relation exists between statistics and mathematics.

Lab4:-

stat	math
78.00	84.00
36.00	51.00
98.00	91.00
25.00	60.00
75.00	68.00
82.00	62.00
90.00	86.00
62.00	58.00
65.00	63.00
39.00	25.00
80.00	47.00

```
1  * Encoding: UTF-8.
2
3  DATASET ACTIVATE DataSet0.
4
5  CORRELATIONS
6    /VARIABLES=stat math
7    /PRINT=TWOTAIL NOSIG
8    /STATISTICS DESCRIPTIVES
9    /MISSING=PAIRWISE.
10 NONPAR CORR
11   /VARIABLES=stat math
12   /PRINT=SPEARMAN TWOTAIL NOSIG
13   /MISSING=PAIRWISE.
14
```

```
CORRELATIONS
  /VARIABLES=stat math
  /PRINT=TWOTAIL NOSIG
  /STATISTICS DESCRIPTIVES
  /MISSING=PAIRWISE.
```

Correlations

Descriptive Statistics

	Mean	Std. Deviation	N
stat	66.3636	23.66970	11
math	63.1818	19.12495	11

Correlations

		stat	math
stat	Pearson Correlation	1	.670*
	Sig. (2-tailed)		.024
	N	11	11
math	Pearson Correlation	.670*	1
	Sig. (2-tailed)	.024	
	N	11	11

*. Correlation is significant at the 0.05 level (2-tailed).

```
NONPAR CORR
  /VARIABLES=stat math
  /PRINT=SPEARMAN TWOTAIL NOSIG
  /MISSING=PAIRWISE.
```

Nonparametric Correlations

Correlations

			stat	math
Spearman's rho	stat	Correlation Coefficient	1.000	.645*
		Sig. (2-tailed)	.	.032
		N	11	11
	math	Correlation Coefficient	.645*	1.000
		Sig. (2-tailed)	.032	.
		N	11	11

*. Correlation is significant at the 0.05 level (2-tailed).

```
SAVE OUTFILE='C:\Users\Hp\Documents\Statpractical4.sav'
  .
```

Result:-

$r = 0.670$, $n = 11$, $S^{\wedge}$ (upper limit) $= 0.558$, S (lower limit) $= 0.782$, $r^2 = 0.4489$

Interpretation of r

Since, $r = 0.670 > 0$, positive correlation exists between Statistics and Mathematics.

$$PE = \frac{0.6745(1-0.670^2)}{\sqrt{11}} = 0.122$$

$6 \times PE = 6 \times 0.122 = 0.672$

Since $|r| < 6PE$, there is no significant relationship between Statistics & Mathematics.

082/11/17

Patan Multiple Campus

Practical No. 5: Correlation and Regression

Level:- TU BCA 3rd Semester

Date:- 2081/01/16

Subject: Probability and Statistics (CACS-202)

Roll No:- 13/080

Name: Bimal Acharya

Question:-

In a study between the amount of daily rainfall(X) and the quantity of air pollution removed(Y), following data were collected:

X	4.3	4.5	5.9	5.6	6.1	5.2	3.8	2.1	2	1.5
Y	12.6	12.1	11.6	11.8	11.4	11.8	11.2	13.2	15.1	18.2

i) Estimate the regression equation of pollution removed on daily rainfall.
ii) Estimate the population removed when rainfall is 6.2.
iii) Interpret the regression coefficient.
iv) Find TSS, SSR and SSE.
v) Find coefficient of determination and interpret it.
vi) Find Standard error of estimate.

Working Expression:

Let rainfall (x) be independent variable and pollution(y) be dependent variable.

Then,

i) Let regression model of pollution (Y) on rainfall (x) be

$$y = a + bx$$

where,

$$b = b_{yx} = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \text{Regression Coefficient}$$

$$a = \bar{y} - b\bar{x} = \text{General mean effect}$$

$$\bar{x} = \frac{\Sigma x}{n}, \quad \bar{y} = \frac{\Sigma y}{n}$$

ii) The estimation of pollution removed ($\hat{y}$) for a specific value of rainfall ($x = 6.2$) is obtained by substituting the values into regression equation.

$$\hat{y} = a + bx.$$

iii) Interpretation of b:

a) If $b = 0$, there is no any effects.

b) If $b > 0$, then dependent variable y is increased by $|b|$ times per unit change in x.

c) If $b < 0$, then dependent variable y is decreased by $|b|$ times per unit change in independent variable x.

iv) TSS = Total Sum of Square = $\Sigma(y - \bar{y})^2$ = Total variation

SSE = Sum of Square due to error = $\Sigma(y - \hat{y})^2$ = Unexplained variation.

SSR = Sum of Square due to regression = TSS − SSE = Explained variation

v) Coefficient of determination (R^2) = $\frac{SSR}{TSS}$

Interpretation of R^2

If $R^2 \times 100\%$ of variation explained by total variation.

Standard Error (S_e) = $\sqrt{\frac{SSE}{n-k-1}}$

$= \sqrt{\frac{14.98}{10-1}}$

$= 1.29$

Lab5:-

X	Y	var
4.30	12.60	
4.50	12.10	
5.90	11.60	
5.60	11.80	
6.10	11.40	
5.20	11.80	
3.80	11.20	
2.10	13.20	
2.00	15.10	
1.50	18.20	

```
1   Encoding: UTF-8.
2
3   DATASET ACTIVATE DataSet0.
4   REGRESSION
5     /DESCRIPTIVES MEAN STDDEV CORR SIG N
6     /MISSING LISTWISE
7     /STATISTICS COEFF OUTS R ANOVA CHANGE
8     /CRITERIA=PIN(.05) POUT(.10)
9     /NOORIGIN
10    /DEPENDENT Y
11    /METHOD=ENTER X.
12
```

```
DATASET ACTIVATE DataSet0.
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT Y
  /METHOD=ENTER X.
```

➔ **Regression**

[DataSet0]

Descriptive Statistics

	Mean	Std. Deviation	N
Y	12.9000	2.18581	10
X	4.1000	1.70489	10

Correlations

		Y	X
Pearson Correlation	Y	1.000	-.807
	X	-.807	1.000
Sig. (1-tailed)	Y	.	.002
	X	.002	.
N	Y	10	10
	X	10	10

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	X[b]	.	Enter

a. Dependent Variable: Y

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.807[a]	.651	.608	1.36877	.651	14.951	1	8	.005

a. Predictors: (Constant), X

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	28.012	1	28.012	14.951	.005[b]
	Residual	14.988	8	1.874		
	Total	43.000	9			

a. Dependent Variable: Y

b. Predictors: (Constant), X

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17.143	1.180		14.534	.000
	X	-1.035	.268	-.807	-3.867	.005

a. Dependent Variable: Y

Lab 6

Results:-

i) $a = 17.143$

$b = 1.035$

∴ Required estimated regression equation of pollution on rainfall is :-

$$\hat{y} = 17.143 + (-1.035)x$$
$$= 17.143 - 1.035x$$

ii) When rainfall $(x) = 6.2$,

$$\hat{y} = 17.143 \quad 1.035 \times 6.2$$
$$= 10.726$$

iii) Since $b = -1.035 < 0$, then pollution is decreased by per unit change in rainfall.

iv) $TSS = 43$

$SSR = 28.012$

$SSE = 14.988$

v) Since $R^2 = 0.807^2 = 0.651 \times 100\% = 65.1\%$ of variation is explained by total variation, it means only 65.1% of relation exists between rainfall and pollution.

vi) Standard error $(se) = \sqrt{\frac{SSE}{n-k-1}}$

$$= \sqrt{\frac{14.988}{10-1}}$$

$$= 1.29$$

Eg 2 08/11/17

Patan Multiple Campus

Practical No. 6: Design of Experiments (one-way ANOVA)

Level:- TU BCA 3rd Semester

Date:- 2081/01/16

Subject: Probability and Statistics (CACS-202)

Roll No:- 13/080

Name: Bimal Acharya

Question:-

The following data show the number of claims processed per day for a group of insurance company employees observed for a number of days. Test the hypothesis is that the employee's mean claim per day are all the same. Use the **0.05** level of significance.

Employee1	15	17	14	12			
Employee2	12	10	13	17	18	15	
Employee3	11	14	13	15	12		
Employee4	13	12	12	14	10	9	8

Working Expression:

1) NULL Hypothesis (H_0): $\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$, Performance of all the k-Sample means are similar.

2) Alternative Hypothesis (H_1): $\mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_k$, Performance of at least one sample mean is different.

3) Test Statistics: Under H_1, test statistics is,

$$F = \frac{MST}{MSE}$$

where, $MST = \frac{SST}{k-1}$, $SST = \sum \frac{T_i^2}{n_i} - C.F$ = Sum of Square due to treatment.

$MSE = \frac{SSE}{N-k}$, $SSE = TSS - SST$ = Sum of Square due to error.

$TSS = \sum y_j^2 - C.F$ = Total sum of Square

$k-1$ = Degree of Freedom of treatment

$N-k$ = Degree of Freedom of error

$N-1$ = Degree of Freedom of total

4) Critical value = The tabulated value of F at α% level of significance and $(k-1, N-k)$ degree of freedom is $F_{\alpha,(k-1,N-k)}$

5) Decision = If $F_{cal} \leq F_{tab}$:- We accept H_1 otherwise reject H_0.

Lab6:-

Score	Emp	var
15.00	1.00	
17.00	1.00	
14.00	1.00	
12.00	1.00	
12.00	2.00	
10.00	2.00	
13.00	2.00	
17.00	2.00	
18.00	2.00	
15.00	2.00	
11.00	3.00	
14.00	3.00	
13.00	3.00	
15.00	3.00	
12.00	3.00	
13.00	4.00	
12.00	4.00	
12.00	4.00	
14.00	4.00	
10.00	4.00	
9.00	4.00	
8.00	4.00	

```
1  * Encoding: UTF-8.
2
3  DATASET ACTIVATE DataSet1.
4  ONEWAY Score BY Emp
5    /MISSING ANALYSIS.
6
```

```
DATASET ACTIVATE DataSet1.
ONEWAY Score BY Emp
  /MISSING ANALYSIS.
```

➧ Oneway

ANOVA

Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	41.310	3	13.770	2.511	.091
Within Groups	98.690	18	5.483		
Total	140.000	21			

Results :-

$k = 4$

$SST = 41.30$

$SSE = 98.690$

$TSS = 140.00$

$k-1 = 3$

$N-K = 18$

$N-1 = 21$

$MST = 13.770$

$MSE = 5.483$

$F_{cal} = \frac{MST}{MSE} = 2.511$

Now,

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$

Test statistics : Under H_0, test statistics is

$F = 2.511$

Critical value : The tabulated value of F at $\alpha = 0.05$ and $(k-1, N-k) = (3, 18)$ is $F_{\alpha,(k-1,N-k)}$

$= F_{0.05,(3,18)}$

Decision :- Since $F_{cal} < F_{0.05,(3,18)}$, we accept H_0. It is not significant.

08/11/19

Patan Multiple Campus

Practical No. 7: Design of Experiments (Two-way ANOVA)

Level:- TU BCA 3rd Semester

Date:- 2081/01/22

Subject: Probability and Statistics (CACS-202)

Roll No:- 13/080

Name: Bimal Acharya

Question:-

The following table gives the data on the performance of three different detergents at three different Water temperatures. The performance was obtained on the 'Whiteness' reading based on specially designed equipment for twelve loads of washing.

Water \ Detergents	A	B	C	D
Cold Water	45	43	55	60
Warm Water	37	40	56	62
Hot Water	42	44	46	55

Perform two-way ANOVA, using level of significance $\alpha = 0.05$

Working Expression:

Procedure of two-way ANOVA test

For water:

Null hypothesis (H_{0R}): $\mu_1 = \mu_2 = \mu_3$; performance of all the three types water is similar.

Alternative hypothesis (H_{1R}): $\mu_1 \neq \mu_2 \neq \mu_3$; performance of at least one type of water is different.

For detergent:

Null hypothesis (H_{0c}): $\mu_1 = \mu_2 = \mu_3 = \mu_4$; performance of all the detergent is similar.

Alternative hypothesis (H_{1c}): $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$; performance of at least one detergent is different.

Test statistics: Under H_0, test statistic is

(i) $F_R = \frac{MSR}{MSE}$

(ii) $F_C = \frac{MSC}{MSE}$

Critical value:

(i) The tabulated value of F at α% of level for row is $F_{\alpha,[t-1,(t-1)(r-1)]}$

(ii) The tabulated value of F at α% level of significance for column is $F_{\alpha,[(r-1),(t-1)(r-1)]}$

Decision:-

(i) If $F_R \leq F_{\alpha,[t-1,(t-1)(r-1)]}$, we accept H_{0R} otherwise reject H_{0R}

(ii) If $F_C \leq F_{\alpha}[(r-1),(t-1)(r-1)]$, we accept H_{0c} otherwise reject H_{0c}.

where,

t = No. of treatment = No. of water

r = No. of block = No. of detergent

$CF = \frac{G^2}{N}$ = Correction factor

$TSS = \sum\sum y_{ij}^2 - C.F$ = Total sum of square

$SSR = \frac{\sum T_R^2}{r} - C.F$ = Sum of square due to treatment

$SSC = \frac{\sum T_c^2}{r} - C.F$ = Sum of square due to block (detergent)

(obj)

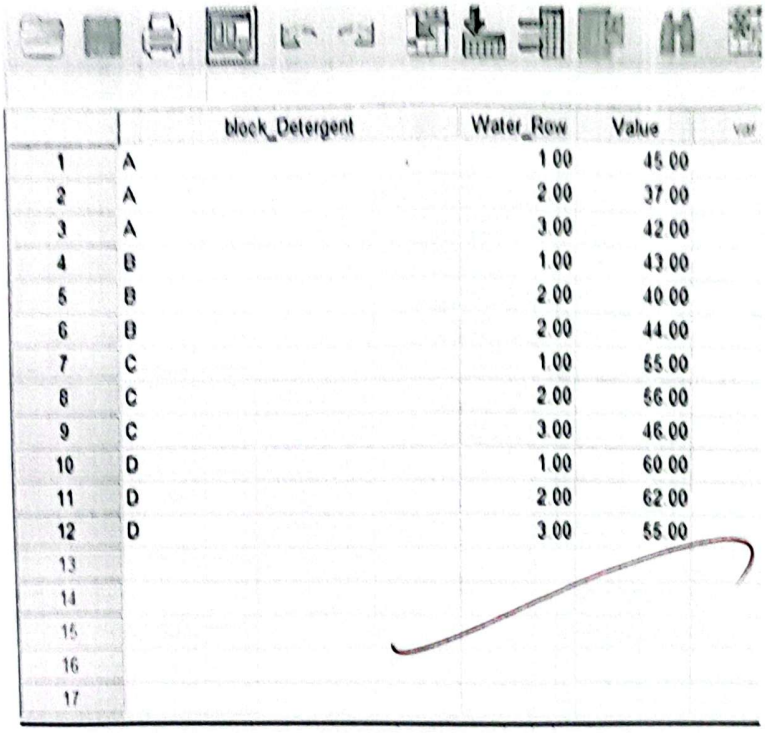
SSE = TSS − SSR − SSC = sum of square due to error

$MSR = \frac{SSR}{t-1}$, $t-1$ = Of. of treatment

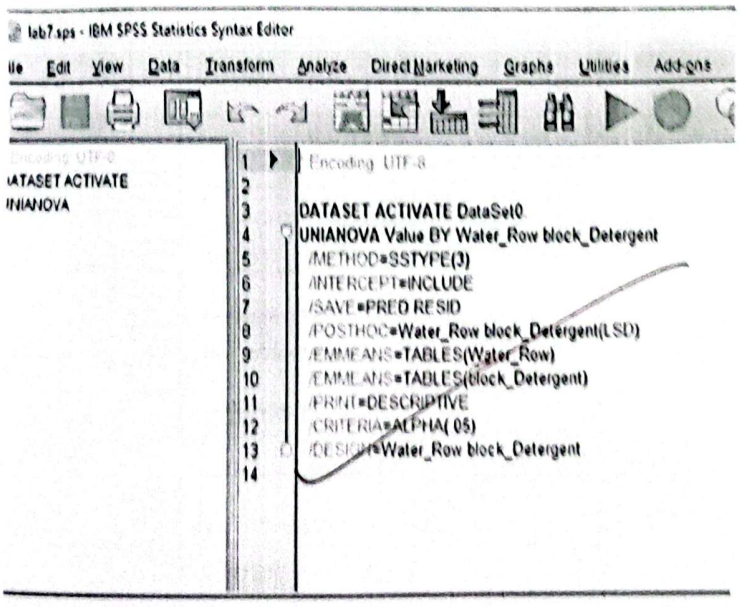
$MSC = \frac{SSC}{r-1}$, $r-1$ = Of. of block

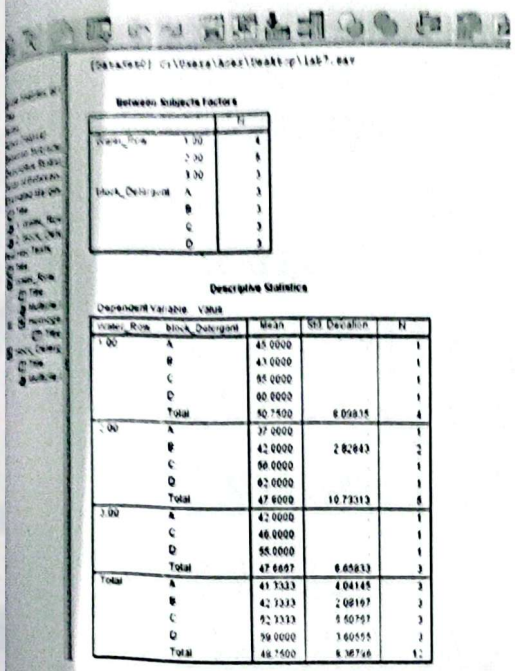
$MSE = \frac{SSE}{(t-1)(r-1)}$; $(t-1)(r-1)$ = Of. of error

Result of lab 7:



	block_Detergent	Water_Row	Value
1	A	1.00	45.00
2	A	2.00	37.00
3	A	3.00	42.00
4	B	1.00	43.00
5	B	2.00	40.00
6	B	2.00	44.00
7	C	1.00	55.00
8	C	2.00	56.00
9	C	3.00	46.00
10	D	1.00	60.00
11	D	2.00	62.00
12	D	3.00	55.00
13			
14			
15			
16			
17			





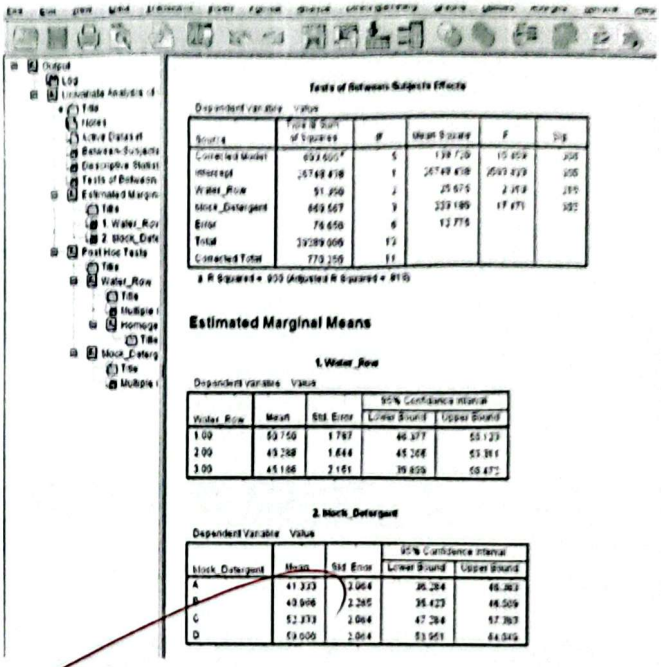
Between-Subjects Factors

		N
Water_Row	1.00	4
	2.00	5
	3.00	3
block_Detergent	A	3
	B	3
	C	3
	D	3

Descriptive Statistics

Dependent Variable: Value

Water_Row	block_Detergent	Mean	Std. Deviation	N
1.00	A	45.0000		1
	B	43.0000		1
	C	55.0000		1
	D	60.0000		1
	Total	50.7500	8.09835	4
2.00	A	37.0000		1
	B	42.0000	2.82843	2
	C	56.0000		1
	D	62.0000		1
	Total	47.6000	10.73313	5
3.00	A	42.0000		1
	C	46.0000		1
	D	55.0000		1
	Total	47.6667	6.65833	3
Total	A	41.3333	4.04145	3
	B	42.3333	2.08167	3
	C	52.3333	5.50757	3
	D	59.0000	3.60555	3
	Total	48.7500	8.36746	12



Tests of Between-Subjects Effects

Dependent Variable: Value

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	693.600[a]	5	138.720	10.859	.006
Intercept	26749.838	1	26749.838	2093.829	.000
Water_Row	91.350	2	45.675	3.576	.095
block_Detergent	669.567	3	223.189	17.471	.002
Error	76.656	6	12.776		
Total	39289.000	12			
Corrected Total	770.256	11			

a. R Squared = .900 (Adjusted R Squared = .817)

Estimated Marginal Means

1. Water_Row

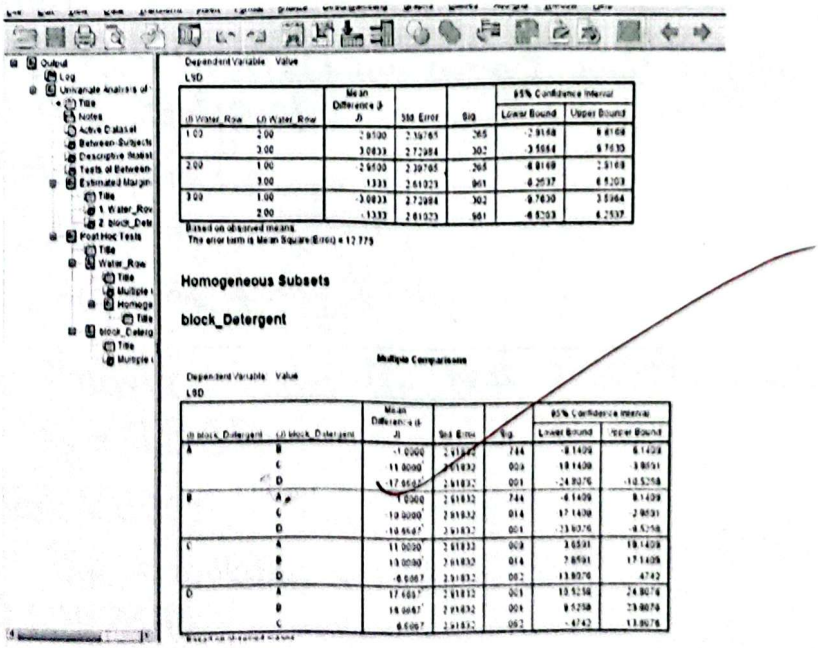
Dependent Variable: Value

Water_Row	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1.00	50.750	1.787	46.377	55.123
2.00	45.200	1.644	41.206	49.194
3.00	45.186	2.161	39.899	50.472

2. block_Detergent

Dependent Variable: Value

block_Detergent	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
A	41.333	2.064	36.284	46.383
B	43.000	2.065	37.923	48.077
C	52.333	2.064	47.284	57.383
D	59.000	2.064	53.951	64.049



Dependent Variable: Value

LSD

(I) Water_Row	(J) Water_Row	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	5.5500	2.59761	.065	-.3168	11.4168
	3.00	5.5833	2.72984	.082	-.5564	11.7231
2.00	1.00	-5.5500	2.59761	.065	-11.4168	.3168
	3.00	.0333	2.61923	.991	-5.9537	6.0203
3.00	1.00	-5.5833	2.72984	.082	-11.7231	.5564
	2.00	-.0333	2.61923	.991	-6.0203	5.9537

Based on observed means.
The error term is Mean Square(Error) = 12.776.

Homogeneous Subsets

block_Detergent

Multiple Comparisons

Dependent Variable: Value

LSD

(I) block_Detergent	(J) block_Detergent	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
A	B	-1.6667	2.61832	.744	-8.0703	4.7370
	C	-11.0000*	2.61832	.003	-17.4036	-4.5964
	D	-17.6667*	2.61832	.001	-24.0703	-11.2630
B	A	1.6667	2.61832	.744	-4.7370	8.0703
	C	-9.3333*	2.61832	.014	-15.7370	-2.9297
	D	-16.0000*	2.61832	.001	-22.4036	-9.5964
C	A	11.0000*	2.61832	.003	4.5964	17.4036
	B	9.3333*	2.61832	.014	2.9297	15.7370
	D	-6.6667	2.61832	.062	-13.0703	.7370
D	A	17.6667*	2.61832	.001	11.2630	24.0703
	B	16.0000*	2.61832	.001	9.5964	22.4036
	C	6.6667	2.61832	.062	-.7370	13.0703

lab7

Result:-

$t=3$, $r=4$

$N = tr = 3\times4 = 12$

SSR = 32

SSC = 642.250

SSE = 96

TSS = 770.250

Now,

for water

$H_{0R}: \mu_1 = \mu_2 = \mu_3$

$H_{1R}: \mu_1 \neq \mu_2 \neq \mu_3$

Test statistic: Under H_0, test statistics is: $F_R = 1$

Critical value:

The tabulated value of F at $\alpha = 0.05$ is $F_{0.05}(2,6) =$ 5.1433

Decision:

Since $F_R < F_{0.05}(3,6)$, we accept null hypothesis and it is not significant.

for Deteregent:

$H_{0C}: \mu_1 = \mu_2 = \mu_3 = \mu_4$

$H_{1C}: \mu_1 \neq \mu_1 \neq \mu_3 \neq \mu_4$

Test statistic: Under H_0, test statistic is:

$F_C = 13.380$

Critical value:

The tabulated value of F at $\alpha = 0.05$ is $F_{0.05}$ (3,6) = 4.7571

Decision

Since $F_C > F_{0.05}(3,6)$, we regret null hypothesis.

PATAN MULTIPLE CAMPUS

Practical No. 8: Binomial Distribution

Level: **BCA 3rd Semester** Date: **2082-01-25**

Subject: **Probability and Statistics (STA-202)**

Roll no: **13/80**

Name: **Bimal Acharya**

Question:

Fit binomial distribution of the following Observation.

X	0	1	2	3	4	5	6	7	8
F	2	8	75	250	210	70	50	20	5

Also, Find

i) P(X <= 2)
ii) P(X >= 6)
iii) P(4 <= X <= 6)

Working Expression:

Working Expression :-

The probability mass function of x is

$$P(x=x) = B(x,n,p) = \begin{cases} {}^{n}C_{x}\, p^{x} q^{n-x} & , x = 0,1,2,\ldots n. \\ 0 & , \text{otherwise} \end{cases}$$

where,

n = parameter = No. of trials

$p = \frac{\bar{x}}{n}$ = parameter = probability of succes events in a single trial

$q = 1-p$

$\bar{x} = \frac{\Sigma f x}{N}$, $N = \Sigma f$

Expected frequency $(Ef) = N \cdot p(x=x)$

Result of lab8:-

	x	f	p	pdf	EF	REF
1	.00	2.00	.4670	.0065	4.49	4.00
2	1.00	8.00	.4670	.0456	31.49	31.00
3	2.00	75.00	.4670	.1400	96.59	97.00
4	3.00	250.00	.4670	.2453	169.27	169.00
5	4.00	210.00	.4670	.2687	185.41	185.00
6	5.00	70.00	.4670	.1884	129.97	130.00
7	6.00	50.00	.4670	.0825	56.95	57.00
8	7.00	20.00	.4670	.0207	14.26	14.00
9	8.00	5.00	.4670	.0023	1.56	2.00
10						
11						

Encoding: UTF-8

```
DATASET ACTIVATE DataSet0.
WEIGHT BY f.

FREQUENCIES VARIABLES=x
  /STATISTICS=MEAN
  /ORDER=ANALYSIS.

COMPUTE p=3.7362 / 8.
EXECUTE.

COMPUTE pdf=PDF.BINOM(x,8,p).
EXECUTE.

COMPUTE EF=690 * pdf.
EXECUTE.

COMPUTE REF=RND(EF).
EXECUTE.
```

Result of lab8:-

```
WEIGHT BY f.
FREQUENCIES VARIABLES=x
  /STATISTICS=MEAN
  /ORDER=ANALYSIS.
```

Frequencies

[DataSet0]

Statistics

x

N	Valid	690
	Missing	0
Mean		3.7362

x

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	2	.3	.3	.3
	1.00	8	1.2	1.2	1.4
	2.00	75	10.9	10.9	12.3
	3.00	250	36.2	36.2	48.6
	4.00	210	30.4	30.4	79.0
	5.00	70	10.1	10.1	89.1
	6.00	50	7.2	7.2	96.4
	7.00	20	2.9	2.9	99.3
	8.00	5	.7	.7	100.0
	Total	690	100.0	100.0	

```
COMPUTE p=3.7362 / 8.
EXECUTE.
COMPUTE pdf=PDF.BINOM(x,8,p).
EXECUTE.
COMPUTE EF=690 * pdf.
EXECUTE.
COMPUTE REF=RND(EF).
EXECUTE.

SAVE OUTFILE='E:\parctical8.sav'
  /COMPRESSED.
```

Lab6

Result

$\bar{x} = 3.7362$

$p = 0.4670$

$N = 690$

x	f	p	pdf	Ef	REf
0	2.00	.467	0.0065	4.49	4.00
1	8.00	.467	0.0456	31.49	31.00
2	75.00	.467	0.1400	96.59	97.00
3	250.00	.467	0.2453	169.27	169.00
4	210.00	.467	0.2687	185.41	185.00
5	70.00	.467	0.1884	129.07	130.00
6	50.00	.467	0.825	56.95	57.00
7	20.00	.467	0.0207	14.26	14.00
8	5.00	.467	0.0023	1.56	2.00

1) $P(x \leq 2)$

$= P(x=0) + P(x=1) + P(x=2)$

$= 0.0065 + 0.0456 + 0.1400$

$= 0.19$

2) $P(x \geq 6)$

$= P(x=6) + P(x=7) + P(x=8)$

$= 0.825 + 0.00207 + 0.0023$

$= 0.11$

3) $P(4 \leq x \leq 6)$

$= P(x=4) + P(x=5) + P(x=6)$

$= 0.2453 + 0.2687 + 0.1884$

$= 0.70$

082/1126

PATAN MULTIPLE CAMPUS

Practical No. 9: Poisson Distribution

Level: **BCA 3rd Semester** Date: **2082-01-30**

Subject: **Probability and Statistics (STA-202)**

Roll no: **13/80**

Name: **Bimal Acharya**

Question:

Fit Poisson distribution of the following Observation.

X	0	1	2	3	4	5	6
F	211	250	70	21	7	2	1

Also, Find

i) P(X <= 2)
ii) P(X >= 4)
iii) P(X >4)
iv) P(2 < X <= 4)

Working Expression:

Working Expression:-

1) The probability mass function of poisson distribution is

$$P(x=x) = P(x,\lambda) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & x = 0,1,2,\ldots 6 \\ 0, & \text{otherwise} \end{cases}$$

when, λ = parameter = mean ($\bar{x}$) = variance (σ^2) = $\frac{\Sigma fx}{N}$

2) Expected frequency (Ef) = $Np(x=x)$

where,

$N = \Sigma f$

$P(x=x)$ = Pmf (Pdf)

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

4) $P(x \geq 4) = P(x=4) + P(x=5) + P(x=6)$

5) $P(x > 4) = P(x=5) + P(x=6)$

6) $P(2 < x \leq 4) = P(x=3) + P(x=4)$

sult of lab9:-

x	f	Pdf	EF	REF	var
.00	211.00	.413003	232.1078	232.00	
1.00	250.00	.365219	205.2529	205.00	
2.00	70.00	.161481	90.7526	91.00	
3.00	21.00	.047599	26.7508	27.00	
4.00	7.00	.010523	5.9139	6.00	
5.00	2.00	.001861	1.0459	1.00	
6.00	1.00	.000274	.1542	.00	

```
1   ▶  Encoding: UTF-8
2
3       DATASET ACTIVATE DataSet0.
4       WEIGHT BY f.
5
6       FREQUENCIES VARIABLES=x
7         /STATISTICS=MEAN
8         /ORDER=ANALYSIS.
9
10      COMPUTE Pdf=PDF.POISSON(x,0.8843)
11      EXECUTE.
12
13      COMPUTE EF=562 * Pdf.
14      EXECUTE.
15
16      COMPUTE REF=RND(EF).
17      EXECUTE.
18
```

esult of lab9:-

```
DATASET ACTIVATE DataSet0.
WEIGHT BY f.
FREQUENCIES VARIABLES=x
  /STATISTICS=MEAN
  /ORDER=ANALYSIS.
```

Frequencies

[DataSet0]

Statistics

x

N	Valid	562
	Missing	0
Mean		8843

x

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	211	37.5	37.5	37.5
	1.00	250	44.5	44.5	82.0
	2.00	70	12.5	12.5	94.5
	3.00	21	3.7	3.7	98.2
	4.00	7	1.2	1.2	99.5
	5.00	2	.4	.4	99.8
	6.00	1	.2	.2	100.0
	Total	562	100.0	100.0	

```
COMPUTE Pdf=PDF.POISSON(x,0.8843).
EXECUTE.
COMPUTE EF=562 * Pdf.
EXECUTE.
COMPUTE REF=RND(EF).
EXECUTE.

SAVE OUTFILE='E:\New folder\9\pre
  /COMPRESSED.
```

Double-click to activate

Lab 9

Result

$N = 562$

$\lambda = \bar{x} = 0.884$

x	f	Pdf	Ef	REf
0.00	217.00	.413003	232.207	252.00
1.00	250.00	.365219	205.52	206.00
2.00	70.00	.161481	90.75	91.00
3.00	27.00	.047599	26.75	27.00
4.00	7.00	.010523	5.91	6.00
5.00	2.00	.001861	1.04	1.00
6.00	1.00	.000274	0.15	0.00

Also,

i) $P(x \leq 2) = 0.413003 + 0.365219 + 0.161481$
$= 0.94$

ii) $P(x \geq 4) = 0.010523 + 0.001861 + 0.000274$
$= 0.127$

iii) $P(x > 4) = 0.001861 + 0.000274$
$= 0.0021$

iv) $P(2 < x \leq 4) = 0.047599 + 0.010523$
$= 0.0581$

Eg 12
082/11/31

Patan Multiple Campus

Practical No. 10: Normal Distribution

Level:- TU BCA 3rd Semester

Date:- 2081/01/31

Subject: Probability and Statistics (CACS-202)

Roll No:- 13/080

Name: Bimal Acharya

Question:-

Fit normal distribution of the following observations:-

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No.of students	5	20	10	100	150	125	50	20	5	2

Also find:-

1. $P(x \leq 30)$
2. $P(x \geq 50)$
3. $P(20 < x < 30)$
4. $P(35 < x < 65)$

Working Expression:

Result :-

$N = 487$

$\mu = 47.3614$

$\sigma = 14.40053$

$\sigma^2 = 207.375$

Then,

Marks	No. of student (f)	$P(n_1 < n < n_2)$ pdf	Ef	REf
0-10	5	.004234	2.06	2.00
10-20	20	.023978	11.68	12.00
20-30	10	.085269	41.53	42.00
30-40	100	.190625	92.83	93.00
40-50	150	.268082	130.56	131.00
50-60	125	.237242	115.54	116.00
60-70	50	.132100	64.53	64.00
70-80	20	.046253	22.53	23.00
80-90	5	.016777	4.96	5.00
90-100	2	.001405	0.68	1.00

Also,

i) $P(n \leq 30) = 0.1140$

ii) $P(n \geq 50) = 0.4273$

iii) $P(20 < n < 30) = 0.0853$

iv) $P(35 < n < 65) = 0.6943$

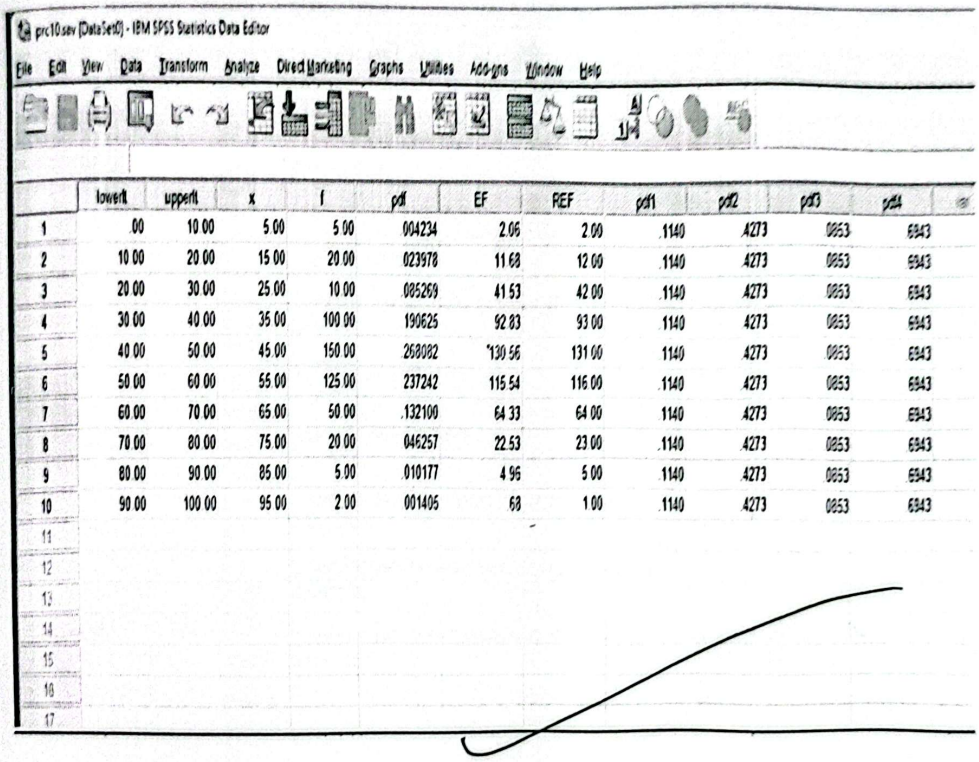
210>

DATA

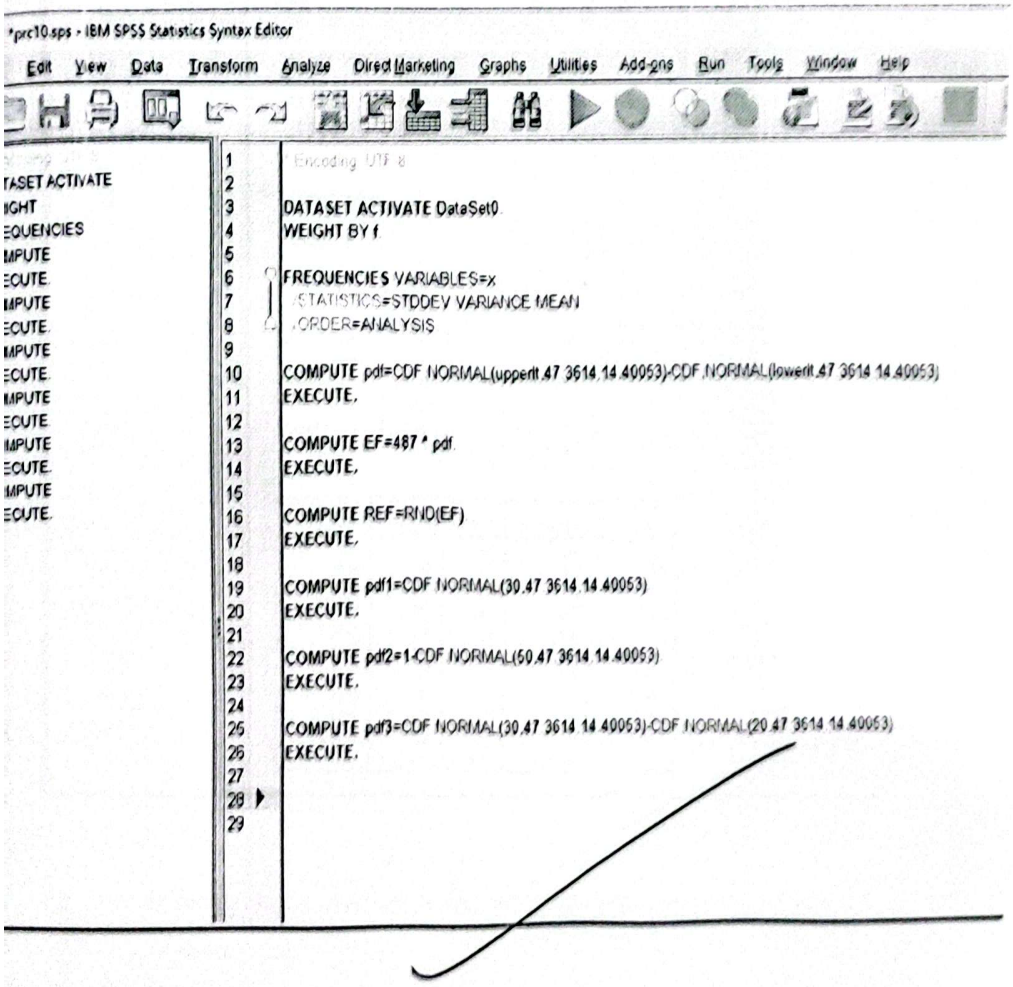
prc10.sav [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

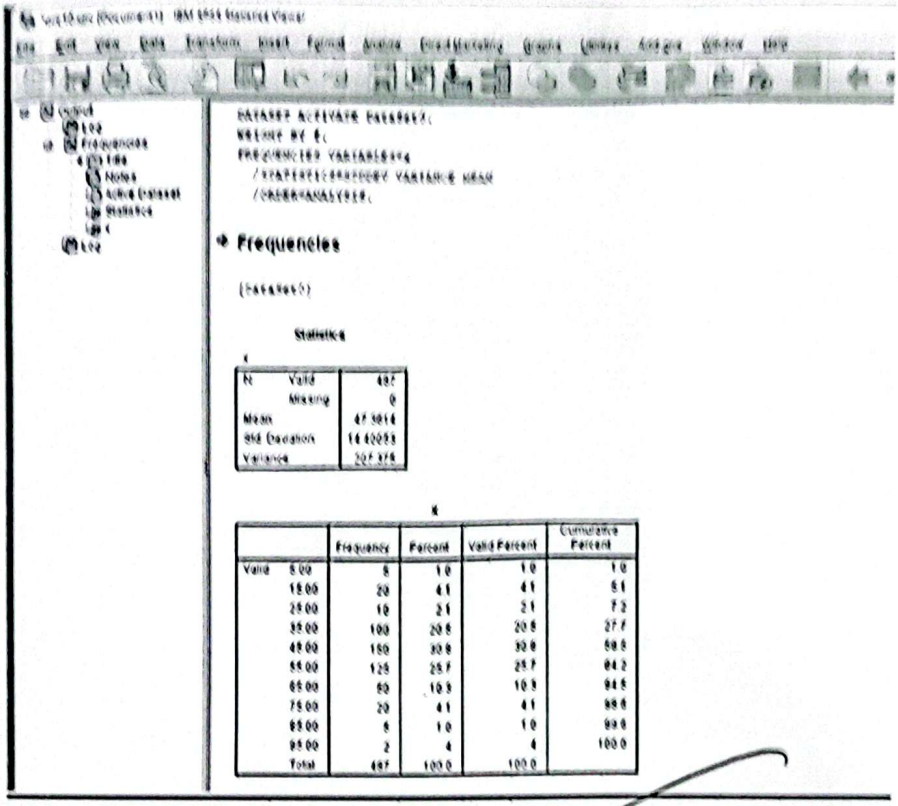
	lowerlt	upperlt	x	f	pdf	EF	REF	pdf1	pdf2	pdf3	pdf4
1	.00	10.00	5.00	5.00	004234	2.06	2.00	.1140	.4273	0853	6943
2	10.00	20.00	15.00	20.00	023978	11.68	12.00	.1140	.4273	0853	6943
3	20.00	30.00	25.00	10.00	085269	41.53	42.00	.1140	.4273	0853	6943
4	30.00	40.00	35.00	100.00	190625	92.83	93.00	.1140	.4273	0853	6943
5	40.00	50.00	45.00	150.00	268082	130.56	131.00	.1140	.4273	0853	6943
6	50.00	60.00	55.00	125.00	237242	115.54	116.00	.1140	.4273	0853	6943
7	60.00	70.00	65.00	50.00	.132100	64.33	64.00	.1140	.4273	0853	6943
8	70.00	80.00	75.00	20.00	046257	22.53	23.00	.1140	.4273	0853	6943
9	80.00	90.00	85.00	5.00	010177	4.96	5.00	.1140	.4273	0853	6943
10	90.00	100.00	95.00	2.00	001405	.68	1.00	.1140	.4273	0853	6943
11											
12											
13											
14											
15											
16											
17											



SYNTAX



OUTPUT



Frequencies

Statistics

N	Valid	487
	Missing	0
Mean		47.3614
Std. Deviation		14.40053
Variance		207.378

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5.00	5	1.0	1.0	1.0
	15.00	20	4.1	4.1	5.1
	25.00	10	2.1	2.1	7.2
	35.00	100	20.5	20.5	27.7
	45.00	150	30.8	30.8	58.5
	55.00	125	25.7	25.7	84.2
	65.00	50	10.3	10.3	94.5
	75.00	20	4.1	4.1	98.6
	85.00	5	1.0	1.0	99.6
	95.00	2	.4	.4	100.0
	Total	487	100.0	100.0	

```
COMPUTE pdf=CDF.NORMAL(upperlt,47.3614,14.40053)-CDF.NORMAL(lowerlt,47.3614,14.40053).
EXECUTE.
COMPUTE EF=487 * pdf.
EXECUTE.
COMPUTE REF=RND(EF).
EXECUTE.

SAVE OUTFILE='E:\New folder\10\pro10.sav'
  /COMPRESSED.
COMPUTE pdf1=CDF.NORMAL(30,47.3614,14.40053).
EXECUTE.
COMPUTE pdf2=1-CDF.NORMAL(50,47.3614,14.40053).
EXECUTE.
COMPUTE pdf3=CDF.NORMAL(30,47.3614,14.40053)-CDF.NORMAL(20,47.3614,14.40053).
EXECUTE.
COMPUTE pdf4=CDF.NORMAL(65,47.3614,14.40053)-CDF.NORMAL(35,47.3614,14.40053).
EXECUTE.
DATASET ACTIVATE DataSet0.

SAVE OUTFILE='E:\New folder\10\pro10.sav'
  /COMPRESSED.
```

working expression :-

1) The probability density function of normal distribution is

$$f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}, & -\infty < x < \infty \\ 0, & \text{otherwise} \end{cases}$$

where, μ = parameter : mean $(\bar{x}) = \frac{\Sigma fx}{N}$

σ^2 = parameter = variance $(\sigma^2) = \frac{\Sigma f(x-\bar{x})^2}{N}$

σ = standard deviation

2) Expected frequency (Ef) = $N \cdot P(x_1 \leq x \leq x_2)$

where,

x_1 = lower limit of the class

x_2 = upper limit of the class

Exp 2
08/12/16