THE NEOTIA UNIVERSITY

ज्ञानम् आत्म प्रदीपाय

STATISTICAL METHODS

Practical Manual

Course Code: CC-AGP 324,Credits, 2(1+1)

BSc (Agriculture)

Compiled by

Mr. Mostaid Ahmed

Department of Agricultural Economics,Statistics
School of Agriculture and Allied Sciences

The Neotia University
Jhinger Pole, Diamond Harbour Rd, Sarisha, West Bengal 743368

Problems:

1. Calculation of Arithmetic Mean for Continuous Frequency Distribution
2. Calculation of Median for Continuous Frequency Distribution
3. Calculation of Mode for Continuous Frequency Distribution
4. Calculation of Standard Deviation for Continuous Frequency Distribution
5. Calculation of Pearson's Correlation Coefficient
6. Calculation of Pearson's Correlation Coefficient
7. Calculation of Spearman's Rank Correlation Coefficient
8. Problem on Chi-Square Goodness of fit

1) Calculate the mean for the following frequency distribution.

Class interval :	0-8	8-16	16-24	24-32	32-40	40-48
Frequency :	8	7	16	24	15	7

=> Solution: Here, we take $A = 28$ and $h = 8$

Computation of Mean.

Class interval	Mid-Value (x)	Frequency (f)	$d = \frac{x-A}{h}$	fd
0-8	4	8	-3	-24
8-16	12	7	-2	-14
16-24	20	16	-1	-16
24-32	28	24	0	0
32-40	36	15	1	15
40-48	44	7	2	14
Total		77		-25

$$\bar{x} = A + \frac{h \Sigma fd}{N}$$

$$= 28 + \frac{8 \times (-25)}{77} = 28 - \frac{200}{77} = 25.404.$$

2) Find the median wage of the following distribution:-

Wages (in Rs) :	2000-3000	3000-4000	4000-5000	5000-6000	6000-7000
No. of Workers :	3	5	20	10	5

=> Solution:

Computation of median.

Wages (in Rs)	No. of employees	c.f.
2000 - 3000	3	3
3000 - 4000	5	8
4000 - 5000	20	28
5000 - 6000	10	38
6000 - 7000	5	43

Here, $\frac{1}{2} N = \frac{1}{2}(43) = 21.5$

Cumulative frequency just greater than 21.5 is 28 and the corresponding class is 4000 - 5000. Thus median class is 4000 - 5000

Hence using median, $= 4000 + \frac{1000}{20}(21.5 - 8) = 4000 + 675 = 4675$

Thus, the median wage is Rs. 4675.

3) Calculate mode for the following distribution:

Class interval :	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency :	5	8	7	12	28	20	10	10

=> Solution - Here, the maximum frequency is 28.
Thus, the class interval 40-50 is the modal class. Using mode formulae, the value of the mode is given by :-

$$\text{Mode} = 40 + \frac{10(28-12)}{(2\times 28 - 12 - 20)} = 40 + 6.666 = 46.67 \text{ (appx)}$$

4) Calculation of standard deviation for continuous frequency distribution :-

Age (in years) :	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No of members :	3	61	132	153	140	51	2

=> Solution :- Here, we take $d = \frac{x - A}{h} = \frac{x - 55}{10}$

Age group	Mid value (x)	Frequency (f)	$d = \frac{x-55}{10}$	fd	fd^2
20-30	25	3	-3	-9	27
30-40	35	61	-2	-122	244
40-50	45	132	-1	-132	132
50-60	55	153	0	0	0
60-70	65	140	1	140	140
70-80	75	51	2	102	204
80-90	85	2	3	6	18
Total		542		-15	765

$$\bar{x} = A + h\frac{\Sigma fd}{N}$$

$$= 55 + \frac{10 \times (-15)}{542} = 55 - 0.28 = 54.72 \text{ years.}$$

$$\sigma^2 = h^2\left[\frac{1}{N}\Sigma f d^2 - \left(\frac{1}{N}\Sigma f d\right)^2\right] = 100\left[\frac{765}{542} - (0.028)^2\right]$$

$$= 100 \times 1.4107$$

$$= 141.07.$$

$\therefore \sigma$ (Standard Deviation) $= \sqrt{141.07} = 11.88$ years.

5. Calculate the co-relation co-efficient for the following heights (in inches) for father's (x) and their sons (y):

x :	65	66	67	67	68	69	70	72
y :	67	68	65	68	72	72	69	71

⇒ Solution: Calculation for co-rrelation co-efficient.

	x	y	x^2	y^2	xy
	65	67	4225	4489	4355
	66	68	4356	4624	4488
	67	65	4489	4225	4355
	67	68	4489	4624	4556
	68	72	4624	5184	4896
	69	72	4761	5184	4968
	70	69	4900	4761	4830
	72	71	5184	5041	5112
Total	544	552	37028	38132	37560

$$\bar{x} = \frac{1}{n}\Sigma x = \frac{544}{8} = 68, \quad \bar{y} = \frac{1}{n}\Sigma y = \frac{1}{8} \times 552 = 69$$

$$r(x,y) = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{\frac{1}{n}\Sigma xy - \bar{x}\bar{y}}{\sqrt{\left(\frac{1}{n}\Sigma x^2 - \bar{x}^2\right)\left(\frac{1}{2}\Sigma y^2 - \bar{y}^2\right)}}$$

$$= \frac{\frac{1}{8} \times 37560 - 68 \times 69}{\sqrt{\left\{\frac{37028}{8} - (68)^2\right\}\left\{\frac{38132}{8} - (69)^2\right\}}}$$

$$= \frac{4695 - 4692}{\sqrt{(4628.5 - 4624)(4766.5 - 4761)}} = \frac{3}{\sqrt{4.5 \times 5.5}} = 0.603.$$

6) The ranks of same 16 students in Mathematics and Physics are as follows. Two numbers within brackets denote the ranks of the students in Mathematics and Physics.:

(1, 1), (2, 10) (3, 3) (4, 4) (5, 5) (6, 7) (7, 2) (8, 6) (9, 8)
(10, 11) (11, 15) (12, 9) (13, 14) (14, 12) (14, 12) (15, 16) (16, 13)

Calculate the rank correlation co-efficient for proficiencies of this group in Mathematics and Physics.

⇒ Solution :-

Ranks in Maths (x)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Ranks in Physics (y)	1	10	3	4	5	7	2	6	8	11	15	9	14	12	16	13
$d = x - y$	0	-8	0	0	0	-1	5	2	1	-1	-4	3	-1	2	-1	3
d^2	0	64	0	0	0	1	25	4	1	1	16	9	1	4	1	9

Total $(d^2) = 136$

∴ Rank correlation co-efficient is given by:

$$\rho = 1 - \frac{6 \Sigma d^2}{n(n^2-1)} = 1 - \frac{6 \times 136}{16 \times 255} = 1 - \frac{1}{5} = \frac{4}{5} = 0.8$$

7) Find the equation of the least square straight line (considering x as independent variable) using the following data:

x : 1 3 4 6 8 9 11 14
y : 1 2 4 4 5 7 8 9

⇒ Solution :

	x	y	x^2	xy	y^2
	1	1	1	1	1
	3	2	9	6	4
	4	4	16	16	16
	6	4	36	24	16
	8	5	64	40	25
	9	7	81	63	49
	11	8	121	88	64
	14	9	196	126	81
$\Sigma x =$	56	40	524	364	251

Therefore, the normal equation becomes,

$$40 = 8a + 56b$$

$$364 = 56a + 524b.$$

Finding the solution of the above equations we get,

$$a = \frac{6}{11}, \quad b = \frac{7}{11}$$

∴ The equation of least square st. line is

$$y = \frac{6}{11} + \frac{7}{11}x.$$

8) Test the hypothesis that the number of Parts demanded does not depend on the day of the week. (Given, the values of chi-square significance at 5, 6, 7 d.f are respectively 11.07, 12.59, 14.07, at the 5% level of significance.)

=> Solution :-

Here we set up the null hypothesis, so that the number of parts demanded does not depend on the day of week.

Under the null hypothesis, the expected frequencies of the spare part demanded on each of the six days should be :- $\frac{1}{6}(1124 + 1125 + 1110 + 1120 + 1126 + 1115)$

$$= \frac{6720}{6}$$

$$= 1120$$

P.T.O

Calculations for χ^2

Days	Frequency Observed (f_i)	Frequency Expected (e_i)	$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
Mon	1124	1120	16	0.014
Tues	1125	1120	25	0.022
Wed	1110	1120	100	0.089
Thurs	1120	1120	0	0
Fri	1126	1120	36	0.032
Sat	1115	1120	25	0.022
Total	6720	6720		0.179

$$\chi^2 = \Sigma \frac{(f_i - e_i)^2}{e_i} = 0.179$$

The number of degrees of freedom = 6 − 1 = 5 (since we are given 6 frequencies subjected to only one linear constrain!

$\Sigma f_i = \Sigma e_i = 6720$)

The tabulated $\chi^2_{0.05}$ for 5 d.f = 11.07

Since calculated value of χ^2 is less than the tabulated value it is not significant and the null hypothesis may be ~~ac~~ accepted at 5% level of significance. Hence we conclude that the number of parts demanded are same over the 6-day period.