

**COVENANT UNIVERSITY  
NIGERIA**

*TUTORIAL KIT  
OMEGA SEMESTER*

**PROGRAMME:  
DEMOGRAPHY AND SOCIAL STATISTICS**

**COURSE: DSS 222**

## **DISCLAIMER**

The contents of this document are intended for practice and leaning purposes at the undergraduate level. The materials are from different sources including the internet and the contributors do not in any way claim authorship or ownership of them. The materials are also not to be used for any commercial purpose.

**DSS222**  
**STATISTICAL METHODS & SOURCES II**

**BY**  
**OLAWOLE-ISAAC, A.**

1. What is Hypothesis?
2. Discuss the types of hypotheses
3. Explain type I and type II error
4. What are the procedures for testing a statistical hypothesis
5. The mean weight of 64 men is 63kg. The mean weight for the population is 60kg, with a standard deviation of 12kg. Is the mean of the sample significantly different from the population mean weight at the 5% level.
6. A firm ordered sack of chemical with a normal weight of 60 kg and a random of 10 sacks was taken and it was found that that the mean is 50kg with a standard deviation of 2.0kg. The firm wishes to test whether the mean weight of sample is less than the nominal weight at 1% and 5% level of significance.
7. What is correlation analysis?
8. Explain correlation coefficient
9. Explain correlation of qualitative data and reasons for ranking
10. The mean life of an electric bulbs produced by a company has in the past been 1120h with a standard deviation of 125h. A sample of eight electric light bulbs recently chosen from a supply of newly produced bulbs showed a mean life of 1070h. Test the hypothesis that the mean life of the bulbs has not changed, using significance levels of (a) 0.05 and (b) 0.01.
11. What is rank correlation and what are the procedures of calculation?
12. An experiment was conducted on 8 children to determine how a child reading ability varied with his ability to write. The points awarded were as follows:

child	A	B	C	D	E	F	G	H
Writing	7	8	4	0	2	6	9	5
Reading	8	9	4	2	3	7	6	5

Calculate the coefficient of rank correlation and interpret your result.

13. An estate valuer, Feyi Akin and associates on visual analysis suspect the existence of a linear relationship between the rate of deterioration of an estate and the mode of payment of rent by tenants. The table is presented below:

Rate of deterioration	10	9	7	11	8	9	11	8
Mode of payment	4	6	3	5	7	11	8	6

Using 0.05 test whether there exist a linear relationship between the variables

14. Ten plots of land were treated with fertilizers produced by Miracle and co. and 12 by Esther and co. The yield of the first plots have a mean yield of 6.00 bushels with a variance of  $9 \times 10^{-4}$  bushels. The yield of the second plot have a mean yield of 5.95 bushels with a variance of  $1.6 \times 10^{-3}$  bushels. Is there a difference in the effects of the fertilizers at 0.01 level of significance.

15. Three coins were tossed a total of 240 times, and each time the number of heads turning up was observed. The results are shown in the table below, together with the results expected under the hypothesis that the coins are fair.

	0 Heads	1 Head	2 Heads	3 Heads
Observed frequency	24	108	95	23
Expected frequency	30	90	90	30

State the null and alternative hypothesis and test this hypothesis at a significant level of 0.05.

16. Given that  $\sum XY = 242$ ,  $\sum X^2 = 96.25$ ,  $\sum Y^2 = 640$ ,  $n = 10$ ,  $\bar{Y} = 27.5$  and  $\bar{X} = 7.6$

Obtain the first order regression equation.

Determine the strength, direction of the linear relationship and interpret your result.

17. The prices and commodity of three commodities milk, butter and sugar from 1980-1983 are given below:

Commodity	Prices				Quantity			
	1980	1981	1982	1983	1980	1981	1982	1983
Milk	20	30	25	40	120	140	180	200
Butter	35	15	32	30	190	110	105	100
Sugar	43	38	30	20	70	75	80	62

Find a simple price index of 1981 and 1983 using 1982 as the base year.

18. The following data are the wheat prices per bushel in the United States for the years 1965 to 1974.

1965	1966	1967	1968	1969	1970	1971	1972	1973
2.58	2.12	2.33	2.13	2.31	2.39	2.44	2.52	2.54

Construct a simple index of the wheat prices using 1968 as the base year. What is the index number for 1973?

19. Using the table in question 17 determine the weighted price index of 1983 using 1981 as the base year using paache price index
20. Determine the simple price index of 1981 using 1982-1983 as the base year.

## ANSWER

1 Statistical hypothesis is a statistical statement which may or may not be true concerning one or more population.

3 Type 1 and type 11 error: if we reject a hypothesis when it should be accepted we say that type 1 error has been made else when we accept a hypothesis when should be rejected we say type 11 error has occurred.

The probability of committing type 1 error is called alpha. The probability of risking a type 1 error is denoted as  $\alpha$  while a type two error is  $\beta$  and is referred to level of significance either 5% or 1 %.

5

$H_0$ : mean weight = 60kg

$H_1$ : mean weight  $\neq$  60kg

$\alpha = 0.05$ ,  $\bar{x} = 63$ ,  $\mu = 60$ ,  $\sigma = 12$ kg and  $N = 64$

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{N}} = \frac{63 - 60}{12 / \sqrt{64}} = 3 / 1.5 = 2.0$$

Since it is two tail test check z table for  $(1 - \alpha/2) = 1 - 0.05/2 = 1 - 0.025 = 0.975$ . Therefore Z tabulated = 1.96.

Conclusion: since the calculated Z value of 2.0 is greater than the tabulated Z value of 1.96 at 5 % level of significance, we reject the null hypothesis  $H_0$  and conclude that the mean weight of the sample is not 60 kg.

7 Answer: correlation provides a single number that summarizes the degree of relationship between two variables.

9 The dependent variable is denoted as y and independent variable as x.

It comprise of ordinal level of variables where ranking are used instead of instead of the initial values of the variables.

Reasons for ranking

It is easier to compute without distorting much the degree of association

Ranks are sometimes more meaningful to work with

Certain characteristics like the qualitative data can be ranked.

11 Rank correlation is used to compute relationship between variables when the precise values of the variables are not available.

The data are therefore ranked in their order of importance or size.  
procedure

1. Rank the values of X from 1 to n where n is the numbers of pairs of values of X and Y in the sample.
2. Rank the values of Y from 1 to n.
3. Compute the value of  $d_i$  for each pair of observation by subtracting the rank of  $Y_i$  from the rank of  $X_i$
4. Square each  $d_i$  and compute  $\sum d_i^2$  which is the sum of the squared values.

13

X	10	9	7	11	8	9	11	8
Y	4	6	3	5	7	11	8	6

XY	40	54	21	55	56	99	88	48
Y <sup>2</sup>	16	36	9	25	49	121	64	36
X <sup>2</sup>	100	81	49	121	64	81	121	64

$$\sum Y/n = 6.25, \sum X/n = 9.13$$

H<sub>0</sub>: there is no linearity between X and Y

H<sub>1</sub>: there is linearity between X and Y

$$\begin{aligned} SYX &= \sum XY - n (\sum Y/n) (\sum X/n) \\ &= 461 - 8 (9.13)(6.25) = 4.75 \end{aligned}$$

$$\begin{aligned} SXX &= \sum X^2 - n (\sum X/n)^2 \\ &= 681 - 8 (9.13)^2 = 14.875 \end{aligned}$$

$$\begin{aligned} SYY &= \sum Y^2 - n (\sum Y/n)^2 \\ &= 356 - 8(6.25)^2 = 43.5 \end{aligned}$$

Source of Variation      degree of freedom      sum of square      mean of square      F calculated

Due to Regression	1	1.517	1.517	0.2168
Due to Residual	6	41.983	6.997	
Total	7	43.5		

$$F_{0.05}^T (2-1)(8-1)$$

$$F \text{ tabulated} = 5.99$$

15

Observed	Expected	(O-E)	(O-E) <sup>2</sup>	(O-E) <sup>2</sup> / E
24	30	-6	36	1.2
108	90	18	324	3.6
95	90	5	25	0.28
23	30	-7	49	1.63

H<sub>0</sub> = there is no association between the three coins tossed

H<sub>1</sub> = there is association between the three coins tossed

$$\chi^2 = \sum (o-e)^2 / e = 6.71$$

$$\chi^2 = 1 - 0.05 = 0.95$$

$$= 0.95 (i-1)(j-1)$$

$$= 0.95 (2) (2)$$

$$= 9.49$$

$$\chi^2 \text{ tab} = 9.49$$

$$17 \text{ P}_{1982/1981} = \frac{\sum P_{1981}}{\sum P_{1981}} \times 100 = 95.4\% \text{ Interpretation: the price decreases by 4.6\%}$$

$$P_{1982/1983} = \frac{\sum P_{1983}}{\sum P_{1982}} \times 100 = 103.4\% \text{ Interpretation: the price increases by 3.4\%}$$

$$19: \frac{\sum P_{1983} * q_{81}}{\sum P_{1981} * q_{83}} \times 100 = 117.7\% \text{ the price index increases by 17.7\%}$$

$P_{83} q_{83}$	8000	3000	1240	12240
$P_{83} q_{81}$	5600	3300	1500	10400
$P_{81} q_{81}$	4200	1650	2850	8700