CHAPTER 7

HYPOTHESIS TESTING

They are two variables to be tested statistically:(1) The effects of schistosomiasis morbidity on labour productivity.(2) The effects of schistosomiasis morbidity on school performance of children.

7.1 THE EFFECTS OF SCHISTOSOMIASIS MORBIDITY ON LABOUR PRODUCTIVITY.

The aim is to test whether the different labour productivity values produced by the index under the three health status and four age groups, if are statistically significant.

7.1.2 Statement of the null hypothesis :

There is no difference of labour productivity among household members in the three health status i.e. (h_y, h_i and h_m). Symbolically is:

$$HO : \alpha_i^j(h_y) = \alpha_i^j(h_i) = \alpha_i^j(h_m)$$

Alternative hypothesis is:

$$H1: \alpha_i^j(h_y) \neq \alpha_i^j(h_i) \neq \alpha_i^j(h_m)$$

F TEST of household member's labour productivity per day in different health status is provided by applying two way analysis of variance. On the left hand side of table 1.7, it shows the blocks (Observations), which is the household members age groups. The right hand side, represents treatments (health status), which is a total number of groups.

Table 7.1 Statistical Test for labour Productivity

Observat (Blocks)	h_y	h_i	h_m	Total	Mean
1	5	2	0.48	7.48	2.49
2	37	18	3.6	58.6	19.53
3	50	25	4.8	79.8	26.6
4	25	13	1.54	39.54	13.18
Total	117	58	10.42	185.42	
Mean	29.25	14.5	2.6		

From the above table, which shows different productivity values expressed in Tanzania currency, the observations 1, 2, 3, and 4 indicates the four age groups classified in each household. The age groups are 07 - 10, 11 - 14, 15 - 60 and over 60 respectively. On the other hand, variables h_y, h_i, and h_m shows the three different health status faced by household members in schistosomiasis endemic area.

h_y = healthy

h_i = infection

h_m = morbidity

$$SST = \sum_{j=1}^{k} \sum_{i=1}^{n} x^2_{ij} - \frac{T_g^2}{kn}$$

Where:
SST= Total sum of squares.

$\sum_{j=1}^{k}$ = Summation of treatments (groups)

$\sum_{i=1}^{n}$ = Summation of blocks (Observations)

k = Number of treatments (groups)

n = Number of blocks (Observations)

x_{ij} = The i^{th} observation that receives the j^{th} treatment

T_g = Grand total

Therefore,

SST=

$(5^2+2^2+0.48^2+37^2+18^2+3.6^2+50^2+25^2+4.8^2+25^2+4.8^2+25^2+13^2+1.5^2)$

$- (2865)$

SST = 5679.6 - 2865 = 2814.6

$$SSB = \frac{\sum_{i=1}^{n} T_i^2}{k} - \frac{T_g^2}{kn}$$

SSB = Sum of squared blocks

T_i = Total of the i^{th} block

$$\frac{(7.48^2+58.6^2+79.3^2+39.54^2)}{3} - (2865) = 942.12$$

$$SS(Tr) = \frac{\sum_{j=1}^{k} T_j^2}{n} - \frac{T_g^2}{kn}$$

whereby:

SS(Tr) = Sum of the squared treatments

T_j = Total of the j^{th} observation

Therefore, $SS(Tr) = (117^2 + 58^2 + 10.42^2) - (2865)$

= 1425.4

$$SSE = SST - SSB - SS(Tr)$$
$$= 2814.6 - 942.12 - 1425.4$$
$$= 447.1$$

VARIATION	DEGREES OF FREEDOM
SS(Tr) =................	k - 1 = 2
SSB = ...	n - k = 3
SSE = ...	(k-1)(n-1) = 6
SST = ...	kn-1 = 11

$$Ms(Tr) = \frac{SS(Tr)}{k-1} \qquad \frac{1425.4}{2} = 712.7$$

$$MSB = \frac{SSB}{n-1} \qquad \frac{942.12}{3} = 314.04$$

$$MSE = \frac{SSE}{(k-1)(n-1)} \qquad \frac{447.1}{6} = 74.5$$

$$F = \frac{MS(Tr)}{MSE} \qquad \frac{712.7}{74.5} = 9.57$$

Table 7.2 ANOVA TABLE FOR LABOUR PRODUCTIVITY

SOURCE	SS	DF	MS	F
Treatments	1425.4	2	712.7	9.57
Blocks	942.12	3	314.04	
Error	447.1	6	74.5	
Total	2814.6	11		

The computed F value is 9.57. Referring to F table of statistics the critical value of F at 2 degrees of freedom in the numerator and 6 in the denominator is 5.14 at 5% significance level. This shows that, the calculated value is greater than the critical value hence we reject the null hypothesis.

However, by rejecting the null hypothesis it implies that the labour productivity values produced by the index are statistically different. This leads us to conclude that Schistosomiasis in endemic area might be among the factors which affects labour productivity of households. This interpretation of results is only valid if the assumptions in the study holds. Other wise, if real data would have been used, and the results shows that computed F is less than F critical value, then we would be required to accept the null hypothesis and conclude that Schistosomiasis in endemic area does not affect labour productivity of households. This would mean that, productivity values produced by the index are statistically insignificant, thus they do not show real difference. It further suggests that, there is no enough evidence to show whether Schistosomiasis in endemic area affects labour productivity of households.

7.2 THE EFFECTS OF SCHISTOSOMIASIS ON SCHOOL PERFORMANCE OF CHILDREN.

7.2.1 HYPOTHETICAL DATA OF SCHOOL CHILDREN PERFORMANCE IN SCHISTOSOMIASIS ENDEMIC AREA AND FREE SCHISTOSOMIASIS AREA.

Given that, School performance index has identified 220 pupils out of 300 as had poor school performance in Schistosomiasis endemic area and in schistosomiasis free area 100, for the same sample size. This implies that, in Schistosomiasis endemic area 73% of pupils had poor school performance compared to 33% in Schistosomiasis free area. This hypothetical situation shows that, School performance of pupils is affected by Schisto in endemic area than in Schisto free area.

However, we need to test the differences of the two sample proportions, so as to draw conclusion whether the observed differences in school performance is a true one, or has been achieved by chance.

7.2.2 TEST OF DIFFERENCES BETWEEN TWO PROPORTIONS

The two sample proportions are obtained from two independent areas, one from schisto endemic area and the other from schisto free area.

Let p_1 = Proportion of school children in endemic area who had poor school performance.

Let p_2 = Proportion of school children in Schisto free area who had poor school performance.

Let n_1 = Sample size in Schisto endemic area

Let n_2 = Sample size in Schisto free area

The applied statistic test is :

$$Z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}_c \bar{q}_c \left(\frac{1}{n} + \frac{1}{n} \right)}}$$

Whereby:

$$\bar{p}_c = \frac{\text{sum of successes in two samples}}{n_1 + n_2}$$

$$\bar{q}_c = 1 - \bar{p}_c$$

$$\bar{p}_1 = \frac{\text{number of successes in sample 1}}{n_1}$$

$$\bar{p}_2 = \frac{\text{number of successes in sample 2}}{n_2}$$

7.2.3 HYPOTHESIS STATEMENT

The hypothesis and its decision rule is:

$H_o : p_1 - p_2 \leq 0$ $H_a : p_1 - p_2 > 0$ Reject H_o if sample $Z > Z\alpha$

Let α = 0.05 with value 1.645

From the data given :

$$\bar{p}_1 = \frac{220}{300} = 0.73$$

$$\bar{p}_2 = \frac{100}{300} = 0.33$$

$\bar{p}_c$ = 0.53, $\bar{q}_c$ = 0.47 n_1 = 0.0033 n_2 = 0.0033

Substituting the values in the Z formula we get:

Z = 243.31

Thus, calculated Z = 243.3 is greater than $Z_{0.05}$ from the table which is 1.645. We reject the null hypothesis and conclude that the two sample proportions are statistically different. This implies that, in Schistosomiasis endemic area, poor school performance of pupils might be due to the effect of Schistosomiasis. However, this conclusion will be valid if the samples in both areas have been selected randomly and there is no major differences in the socio economic characteristics between the two areas. If other factors such as differences in the availability of teaching materials, availability of good teachers, and conducive study environment are not identical, then these results may not be valid.

However, if the results would have been indicated that Z < $Z\alpha$, then we would have to accept the null hypothesis that the observed difference between the two sample proportions is not a real difference. The difference, has been observed by chance. This implies that, there is no enough evidence to show that in Schisto endemic area poor school performance is due to the presence of Schisto as compared to Schisto free area.