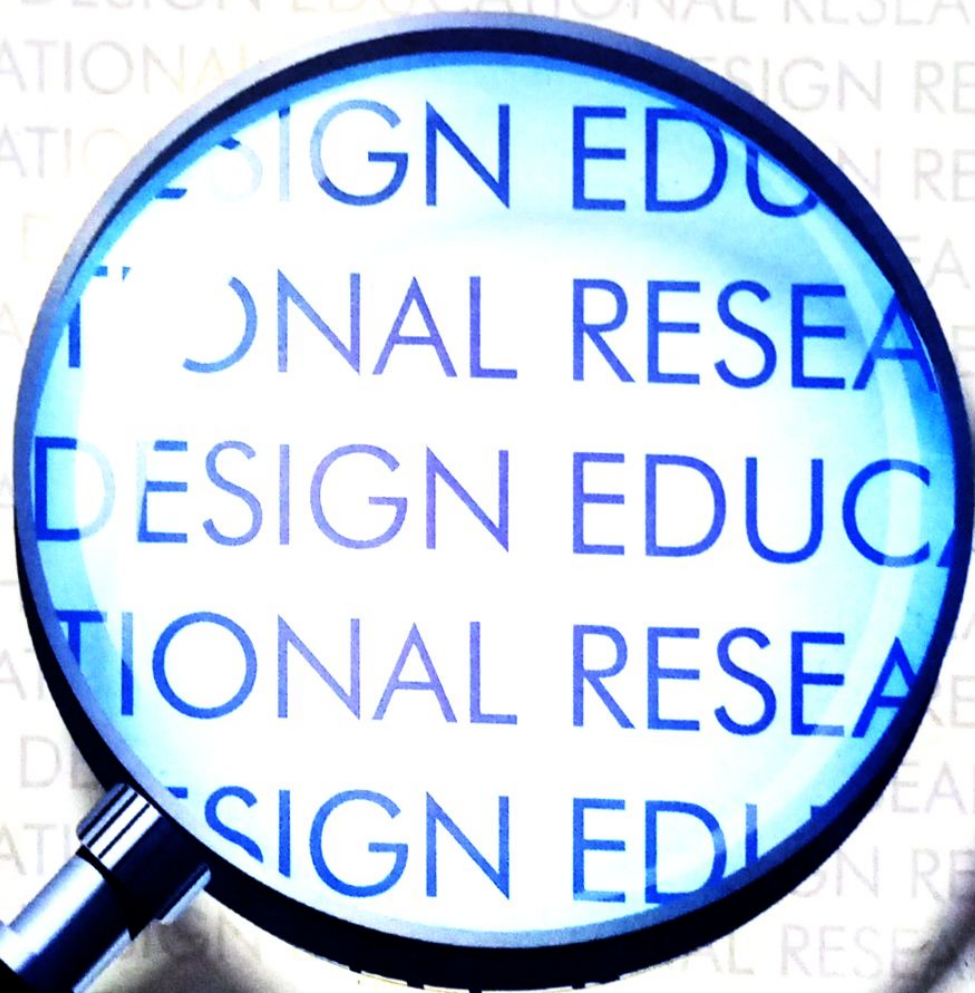


EDUCATIONAL RESEARCH DESIGN



EDITED BY:
H. O. Owolabi

Copyright (c)

Department of Social Sciences Education, University of Ilorin,
July, 2015

All right reserved

No part of this publication may be reproduced or transmitted in any form or by means of electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the publisher

ISBN: 978-978-568-9

Printed in Nigeria by
INDEMAC Publisher(Nig) Ltd.
100B Ibrahim Taiwo Road,
Ilorin

TABLE OF CONTENTS

FOREWORD	iii
NOTES ON CONTRIBUTORS	iv
CHAPTER 1	
Concept of Research	1
CHAPTER 2	
Historical Research In Education	8
CHAPTER 3	
Case Studies Method	17
CHAPTER 4	
Survey Research	21
CHAPTER 5	
Longitudinal Studies	30
CHAPTER 6	
Cross-section Studies	36
CHAPTER 7	
Correlation Studies	39
CHAPTER 8	
Causal Comparative Studies (Ex-Post Facto).....	45
CHAPTER 9	
Quasi-experimental Research	51
CHAPTER 10	
Experimental Research	54
CHAPTER 11	
Basic Steps In The Conduct Of Educational Research.....	61
CHAPTER 12	
Review Of Literature.....	73

CHAPTER 13	
Research Design And Instrumentation.....	78
CHAPTER 14	
Proposal Writing In Social Sciences Research	85
CHAPTER 15	
Research Report Writing	94
CHAPTER 16	
Referencing And Bibliographic Styles	104
CHAPTER 17	
Chi-Square Test	116
CHAPTER 18	
Student's T-TEST.....	129
CHAPTER 19	
Analysis Of Variance (ANOVA).....	138
CHAPTER 20	
Correlation Coefficient.....	153
CHAPTER 21	
Qualitative Research In Education.....	161
CHAPTER 22	
Field Studies In Educational Research.....	171

CHAPTER 17

CHI-SQUARE TEST

J. Oladele

Statistics is a type of mathematical analysis involving the use of quantified representations, models and summaries for a given set of empirical data used in quantitative research or real world observations or a measure based on observations of the characteristics of a sample commonly used in qualitative research. Statistical analysis involves the process of collecting and analyzing data (quantitative or qualitative) and then summarizing the data into a numerical form or facts.

A statistic computed from a sample may be used to estimate a parameter, the corresponding value in the population from which the sample is selected. The information about a population may also be determined by its parameters than a statistical test and this brings about the parametric and non-parametric statistics.

Parametric methods are those whose population is approximately normal; that is, a few students do very well and a few do very poorly which makes a bunch of the scores end up clumped around the mean score. Different tests may result in different means and standard deviations, but the shape of the graph is nearly always the same; commonly called the bell curve. Parametric methods are typically the first methods studied in an introductory statistics course. **Non-parametric methods on the other hand** are statistical techniques for which assumption of normality for the population under study is not made and so this method does not have any dependence on the population of interest and can therefore be referred to as distribution free methods (Alarape, 2005). A few parametric methods include: T-test, Z-test, F-test, and Analysis of Variance (ANOVA) while some commonly used parametric methods are Spearman's rank order correlation test and Chi-Square; Chi-Square being the focus of this chapter.

In the simplest cases, the data are organized into two categories, such as yes and no, high and low, for and against (Frankel & Wallen, 2000). Although, less statically powerful, chi-square statistics is becoming popularized because it becomes one of the safe options when the assumption of normality cannot be made and as such, it is used to test whether the observed are statistically different from what is expected (Coakes & Steed, 2011).

CONDITIONS FOR USING CHI-SQUARE

1. When a researcher uses nominal data for comparison
2. When the observations are not limited by sample size, this makes the test useful in rare situations and by their nature contains few observations
3. When the variable under study is categorical and the research questions involves collecting data on a single variable used for comparison based on how the observations are categorized

4. The population is at least 10 times as large as the sample.

ASSUMPTION OF CHI-SQUARE TEST

1. Observations are randomly selected from a large population else generalizing the result should be done with caution because the sample may not represent the whole population
 2. Independence of observations which ensures that each observation should be generated by a different subject and no subject is counted twice
 3. The number of expected observation in any given category should be large with at least five (5) observations in each category in any given analysis
- Chi-square can be used to test either goodness of fit or test of independence.

COMPUTATION PROCEDURES FOR CHI-SQUARE

1. Determine whether the data were collected using nominal or ordinal scale
2. State the null (H_0) and the alternate (H_1) hypothesis
3. Perform the analysis using statistics and determine whether to reject or retain the hypothesis
4. Interpret results inline with the stated hypothesis
5. Calculate the chi-square score (χ^2)

FORMULA FOR CALCULATING χ^2

The formula for calculating χ^2 is given as:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

f_o : Observed frequency in a given category

f_e : Expected frequency in a given category

Note that the calculation of χ^2 can be easily accomplished by presenting it in tabular form.

CHI-SQUARE (χ^2) GOODNESS OF FIT

Chi-square goodness of fit test method is appropriate when a researcher attempts to fit a statistical model to the observed data if he or she wants to determine how close are the observed values to those which would be expected under the fitted model. This test is commonly used to test association of variables in two-way tables, where the assumed model of independence is evaluated against the observed data.

EXAMPLE USING MANUAL CALCULATION

A researcher is interested in determining whether there are more students who ask questions from the teacher during lessons compared to those who do not. To get the data on the variables, a questionnaire could be developed to collect the information after which assumptively, 44 reported that they sought help from library staff while 30 students did not. Ideally, one would expect that the number of students who sought help would be

similar to the number who did not and so the number will be 37 each as presented in the table below:

Table 1

Number of Students	f_o	f_e
Asked questions	44	37
Did not ask question	30	37

The data presented in Table 1 above deviates markedly from what is expected. The χ^2 enables the researcher to conclude whether the f_o are significantly different from f_e . The calculation of χ^2 can be easily accomplished by presenting it in the tabular form as shown in Table 2.

Table 2

Number of Students	f_o	f_e	$(f_o - f_e)$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
Asked questions	44	37	7	49	1.32
Did not ask question	30	37	-7	49	1.32
					$\sum \frac{(f_o - f_e)^2}{f_e}$ = 2.64

In column 6, the final quotient resulting from the decision of each squared score with the frequency expected is determined by summing each of the quotients.

TEST OF STATISTICAL SIGNIFICANCE

The chi-square statistic may be used to test the hypothesis of no association between two or more groups, populations, or criteria whereby observed counts are compared to expected counts.

Recall that the hypothetical example used for the computation expects that the level of students who ask questions from the teacher during lessons compared to those who do not are the same, the hypothesis can therefore be state as:

H_o : There is no significant difference in the number of students who ask questions from the teacher during lessons compared to those who do not.

The test of statistical significance is therefore to determine whether to reject or retain the null hypothesis (H_o) stated above. To make a decision, the calculated value (χ^2) will be compared with the table value; traced using the degree of freedom (df) for the goodness of fit test given as: $K-1$, therefore:

$$\begin{aligned} df &= k-1 \\ &= 2-1 = 1 \end{aligned}$$

DECISION RULE USING TABLE VALUES

If the calculated value is equal or greater than the critical value, then reject the H_0 and accept H_1 and vice versa.

With the df calculated as 1 and χ^2 value: 2.64, the critical value is: 3.84, H_0 can be statistically presented as: $\chi^2(1) = 2.64 \leq 3.84$

This shows that the calculated value is less than the critical value so the answer is to accept the H_0 stating that: there is no significant difference in the number of students who ask questions from the teacher during lessons compared to those who do not and this is mathematically presented as:

$$\chi^2(1) = 2.64, p > 0.05$$

The non-significance of the χ^2 test statistics implies that there is no statistically significant difference between students who ask questions from the teacher during lessons compared to those who do not.

EXAMPLE USING SPSS METHOD

A researcher is interested in determining whether there are more students who seek ask questions from the teacher during lessons compared to those who do not. To get the data on the variables, a questionnaire could be developed to collect the information after which assumptively, 44 reported that they sought help from library staff while 30 students did not. This can be presented in tabular form as shown in Table 3:

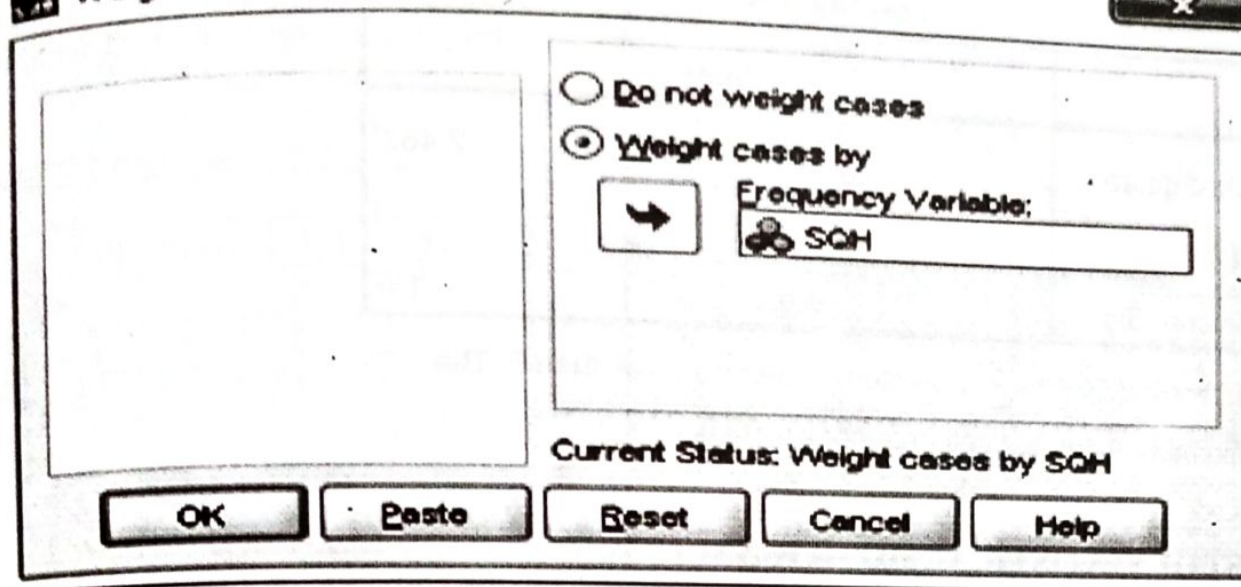
Table 3

Attitude of students towards asking questions during lessons	Frequency of responses
Asked questions	44
Did not ask question	30

Computational Procedure

1. Select the **Data** menu
2. Click on **Weighed Cases** to open the **Weighed cases** dialogue box
3. Click on **Weighed cases** by radio button
1. Select the variables required (i.e freq) and click on the ► button to move the variables into the **Frequency Variable:** box as shown below:

Weight Cases



Weight Cases dialog box. The 'Weight cases by' radio button is selected. The 'Frequency Variable' is 'SQH'. The 'Current Status' is 'Weight cases by SQH'. Buttons: OK, Paste, Reset, Cancel, Help.

☐ Do not weight cases

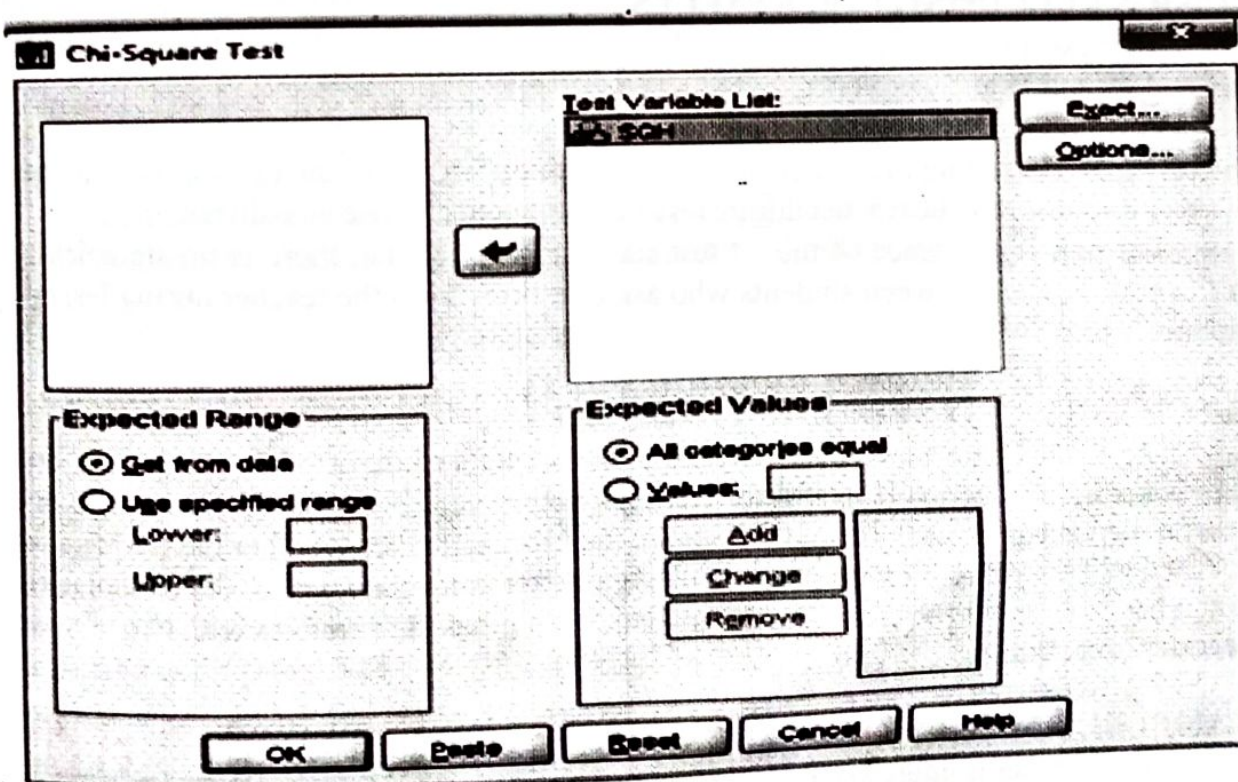
☒ Weight cases by

Frequency Variable: SQH

Current Status: Weight cases by SQH

OK Paste Reset Cancel Help

1. Click on **OK**. The message **Weight On** should appear on the status bar at the bottom of the application window
2. Select the **Analyze** menu
3. Click on **Nonparametric Test** and then **Chi-Square** to open the **Chi-Square Test** dialogue box
4. Select the variable requires (i.e. attitude) and click on the ► button to move the variable into the **Test Variable List** box as shown below:



Chi-Square Test dialog box. The 'Test Variable List' contains 'SQH'. The 'Expected Range' section has 'Get from data' selected. The 'Expected Values' section has 'All categories equal' selected. Buttons: OK, Paste, Reset, Cancel, Help, Expected..., Options...

Test Variable List: SQH

Expected Range

☒ Get from data

☐ Use specified range

Lower:

Upper:

Expected Values

☒ All categories equal

☐ Values:

Add Change Remove

OK Paste Reset Cancel Help

Expected... Options...

9. Click **OK**

Test Statistics

	SQH
Chi-Square	2.462 ^a
df	1
Asymp. Sig.	.117

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 52.0.

TEST OF STATISTICAL SIGNIFICANCE

Recall that the hypothetical example used for the computation expects that the numbers of students who ask questions from the teacher during lessons compared to those who do not are the same, the hypothesis can therefore be state as:

H₀: There is no significant difference in the number of students who ask questions from the

teacher during lessons compared to those who do not.

The test of statistical significance is therefore to determine whether to reject or retain the null hypothesis (**H₀**) stated above. This decision will be made using the output Assump. Sig Value (0.117).

DECISION RULE USING TABLE VALUES

If the assump. Sig. value is equal or less than the significance level 0.05; then rejects the **H₀** and accepts **H₁** and vice versa.

This shows that the assump. Sig. value is greater than the critical value so the answer is to accept the **H₀** stating that: there is no significant difference in the number of students who ask questions from the teacher during lessons compared to those who do not

The non-significance of the χ^2 test statistics implies that there is no statistically significant difference between students who ask questions from the teacher during lessons compared to those who do not.

CHI-SQUARE (χ^2) TEST OF INDEPENDENCE

The Chi-Square test is used where a researcher is interested in more than one categorical variable; which is useful for testing whether variables are independent of one another or dependent of each other. Chi-Square therefore indicates whether the frequencies associated with two variables are statistically significant independent of or dependent upon one another. Note should be taken that this statistics examines the variables with two or more categories in each other.

EXAMPLE USING MANUAL CALULATION

A researcher is interested in finding out whether students in their 1st, 2nd, 3rd and final year ask questions from the teacher during lessons; and so he is interested in finding out

whether this behavior is affected by the student's level of study. To accomplish this task, 267 students were randomly sampled with indications of their level of study and whether they ask questions from the teacher during lessons. The data collected based on their responses are shown in the contingency table 2

TABLE 2: Contingency Table 2a

Level of Study	Asked Questions	Did not ask Questions	Row Total
First	A 53	B 30	83
Second	C 42	D 29	71
Third	E 27	F 37	64
Fourth	G 20	H 29	49
Column Total	142	125	267

The number in each cell is called the observed frequency (f_o), the formula below is used to calculate the expected frequency (f_e).

$$f_o = \frac{\text{Column Total} * \text{Row Total}}{N}$$

$$(f_e)A = \frac{142 * 83}{267} = 44.14$$

$$(f_e)B = \frac{125 * 83}{267} = 38.86$$

$$(f_e)C = \frac{142 * 71}{267} = 37.76$$

$$(f_e)D = \frac{125 * 71}{267} = 35.11$$

$$(f_e)E = \frac{142 * 64}{185} = 30.04$$

$$(f_e)F = \frac{142 * 64}{185} = 29.96$$

$$(f_e)G = \frac{142 * 49}{267} = 26.06$$

$$(f_e)H = \frac{125 * 49}{267} = 22.99$$

Please, note that the sum of expected frequency (f_e) for all the cells as the case may be, 8 in this example must be equal to the total number (N). This is because in Chi-Square, the total sum of the observed frequency must ALWAYS equal the sum of the expected frequency. This also implies that the sum of the row and column total should be equal to N.

Contingency Table 2b shows the expected frequencies.

Table 2b: Contingency Table 2b

Cell	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
A	53	44.14	8.86	78.50	1.78
B	30	38.86	-8.86	78.50	2.02
C	42	37.76	4.24	17.98	0.48
D	29	35.11	-6.11	37.33	1.06
E	27	30.04	-3.04	9.24	0.31
F	37	29.96	7.04	49.56	1.65
G	20	26.06	-6.06	36.72	1.41
H	29	22.94	6.06	36.72	1.60
					$\chi^2 = 10.31$

Therefore, the Chi-Square value is: $\chi^2 = 10.31$

Note that Chi-Square score measures the discrepancy between the f_o and f_e scores. The larger the Chi-Square score, the larger the discrepancy and the more likely it is that the two variables being studied are related.

Recall that the null hypothesis (H_o) assumes that there is no relationship between the two variables, thus the larger the χ^2 , the smaller the probability that H_o is true.

TEST OF STATISTICAL SIGNIFICANCE

Like the Chi-Square goodness of fit was tested for statistical significance, Chi-Square test of independence also needs to be tested for statistical significance. This will enable the researcher to determine whether to reject or retain the stated hypothesis.

Recall that the hypothetical example used for the computation expects that the level of students; whether in their first, second, third and final year does not affect student's questioning habits during lessons; the hypothesis can therefore be state as:

H_o : There is no relationship between the student's level of study and their asking questions from their teacher during lessons.

The test of statistical significance is therefore to decide on rejecting or retaining the null hypothesis (H_o) stated above. To do this, the calculated value (χ^2) will be compared with the critical values; also referred to as table values.

RULE USING TABLE VALUES

Recall that if the calculated value is equal or less than the probability value, then; accept the H_0 .

Having calculated the χ^2 score, the next step is to calculate the degree of freedom using the formula: $df = (r-1)(c-1)$; where: r = number of rows and; c = number of columns

Using the contingency Table 2a, r is 4; representing the 4 levels of students study periods and $c = 2$; representing students who asked questions during lessons and those who do not ask questions during lessons.

$$\text{Therefore: } df = (4-1)(2-1) \\ = 3 \times 1 = 3$$

With the df calculated as 3 and χ^2 value: 10.31, the critical value is: 7.81, H_0 can be statistically presented as: $\chi^2(3) = 10.31 \geq 7.81$

This shows that the calculated value is greater than the critical value so the decision is to reject the H_0 stating that: there is no relationship between the student's level of study and their asking questions from their teacher during lessons and this is mathematically presented as:

$$\chi^2(3) = 10.31, p < 0.05$$

The significance of the χ^2 test statistics implies that there is statistically significant difference between the level of students and their questioning behavior to their teachers during lessons.

Having established that there is relationship between the 2 variables; asking questions from teachers during lessons and level of study, the pattern of relationship should be established and this can be done by comparing observed frequencies to expected frequencies. This can be best established by setting up a single task to compare the f_o and f_e clearly.

Table 3

Level of Study	Asked Questions	Did not ask Questions	Row Total
First	53 (44.14)	30 (38.86)	83
Second	42 (37.76)	29 (35.11)	71
Third	27 (30.04)	37 (29.96)	64
Fourth	20 (26.06)	29 (22.94)	49
Column Total	142	125	267

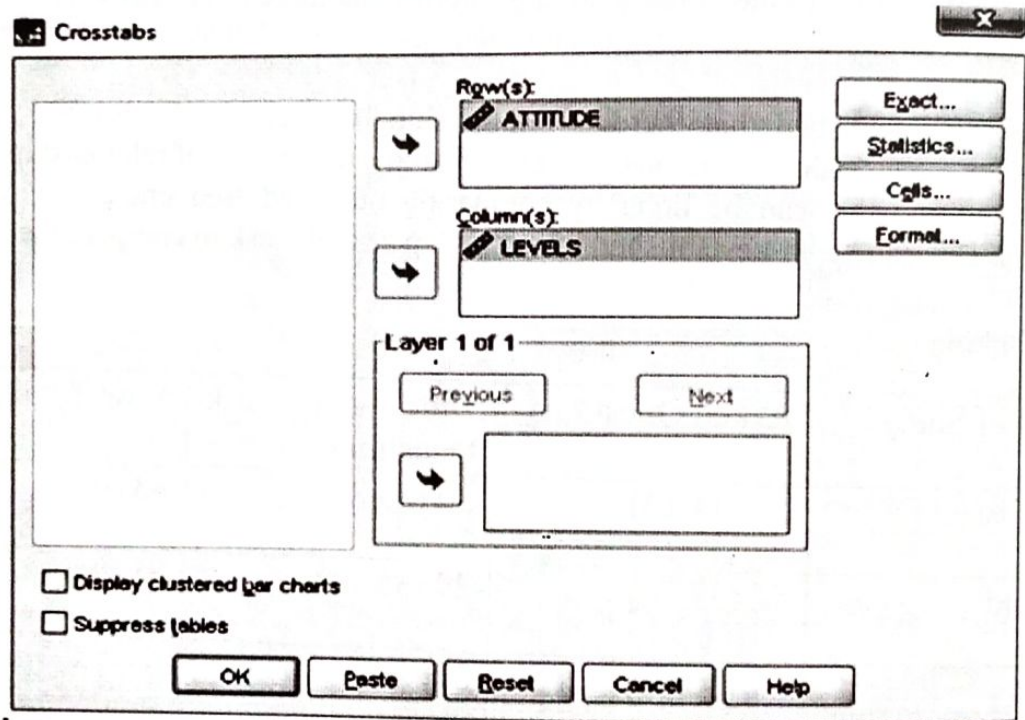
The comparism can either start from those who ask questions and to those who do not.

EXAMPLE USING SPSS METHOD:

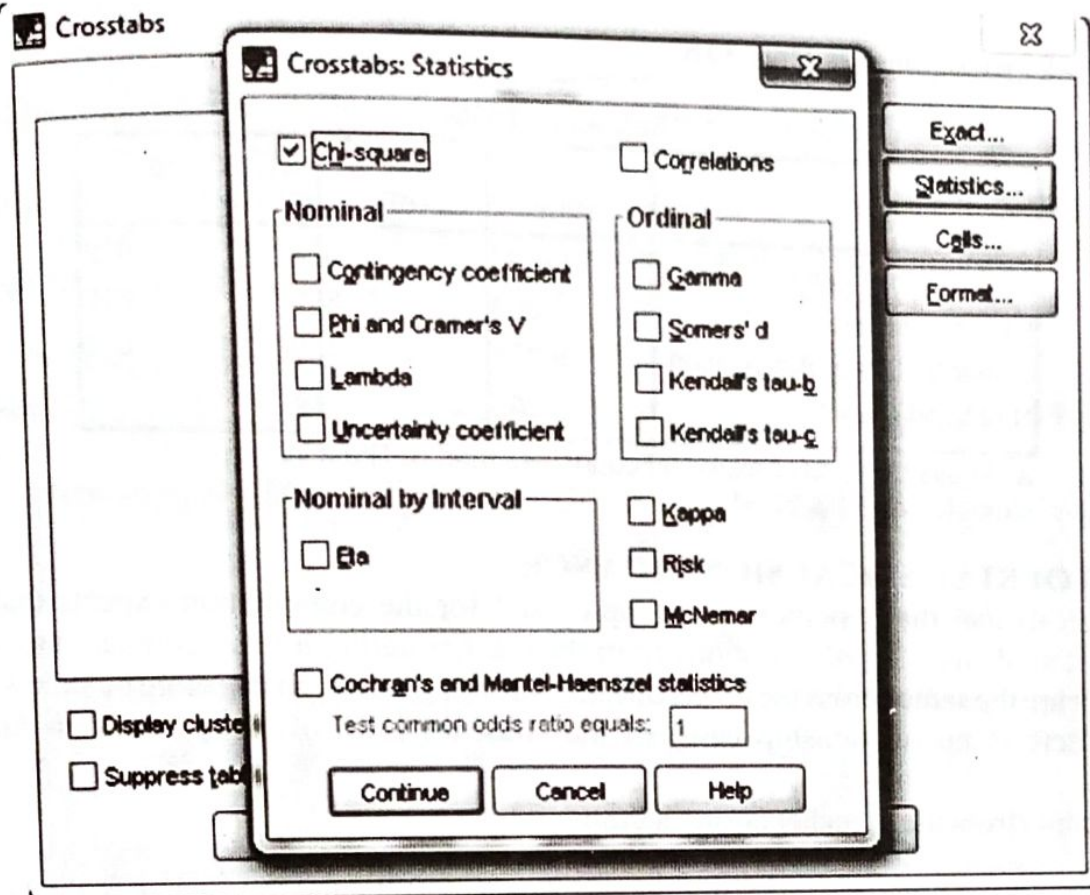
A researcher is interested in finding out whether students in their 1st, 2nd, 3rd and final year ask questions from the teacher during lessons; and so he is interested in finding out whether this behavior is affected by the student's level of study. To accomplish this task, 267 students were randomly sampled with indications of their level of study and whether they ask questions from the teacher during lessons.

Computational Procedure

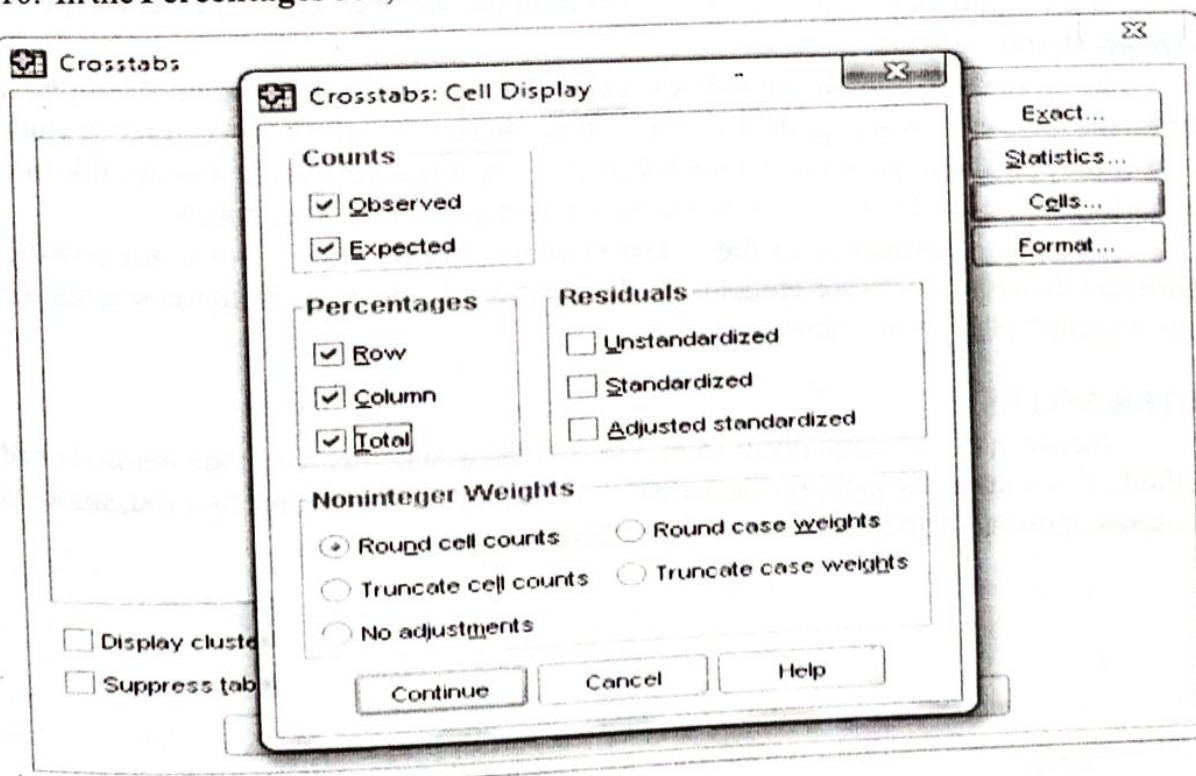
1. Select the **Analyze** menu
2. Click on **Descriptive statistics** and then click on **Crosstabs** to open Crosstabs dialogue box
3. Select a row variable (i.e. attitude) and click on on the ► button to move the variable into the **Row(s):** box
4. Select a column variable (i.e. levels) and click on the ► button to move the variables into the **Column(s):** box



5. Click on the **statistics** command push button to open the **Crosstab: Statistics** sub-dialogue box
6. Click on the Chi-Square check box



7. Click on **Continue**
8. Click on **Cells** command pushdown button to open the **Crosstabs: Cell Display** sub dialogue box
9. In the **Column** box, click on the **Observed** and **Expected** check boxes
10. In the **Percentages** box, click on **Rows**, **Column** and **Total** check boxes



11. Click on Continue and then OK

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.931 ^a	3	.012
Likelihood Ratio	11.000	3	.012
Linear-by-Linear Association	9.825	1	.002
N of Valid Cases	267		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 22.94.

TEST OF STATISTICAL SIGNIFICANCE

Recall that the hypothetical example used for the computation expects that the numbers of students who ask questions from the teacher during lessons compared to those who do not are the same across the levels of study, the hypothesis can therefore be state as:

H₀: There is no relationship between the student's level of study and their asking questions habits from their teacher during lessons

The test of statistical significance is therefore to determine whether to reject or retain the null hypothesis (**H₀**) stated above. This decision will be made using the output Pearson statistics, which is short for Person's chi-square.

In the above example, Pearson has a value of (10.931^a) with a significance of (.012).

DECISION RULE USING TABLE VALUES

If the significant value is equal or less than the alpha level 0.05; then accept the **H₀** and reject **H₁** and vice versa.

From the analysis, the significant value the sig. (.012) is less than the alpha level 0.05 so the answer is to accept the **H₀** stating that: there is a significant difference student's level of study and their questions asking habits from their teacher during lessons; this implies that the higher the level of the students, the better their question asking habits.

The non-significance of the ² test statistics implies that there is no statistically significant difference between students per level who ask questions from the teacher during lessons compared to those who do not.

AFTER NOTES

Please, note the same result should be arrived at whether using the manual or SPSS methods. Current trends supports the use SPSS method over the manual method, students are advised to familiarize themselves with the SPSS packages.

Bibliography

- Abiri, J.O.O. (2007). Elements of Evaluation Measurement & Statistical Techniques in Education. Ilorin: Unilorin Press.
- Alarape, A.I. (2005). Chapter 8: Chi-Square Test. In Gbadegesin, A, Olapoenia, R & Jerome, A. (Eds), Statistics for Social Sciences (Page 98-107). Ibadan: Uni. Press.
- Frankel, J.R. & Wallen, N.E. (2000). How to Design and Evaluate Research in Education. USA: Von Hoffemam Press Inc
- Olufunmilola, O.I.E. (2010). Introduction to Research Methodology. Akoko: Fruit of the Spirit Publishing House.