academicJournals

Vol. 8(15), pp. 1231-1241, 10 August, 2013
DOI: 10.5897/ERR11.0305
ISSN 1990-3839 ©2013 Academic Journals
http://www.academicjournals.org/ERR

Educational Research and Reviews

Full Length Research Paper

Conceptual model on application of chi-square test in education and social sciences

Sureiman Onchiri

Masinde Muliro University of Science and Technology, Kenya.

Accepted 16 May, 2013

Whenever you think you have an idea of how something works, you have a mental model. That is, in effect, a layman's way of talking about having an hypothesis. The hypothesis needs to be tested for how closely it fits reality - and reality is the data collected from an experiment. So the data is collected on the few and compared with a few controls. Is there really a difference between the two groups? It is that sort of question where Chi-square analysis comes in. In short, if the differences between your model and reality are small, that is good; if huge, develop a new model! These differences are denoted as "chi-square", which equals the sum of all the squares of the deviations divided by what was expected. A Chi-square test is one of the most frequently used tests with a number of improper applications. Some of the general causes of the improper applications include researchers not understanding the areas and conditions of application of the Chi-square test. To give solutions to the above problems, this paper explored the existing literature on the main areas of application of Chi-square, that include the test of frequencies (goodness of fit, homogeneity, independence) and the test of population variance. The paper identifies the shortfalls in the existing literature, and fills them by the application of appropriate illustrations and examples. To shield the loopholes in data analysis using Chi-square test, a simplified conceptual model which can be adopted by researchers is finally developed.

Key words: Hypothesis, observed frequency and expected frequency.

INTRODUCTION

The researching process is so important to the university that the credibility of university and its prestige is to an extent tied to the monumental researches that are tied to the university. This brings to fore the question of the quality in the research process. Therefore just as quality is important in the other stages of research process, so is it for data analysis.

According to Onyango and Odebero (2009), data analysis is significant in the following ways:

- Summary: helps in giving a didactic and ingenious summary of the data collected.
- Showing of relationship: After data has been systematically handled, different categories can be compared and contrasted. In this way the relationship between the data shall be known. In quantitative analysis, relationships and predictions among variables could be shown by correlation and regression techniques. In qualitative analysis, axial coding helps in showing relationships between categories.
- Clarity: Some tools of analyzing data make it possible to discern data clearly, like for example, the use of the histograms and bar graphs.
- Answering the research questions: it is after the analysis of data that the researcher will be able to discern whether she/he has adequately answered the research questions,

E-mail: onchirisureiman@yahoo.com.

if she/he has adequately tested the hypothesis of the research and the objectives.

- Data analysis is important in the write up: analysis and findings have influence on write up.
- Data analysis contributes, in instrumentation terms, showing the elements of validity and reliability of the research.
- After a comprehensive data analysis come the research findings that are important to the generalizations that shall be made about the research.

Statistical methods are extensively used in educational research. They provide an indispensable tool for collecting, organizing, analyzing and interpreting data expressed in numeric terms. By synthesizing the data, these methods can facilitate the derivation of conclusions and formulation of generalizations. Though it is an important tool of a researcher, there are a number of limitations that should be recognized in using statistical methods, and in drawing conclusions from statistical treatment of data. Best (1977) has listed them as under;

i) Statistical process is the servant of logic and only has value if it verifies, clarifies, and measures relationships that have been established by clear, logical analysis. Statistics is a mean, never an end of research.
ii) A statistical process should not be employed in the analysis of data unless the basic assumptions and limitations underlying its uses are clearly understood.
iii) The conclusions derived from statistical analysis will no more be accurate or valid than the original data. To use an analogy, no matter how elaborate the mixer, a cake made of poor ingredients will be poor cake. All the refinement of elaborate statistical manipulation will not yield significant truths if the data results from crude or inexact measurement.
iv) All treatment of data must be checked and double checked frequently to minimize the likelihood of errors in measurement, recording, tabulation and analysis.
v) There is a constant margin of error wherever measurement by human beings is involved. This error is increased when qualities or characteristics of human personality are subjected to measurement, or when inferences about the population are made from measurements derived from statistical samples. When comparisons or contrasts are made, a mere number difference is itself, not a valid basis for any conclusion. A test of statistical significance should be employed to weigh the possibility that chance in sample selection could have yielded the apparent difference. To apply these measures of statistical significance is to remove some of the doubt from conclusions.
vi) Statisticians and liars are often equated in humorous equips. There is little doubt that statistical processes can be used to prove nearly anything that one sets out to prove. Starting with false assumptions, using inappropriate procedures, or omitting relevant data, the biased investigator can arrive at false conclusions. These conclusions are often particularly dangerous because of the authenticity that the statistical treatment seems to confer.

Statistical techniques, if rightly employed, enable the researcher to analyse the data meaningfully and exactly; to summarize the results in meaningful and convenient form; and to draw general conclusions. Koul (1984) contends that no statistical method should be applied unless it adds clarity or meaning to the analysis and treatment of data. Data analysis can involve both descriptive and inferential statistics. According to Ingula and Gatimu (1996), descriptive statistics involves tabulating, graphing and describing data using measures of central tendency, variability, relative position, relations and association. Inferential statistics provides statistics that provides the method to be used for making inferences about a large group on the basis of small group findings (Mutai, 2000). Inferential method falls into two major categories: parametric and non-parametric approaches.

Parametric tests are more powerful than non parametric statistics. According to Koul (1984) use of parametric tests is based upon certain assumptions:

i) The variables must be expressed in interval or ratio but not in nominal or ordinal scale of measurement.
ii) The population values are normally distributed.
iii) The subjects are selected independently for the study.
iv) The samples have equal or nearly equal variances. This condition is known as equality or homogeneity of variances and is particularly important to determine when the samples are small. Gay (1976), on complementing the concept of equality, noted that at least the ratio of the variances is known.

According to Gay (1976), nonparametric tests make no assumptions about the shape of the distribution. They are used when the data represent an ordinal or nominal scale, when a parametric assumption has been violated, or when the nature of the distribution is not known (Gay, 1976). Onyango and Odebero (2009) summarizes the assumptions as the population is not known to be normal, sample size is small and the variables expressed as nominal or ordinal.

Non-parametric does not depend on the shape of the distribution of the population and hence are known as distribution free tests. The non-parametric tests are comparatively advantageous to parametric tests in various ways. First, they can be used in situations where stringent assumptions required by parametric tests do not have to be met. Second, they are less cumbersome to use as far as conceptual techniques are concerned. Third, they are most useful when dealing with qualitative values with data that can be classified in order of ranks. Finally, if the sample size is small, the parametric tests are not applicable unless the population distribution is exactly known which is generally not possible.

There are a number of non-parametric tests including Chi-square, Mann-Whitney U, Sign, Wilcoxon, Median, Kolmogorov-Smirnov and McNemar. Chi-square test is used with discrete data in form of frequencies. It is a test of independence and is used to estimate the likelihood that some factor other than chance accounts for observed relationship. The median test is used for testing whether two independent samples differ in central tendencies. It gives information as to whether it is likely that two independent samples have been drawn from populations with the same median. Kolmogorov-Smirnov test is a test of whether two independent samples have been drawn from the same population or from populations with the same distribution. If the two samples have been drawn from the same population distribution, then the cumulative distributions of both samples may be expected to be fairly close to each other, in as much as they both should show only random deviations from the population distribution. The Sign test is the simplest test of significance in the category of non-parametric tests. It makes use of plus and minus signs rather than quantitative measures as its data. It is particularly useful in the situation in which quantitative measurement is impossible or impracticable, but on the basis of superior or inferior performance it is possible to rank with respect to each other the two members of each pair. Wilcoxon matched-pairs signed- ranks test is more powerful than the sign test because it tests not only direction but also magnitude of differences within pairs of matched groups. McNemar test for the significance of changes is particularly applicable to those before and after designs in which each individual is used as his own control and in which measurement is expressed in either nominal or ordinal scale.

This study will consider the Chi-square test, because it is the one commonly used by researchers in comparison to the other non-parametric tests. Its correct application is an uphill task to most researchers (Kothari, 2007). Kothari (2007) notes that possible reasons of the improper application or misuse of this test can be neglect of frequencies of non-occurrence, failure to equalize the sum of observed and sum of expected frequencies, wrong determination of degrees of freedom and wrong computations. Researchers must understand the rationale of this test before using it and drawing inferences concerning the hypothesis. Most researchers start to use computer programmes before they understand the test manually which poses a great problem during the data analysis. Orodho (2005) identifies one of the common mistakes made in processing research data as not anticipating and appropriately choosing type of statistical tools to use such as Chi-square, Pearson's or regression at the start of analysis. This paper tried to address the above challenges in using the Chi-square test.

RESEARCH METHODOLOGY

This study involves the collection and critical analysis of the existing literature and information from various sources. The data was collected from secondary sources. Purposive sampling was used to gather the accessible relevant material. The study identified and filled some of the existing gaps on the analysis of data by Chi-square. Some of the most relevant illustrations were used to clarify the challenges. Finally, the researcher develops the conceptual framework which can guide researchers in the application of the Chi-square to test for frequencies (goodness of fit, homogeneity of a number of frequency distributions, independence) and population variance.

RESULTS AND DISCUSSION

To analyse the research results effectively and efficiently using the Chi-square test the paper addresses what various authors understand by the term Chi-square test. The procedure, area of application and the conceptual model is also outlined.

Definition of Chi-square test

Kothari (2007), Chi-square, symbolized as χ^2, is a non-parametric test of significance appropriate when the data is in form of frequency counts occurring in two or more mutually exclusive categories. A Chi-square test compares proportions actually observed in a study with the expected to establish if they are significantly different. The Chi-square value increases as the difference between observed and expected increase. Whether the calculated Chi-square value is significant is determined by comparing it with the value from table. If the calculated value exceeds the table value, the difference between the observed and expected frequencies is taken as significant otherwise it is considered insignificant.

Kothari (2007) notes that the following conditions should be met before the test can be applied. Unfortunately, Kothari (2007) has not given any explanation on the conditions.

i) Observations recorded and used are collected on a random basis. According to Gay (1976) it is the best single way to obtain a representative sample. In addition, it is required by inferential statistics to permit researchers to make inferences about populations based on behavior of samples. If observations are not randomly collected, then one of the major assumptions of inferential statistics is violated and inferences are correspondingly tenuous.

ii) All the members (or items) in the sample must be independent. This is to ensure that the occurrence of one individual observation (event) has no effect upon the occurrence of any other observation (event) in the sample under consideration.

iii) No group should contain very few items(less than 10). According to Yates (1934) the small data overestimate the statistical significance. Yates (1934) recommends this number as 5. Brown (2004) and Kothari (2007) concurs that controversies surrounds the number, some statisticians take this number as 5, but 10 is regarded as better by most statisticians.

iv) The overall number of items must be reasonably large

Table 1. Calculation expected frequencies (f_e).

	Category 1	Category 2	Category 3	Row Total
Sample X	A	B	C	A+B+C
Sample Y	D	E	F	D+E+F
Column totals	A+D	B+E	C+F	A+B+C+D+E+F=N

Table 2. General notation of a 2×2 contingency table.

		Variable 1		
		Data type 1	Data type 2	Total
Variable 2	Category 1	A	B	A+B
	Category 2	C	D	C+D
	Total	A+C	B+D	A+B+C+D=N

(at least 50). This assumption arises because the distribution of counts under the null hypothesis is multinomial, and the normal distribution can be used to approximate the multinomial distribution if the sample size is sufficiently large and the probability parameters aren't too small. It can be shown via the Central Limit Theorem that the multinomial distribution converges to the normal distribution as the sample size approaches infinity. Although chi-square does not require that the population distribution is normal, the assumption is that the deviations between the observed and expected values are uniform.
v) The constraints must be linear (contains no squares or higher powers of the frequencies).

Procedure for Chi-square test

1. State the hypothesis being tested and the predicted results: Gather the data by conducting the proper experiment. The relative standard to serve as the basis for accepting or rejecting the hypothesis is also determined. The relative standard commonly used in research is $\alpha > 0.05$. The α value is the *probability* that the deviation of the observed from that expected is due to chance alone (no other forces acting). In this case, using $\alpha > 0.05$, you would expect any deviation to be due to chance alone 5% of the time or less.
2. Calculate the Chi-square value. The following steps are involved in calculating the value;

i) Calculate the expected frequencies(fe) if not given
ii) Obtain the difference between observed (fo) and expected frequencies and find out the squares of these differences that is calculate $(fo—fe)^2$
iii) Divide the quantity $(f_0—f_e)^2$ obtained by corresponding frequency to get $(fo\text{-}fe)^2/fe$
iv) Find the sum of $(fo\text{-}fe)^2/fe$ which is $\sum(fo\text{-}fe)^{2/}fe$. This is the calculated chi-square.
vi) To make inferences, the calculated chi-square value is compared with obtained value from the table.

Unless otherwise given the expected frequencies, fe (the theoretical or known frequencies) for each groups or cells in a chi-square frequency table can be obtained or calculated as follows:

Goodness-of-fit: $f_e = np$, where n = sample size and p = proportion for each group within a category (cell).

In case of Independence and Homogeneity tests, the expected frequencies (f_e) can be calculated as illustrated by Table 1.

Now we need to calculate the expected values for each cell in the table and we can do that using the row total times the column total divided by the grand total (N). For example, for cell A the expected frequency (fe) would be (A+B+C)(A+D)/N

The following Alternative formula can be used to find the value of chi-square. It is not applicable uniformly in all cases but can be used only in a 2×2 contingency table which is one dimensional (two expected and two actual frequencies). If we set the 2×2 table to the general notation shown in Table 2, using the letters A, B, C, and D to denote the contents of the cells.

Then the formula for calculating the value of χ^2 will be stated as follows:

$$X^2 = N.\ (AD\text{-}BC)^2/(A+C)(B+D)(A+B)(C+D)$$

Where N is the total frequency (grand total), AD is the larger cross product, BC is the smaller cross product and (A+C), (B+D), (A+B), (C+D) are the marginal totals.

According to Kothari (2007), when some of the frequencies are less than 10, we first regroup the given frequencies until all of them all more than 10 and then work out the value of Chi-square χ^2. If the original (2×2) table has small frequencies (less than 10) or after regrouping up to 2×2 table some of the frequencies remain small the Yate's correction is applied. The rule of the Yate's correction is to adjust the observed frequency in each of a 2×2 table in such away that as to reduce the deviation of the observed from the expected frequency for that cell by 0.5 but this adjustment is made in all the cells without disturbing the marginal totals (Kothari, 2007). This reduces the Chi-square value obtained and thus increases its p-value (Sokal and Rohlf, 1981).

The effect of Yates's correction is to prevent over-estimation of statistical significance for small data. The sample size must be large enough to fairly represent the population from which it is drawn. At least 20 observations should be used, with at least five members in every individual category (Sokal and Rohlf, 1981). Unfortunately, Yates's correction may tend to overcorrect. This can result in an overly conservative result that fails to reject the null hypothesis when it should (a type II error).

When using the definitional formula, $\chi^2 = \sum (fo-fe)^2/fe$, then Yate's correction can be applied under:

$$\chi^2_{(corrected)} = \sum|fo-fe|-0.5|^2/fe.$$

After applying Yate's correction, the χ^2 can be calculated using the computational formula below:

$$\chi^2_{(corrected)} = N.|AD-BC-0.5\ N|^2/(A+C)(B+D)(A+B)(C+D)$$

3. Use the chi-square distribution table to determine significance of the value.

a. Determine degrees of freedom and locate the value in the appropriate column.

The degree of freedom (df) is a parameter used to look up Chi-square values from the Chi-square distribution table. It is related to the sample size or number of classification of data within a category.

If there is one independent variable, df = c - 1 where c is the number of levels of the independent variable.

If there are two independent variables, df = (c- l) (r - l) where c and r are the number of levels of the first and second independent variables, respectively.

If there are three independent variables, df = (c - l) (r - 1) (t - 1) where c, r, and t are the number of levels of the first, second, and third independent variables, respectively.

b. Locate the value closest to your calculated χ^2 on that degrees of freedom *df* row.

c. Move up the column to determine the p value

4. State your conclusion in terms of your hypothesis.

a. If the α value for the calculated χ^2 is $\alpha > 0.05$, accept your hypothesis. The deviation is small enough that chance alone accounts for it. A α value of 0.8, for example, means that there is a 80% probability that any deviation from expected is due to chance only. This is within the range of acceptable deviation.

b. If the α value for the calculated is $\alpha < 0.05$, reject your hypothesis, and conclude that some factor other than chance is operating for the deviation to be so great. For example, a α value of 0.10 means that there is only a 10% chance that this deviation is due to chance alone. Therefore, other factors must be involved.

Areas of application of the Chi-square test

Chi-square test is applicable in large number of problems. The main applications include test of frequencies (test the goodness of fit, test the homogeneity of a number of frequency distributions, test of independence) and test of population variance (test of single sample variance). The applications are discussed below and illustrated with numerical examples.

Test of goodness of fit

Goodness-of-fit test is a Chi-square test technique used to study similarities between proportions or frequencies between groupings (or classification) of categorical data. Mutai (2000) contends that goodness of fit is away of comparing empirically derived data (expected as frequencies) with theoretically expected results. In other words, a research situation may occur in which the experimenter is interested in similarity (or "goodness of fit") between the distribution of a sample of observations, and the distribution of cases that previous research or theory would suggest (Mutai, 2000). According to Kothari (2007), the test assists us to see how well the distribution of observed data fits the assumed theoretical distribution like the normal distribution, binomial distribution or the Poisson. If the calculated value of chi-square is less than the table value at a certain level of significance, the fit is considered to be a good one which means the divergence between the observed and expected frequencies is attributed to fluctuations of sampling. On the other hand, if the calculated value of chi-square is greater than its table value, the fit is not considered to be good one.

Illustration on the test of goodness of fit

Mathematics examination is composed of two papers, 1 and 2. Passing paper 1 is represented A while failing is represented by a and passing paper 2 is represented by B while failing is represented by b. The performance of students represented by AB, Ab, aB and ab are expected to occur in a 2:4:3:1 ratio. The actual frequencies were: AB=10, Ab=30, aB=54 and ab=26. Test the goodness of fit of the performance to the theory at 5% level of significance.

Table 3. Contingency table.

		Paper 1	
		Pass(A)	**Fail (a)**
Paper 2	Pass (B)	10(24)	54(36)
	Fail (b)	30(48)	26(12)

Table 4. Calculation of the Chi-square (χ^2).

Group	Observed frequency (f_0)	Expected frequency (fe)	Fo-fe	(fo –fe)^2	(fo –fe)^2/fe
AB	10	24	-14	196	8.167
Ab	30	48	-18	324	6.75
Ab	54	36	18	324	9
Ab	26	12	14	196	16.33
					χ^2=∑ (fo–fe)^2/fe=-4.726

Table 5. Calculation of the Chi-square (χ^2) value.

Group	Observed frequency (f_0)	Expected frequency (fe)	Fo-fe	(fo –fe)^2	(fo –fe)^2/fe
A	882	900	-18	324	0.360
B	313	300	13	169	0.563
C	287	300	-13	169	0.563
D	118	100	18	324	3.240
					χ^2=∑ (fo–fe)^2/fe=-4.726

Solution

The null hypothesis to be tested is there is no significant difference between the actual and the theoretical performances.

Since the total number of students is 120, then the expected frequencies are: AB=[2/10]×120=24; Ab=[4/10]×120= 48; aB = [3/10]×120= 36 and ab=[1/10]×120=12.

The information can be presented in a contingency table (Table 3).

Therefore degrees of freedom= (c-1) (r-1) = (2-1) (2-1) = 1 where c is the number of columns and r is the number of rows.

The value of χ^2 for one degree of freedom at 5% level of significance is 3.84 (Table 4). The calculated value of χ^2 is higher than the table value which means that the calculated value cannot be said to have arisen just because of chance. This means that there is significant difference between the actual and the theoretical performance.

Test of homogeneity

Test of homogeneity is a Chi-square technique used to study whether different populations are similar (or homogeneous or equal) in reference to some characteristic or attribute. Mutai (2000) notes that you may use a contingency table to find out if two sets of empirical data are alike- or, in statistical language, whether two sets are random samples from the same population. According to Kothari (2007), the test can also be used to explain whether the results worked out on the basis of sample or samples are consistent with the defined hypothesis or the results fail to support the given hypothesis.

Illustration on test of homogeneity

The theory predicts the proportion of beans, in the four groups A, B, C and D should be 9:3:3:1. In an experiment among 1600 beans, the number in the four groups were 882,313, 287 and118. Does the experiment support the theory? Apply χ^2 test (M.B.A., Delhi University, 1975; unpublished).

Solution

The null hypothesis is there is no significant difference between the experiment and the theoretical prediction.
The expected frequencies are:

A= [9/16] X 1600=900; B= [3/16] X1600=300; C= [3/16] X1600=300; and D = [1/16] X1600=100
The value of χ^2 for three degree of freedom at α=.05 level of significance is 7.82 (Table 5). The calculated value

Table 6. The resulting data.

		Major			
		Science	**Humanities**	**Arts**	**Professional**
Gender	Female	30	10	15	45
	Male	80	120	45	155

Table 7. Calculation of the marginal totals and the grand total.

		Major				
		Science	**Humanities**	**Arts**	**Professional**	**Total**
Gender	Female	30	10	15	45	100
	Male	80	120	45	155	400
	Total	110	130	60	200	500

(4.726) of χ^2 is smaller than the table value (7.82) hence there is no significant difference between the experimental data and theoretical prediction. We can thus conclude that the experiment support the theory.

Test of independence

A test of Independence is a chi-square technique used to determine whether two characteristics (such as food spoilage and refrigeration temperature) are related or independent. The test enables us to explain whether or not two attributes are associated. In other words, it is used to estimate the likelihood that some factor other than chance accounts for the observed relationship. Kothari (2007), chi-square is not a measure of degree of relationship or the form of relationship between two attributes but it is simply a technique of judging the significance of such association or relationship between two attributes. Koul (1984) notes that since the null hypothesis states that there is no relationship between the variables under study, the Chi-square test merely evaluates the probability that the observed relationship results from chance. If the calculated chi-square value is less than its table value at a certain level of significance for a given degree of freedom, then we conclude that the two attributes are independent or not associated. But if the calculated chi-square value is greater than its table value at a certain level of significance for a given degree of freedom, then we conclude that the two attributes are associated.

Illustration on test of independence

A researcher would like to know if there is a relationship between a student's gender and his or her choice of college major at α=.05. To test this hypothesis, a sample of 500 students is selected, and each person's gender and college major is recorded. The resulting data are given in Table 6.

a) What is the value of degrees of freedom for this design?
b) Calculate the marginal totals and the grand total.
c) Calculate the expected frequency for each cell.
d) Calculate χ^2_{obs}
e) Use the table of the χ^2 distribution which is provided to test if there is a relationship between the two variables at α=.05 (PhD in Educational Planning and Management 2010 Examination, Masinde Muliro University of Science and Technology, Kenya, Unpublished).

Solution

a) Degrees of freedom= (c-1)(r-1)=(4-1)(2-1)=3
b) The marginal totals and the grand total can calculated as shown in Table 7.
c) Calculation of expected frequency for each cell=product of corresponding marginal totals/Grand total

Female science (FS) = (100X110)/500=22
Female humanities (FH) = (130X100)/500=26
Female arts (FA) = (60X100)/500=12
Female professional (FP) = (200X100)/500=40
Male science (FS) = (110X400)/500=88
Male humanities (MH) = (130X400)/500=104
Male humanities (MH) = (60X400)/500=48
Male professional (MP) = (200X400)/500=160

d) The null hypothesis is there is no significant relationship between a student's gender and his or her choice of college major.

The table value of χ^2 for three degrees of freedom at α=.05 level of significance is 7.82 (Table 8). The calculated value of χ^2 is higher than the table value hence we reject the null hypothesis. We can thus conclude that

Table 8. Calculation of Chi-square observed (χ^2_{obs}).

Group	Observed frequency (f_0)	Expected frequency (fe)	Fo-fe	$(fo - fe)^2$	$(fo - fe)^2/fe$
FS	30	22	8	64	2.909
FH	10	26	-16	256	29.846
FA	15	12	3	9	0.75
FP	45	40	5	25	0.625
MS	80	88	-8	64	0.727
MH	120	104	16	256	2.461
MA	45	48	-3	9	1.688
MP	155	160	-5	25	0.156
					$\chi^2 = \sum (fo - fe)^2/fe = 4.726$

Table 9. Sample of 140 observations classified by two attributes A and B.

	A_1	A_2	A_3	Total
B_1	35	30	25	90
B_2	34	3	3	40
B_3	7	2	1	10
Total	76	35	29	140

Table 10. Regrouping data.

	A_1	A_2 & A_3	Total
B_1	35	45	90
B_2 & B_3	41	9	50
Total	76	54	140

there is significant relationship between a student's gender and his or her choice of college major.

Illustration

A sample of 140 observations classified by two attributes A and B is shown in Table 9.
Use the χ^2 test examine whether A and B are associated at α=.05. (Master of Education 2006 Examination, Kenyatta University, Kenya, Unpublished)

Solution

Taking the hypothesis that there is no significant association between attributes A and B, some of the frequencies are less than 10; we shall first regroup the given data as shown in Table 10.

Since the frequency of 9 is still less than 10, we apply Yate's correction as follows. The expectations of the attributes are calculated as:

A_1B_1=[76X90]/140=48.86; A_2&A_3B_1 = [54X90]/140=34.71; A_1B_2&B_3=[76X50]/140=27.14; A_2&A_3 B_2&B_3=[54X50]/140=19.28.

The value of χ^2 for one degree of freedom at α=.05 level of significance is 3.84 (Table 11). The calculated value of χ^2 is higher than the table value hence we reject the null hypothesis. We can thus conclude that there is significant association between attributes A and B.

Test of population variance

The chi-square test can be used to judge if a random sample has been drawn from a normal population with mean (μ) and with specified variance (σ^2). However, when testing for variance, the assumption that needs to be met is that the populations involved are approximately normal (Ingula and Gatimu, 1996). The chi-square is used to test for a single sample variance. The Chi-square test statistic for the null hypothesis is calculated by the formula:

X^2 = ∑(Xi-X)2σ2 = (N-1)s2σ2

where X_i is the variable, X is the sample mean, σ is the population standard deviation and s is sample standard deviation as in inferential statistics. The degrees of freedom are (N-1).

Illustration on test of population variance

The height of 8 students in centimetres was given as 65, 68, 67, 76, 72, 78, 69, and 81. Test the hypothesis that the standard deviation of the population from which this students were selected from is equal to 5 at 5% level of significance (B.ED 2011 Examination, Bugema University-Nyamira, Kenya, Unpublished).

Solution

X=∑Xi/N=576/8 = 72; $X^2 = \sum(X_i-\bar{X})^2/\sigma^2 = 232/5^2 = 9.28$

Table 11. Calculation of the Chi-square (χ^2) value.

Group	Observed frequency (fo)	Expected frequency (fe)	$\|f0-fe\|-0.5\|^2$	$\|fo-fe\|-0.5\|^2/fe$
A_1B_1	35	48.86	178.49	3.65
A_2&A_3B_1	45	34.71	95.84	2.76
A_1B_2&B_3	41	27.14	178.49	6.58
A_1B_2&B_3	9	19.28	95.65	4.96
				$\chi^2_{(corrected)}$=17.95

Table 12. Calculation of the sum of squared deviations from the mean, $\sum (X_i-\bar{X})^2$

Student	Height(X_i)	$X_i-\bar{X}$	$(X_i-\bar{X})^2$
1	65	-7	49
2	68	-4	16
3	67	-5	25
4	76	4	16
5	72	0	0
6	78	6	36
7	69	3	9
8	81	9	81
	$\sum$ Xi=576		$\Sigma(X_i-\bar{X})^2$ =232

The degrees of freedom are (N-1)=8-1=7

At 5% level of significance the table value of χ^2 is 14.07 which is greater than the calculated value of χ^2 which is 9.28 (Table 12). Therefore we do not reject the hypothesis and conclude that the standard deviation of the population from which this students were selected from is equal to 20 at 5% level of significance.

THE CONCEPTUAL MODEL

From the earlier discussion, it should be noted that the chi-square test is to be applied in testing of frequencies (goodness of fit, homogeneity and independence) and test of population variance. This is against other authors like Kothari (2007) who contended that the chi-square is based on frequencies and not on the parameters like mean and standard deviation. By the use of standard deviation in testing population variances it means that the chi-square test involves variables expressed as interval or ratio.

Testing of frequencies is appropriate when the population is not known to be normal, sample size is small and the variables expressed as nominal or ordinal. The conditions for the application of chi-square in testing frequencies involves observations recorded or used are collected on a random basis, all the members (or items) in the sample must be independent, the overall number of items must be reasonably large, the constraints must be linear and no group should contain very few items that is where the frequencies are less than 10, regrouping is done by combining the frequencies of adjacent groups so that the new frequency is greater than 10. The Yates correction formula is used in case of 2×2 table and some frequencies below 10.

In the test of population variance, the chi-square is used to test for a single sample variance which involves judging if a random sample has been drawn from a normal population with mean and with specified variance. However, when testing for variance, the assumption that needs to be met is that the populations involved are approximately normal.

In both the test of frequencies and population variances, the degrees of freedom and significance levels are used to determine the table value. In the interpretation of results, if the calculated chi-square value is less than its table value at a certain level of significance for a given degree of freedom, then we conclude that the two attributes are independent or not associated. But if the calculated chi-square value is greater than its table value at a certain level of significance for a given degree of freedom, then we conclude that the two attributes are associated.

The above information is presented schematically as shown in Figure 1. In the model the computational (alternative) formula has been excluded because it is not applicable uniformly in all cases but can be used only in a 2×2 contingency table which is one dimensional (two actual and two expected frequencies). In this case the definitional formula has been emphasized.

CONCLUSIONS AND RECOMMENDATIONS

This research highlighted the definition, procedure and applications of Chi-square test. By the use of examples, this research illustrated in details the application of the test on homogeneity, goodness of fit, independence and population variance. To understand the Chi-square test the researcher developed a simplified conceptual model which can be applied in most cases. To ensure appropriate use of the test in research, the researchers must understand the concept of Chi-square test manually before using the computer for analysis; consult a variety of books and experts in the area; and try use the definitional formula which is more convenient than the

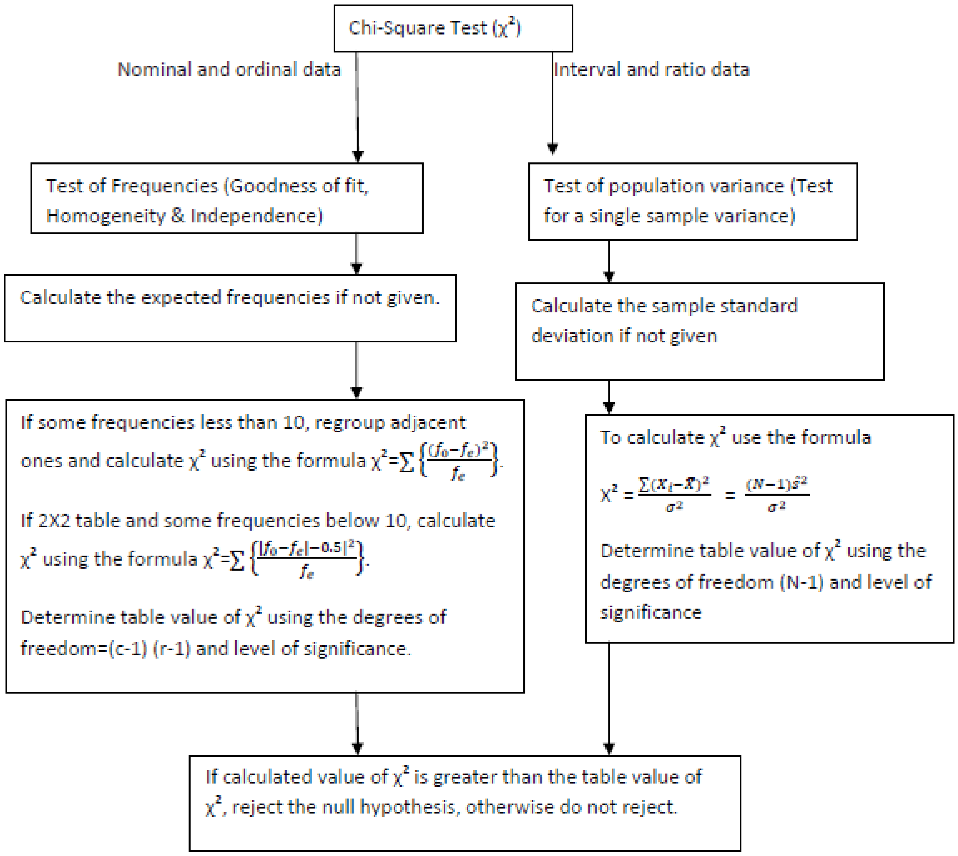

Figure 1. Conceptual model of Chi-square test.

Table 13. Chi-square distribution table.

Df	Level of significance				
	.10	.05	.02	.01	.001
1	2.71	3.84	5.41	6.64	10.83
2	4.60	5.99	7.82	9.21	13.82
3	6.25	7.82	9.84	11.34	16.27
4	7.78	9.49	11.67	13.28	18.46
5	9.24	11.07	13.39	15.09	20.52
6	10.64	12.59	15.03	16.81	22.46
7	12.02	14.07	16.62	18.48	24.32
8	13.36	15.51	18.17	20.09	26.12
9	14.68	16.92	19.68	21.67	27.88
10	15.99	18.31	21.16	23.21	29.59
11	17.28	19.68	22.62	24.72	31.26
12	18.55	21.03	24.05	26.22	32.91
13	19.81	22.36	25.47	27.69	34.53
14	21.06	23.68	26.87	29.14	36.12
15	22.31	25.00	28.26	30.58	37.70
16	23.54	26.30	29.63	32.00	39.25
17	24.77	27.59	31.00	33.41	40.79
18	25.99	28.87	32.35	34.80	42.31
19	27.20	30.14	33.69	36.19	43.82

Table 13. Contd.

20	28.41	31.41	35.02	37.57	45.32
21	29.62	32.67	36.34	38.93	46.80
22	30.81	33.92	37.66	40.29	48.27
23	32.01	35.17	38.97	41.64	49.73
24	33.20	36.42	40.27	42.98	51.18
25	34.38	37.65	41.57	44.31	52.62
26	35.56	38.88	42.86	45.64	54.05
27	36.74	40.11	44.14	46.96	55.48
28	37.92	41.34	45.42	48.28	56.89
29	39.09	42.56	46.69	49.59	58.30
30	40.26	43.77	47.96	50.89	59.70

Source: Fisher and Yates (1974).

alternative formula.

For χ^2 obtained from the data to be significant it must be equal to or larger than the value shown in Table 13.

From Table 13 for degrees of freedom greater than 30, the quantity 2x2-2v-1 may be used as a normal variate with unit variance.

REFERENCES

Best JW (1977). Research in education. New Delhi, Prentice-Hall of India Private limited.

Brown JD (2004). Yates correction factor. SHIKEN, The JALT Testing & Evaluation SIG Newsletter 8(1):22-27).

B.ED (2011). Examination, Bugema University-Nyamira, Kenya, Unpublished).

Fisher RA, Yates F (1974). Statistical tables for biological, agricultural and medical research, Sixth edition, Addison Wesley Longman Ltd.

Gay LR (1976). Educational research: Competencies for analysis and application. Ohio, Bell & Howell Company.

Ingula F, Gatimu H (1996). Essential of educational statistics. Nairobi, East African Educational Publishers.

Kothari CR (2007). Quantitative techniques. New Delhi, UBS Publishers LTD.

Koul L (1984). Methodologies of educational research. New Delhi, Vikas.

M.B.A. (1975). Delhi University, 1975; unpublished).

Mutai BK (2000). How to write quality research proposal: A complete and simplified recipe. New Delhi, Thelley Publications.

Onyango JO, Odebero SO (2009). Data analysis for educational and social research: Quantitative and qualitative approaches. Kenya J. Educ. Planning, Econ. Management. Vol.1 (1): 44-53.

Orodho JA (2005). Elements of education and social science research methods. Nairobi, Masola Publishers.

PHD in Educational Planning and Management (2010). Examination, Masinde Muliro University of Science and Technology, Kenya, Unpublished).

Sokal RR, Rohlf FJ (1981). Biometry: The Principles and Practice of Statistics in Biological Research. Oxford: W.H. Freeman, ISBN 0-7167-1254-1257.

Yates F (1934). Contingency table involving small numbers and the χ^2 test: Supplement J.Royal Statistical Soc. 1(2):217-235.