IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 8, August 2015.
www.ijiset.com
ISSN 2348 – 7968

Multi-Level Test of Independence for 2 X 2 Contingency Table using Cochran and Mantel–Haenszel Statistics

Adetunji, A. A.
Department of Mathematics & Statistics,
Federal Polytechnic,
Ado-Ekiti, Nigeria.
adecap4u@yahoo.com

Jemilohun V. G.
Department of Business Administration,
Afe Babalola University,
Ado-Ekiti, Nigeria
jemvinga@yahoo.com

Adaraniwon, A. O.
Department of Mathematics & Statistics,
Federal Polytechnic,
Ado-Ekiti, Nigeria.
lekanamoz@gmail.com

Abstract
Cochran (1954) introduced a test statistic to extend the chi-square test of independence in a 2 X 2 table to multiple 2 X 2 tables where each table corresponds to a different level of an intervening variable. He proposed a test of conditional independence of the variables forming the rows and columns of the tables, conditional on the levels of a third variable. Results of graduates from ten randomly selected departments from the federal Polytechnic, Ado-Ekiti were examined in this research. Students final grade are categorised in to two: (i) Low Grade (Pass and Lower Credit) and (ii) High Grade (Upper Credit and Distinction). This is combined with sex (male and female) to form 2 x 2 x h contingency table. Breslow-Day and Tarone's statistics show that the null hypothesis of homogeneity of odds ratio across the departments is not rejected implying that the odds ratio across the ten departments is not significantly different. Cochran's and Mantel-Haenszel statistics reveals the final grade of students (Low or High) is not associated with sex.

Keywords: Cochran-Mantel-Haenszel, Breslow-Day, Tarone, Tests of repeated measure, Chi-Square test of Independence, Students' Performance

INTRODUCTION
The development of methods for categorical variables was stimulated by research studies in the social and biomedical sciences. Categorical scales are pervasive in the social sciences for measuring attitudes and opinions. Although categorical data are common in the social and biomedical sciences, they are by no means restricted to those areas. They frequently occur in the behavioural sciences (e.g., type of mental illness, with the categories schizophrenia, depression, neurosis), epidemiology and public health (e.g., contraceptive method at last intercourse, with the categories none, condom, pill, IUD, other), genetics (type of allele inherited by an offspring), zoology (e.g., alligators' primary food preference, with the categories fish, invertebrate, reptile), education (e.g., student responses to an exam question, with the categories correct and incorrect), and marketing (e.g., consumer's preference among leading brands of a product, with the categories brand A, brand B, and brand C). They even occur in highly quantitative fields such as engineering sciences and industrial quality control. Examples are the classification of items according to whether they conform to certain standards, and subjective evaluation of some characteristic: how soft to the touch a certain fabric is, how good a particular food product tastes, or how easy to perform a worker finds a certain task to be.
Research data can easily be clumped into categories. Virtually every research project categorizes some of its observations into various categories: sex (male or female); students mode of entry into higher institution (UME, REM, DE); class of degree (Pass, Third Class, Second Class Lower, Second Class Upper, First Class), and so on. When we collect data by categories, we record counts, that is, how many observations fall into a particular bin. Categorical variables are usually classified as being of two basic types: nominal and ordinal. Nominal variables involve categories

ISSN 2348 – 7968

that have no particular order such as sex while the categories associated with an ordinal variable have some inherent ordering (class of degree).

Categorical variable have a measurement scales consisting of a set of categories. They are variables that can be measured using a limited number of values or categories (Daniel and Yu, 1999). This distinguishes categorical variable from continuous variables, however, continuous variables are sometimes treated as categorical. When a continuous variable is treated as a categorical variable, it is called CATEGORIZATION of the continuous variable. In this research work, final grade of students is categorized, although continuous, grade is often treated as categorical in actual research for substantive and practical reasons.

Unlike methods for continuous variables, methods for analyzing categorical data require close attention to the type of measurement of the dependent variable. Methods for analyzing one type of categorical dependent variable may be inappropriate for analyzing another type of variable.

RESEARCH METHODOLOGY

Contingency Table Analysis ($r \times c$)

Contingency table analysis is a common method of analyzing the association between two categorical variables. Considering a categorical variable that has r possible response categories and another categorical variable with c possible categories. In this case, there are $r \times c$ possible combinations of responses for these two variables. The $r \times c$ cross-tabulation or contingency table has r rows and c columns consisting of $r \times c$ cells containing the observed counts (frequencies) for each of the $r \times c$ combinations. This type of analysis is called a contingency table analysis and is usually accomplished using a chi-square statistic that compares the observed counts with those that would be expected if there were no association between the two variables.

Combining Several 2 x 2 Tables

Often, associations between two categorical variables are examined across two or more populations. The resulting data usually lead to several (say, H) 2 x 2 contingency tables. Categorising final grade of student at graduation in to *High Grade* (Distinction and Upper Credit) and *Low Grade* (Pass and Lower Credit) based on the result obtained for the data from various departments of a faculty. These tables can sometime come from a single study that has been stratified by some factor (in this case, department) that might be a confounder. The goal is usually to be able to combine the tables in order to have unified information across the tables. We would like to combine the evidence from the all the departments in the School of Science and Computer Studies to make an overall statement about whether final grade is independent of sex. If the conditional test for these data within each department can be obtained by computing Fisher's exact test separately for each sub-table or obtain Pearson's χ^2 for each table and the computed values suggest that neither of them is significant (which suggests that final grade is independent of sex) within each department. The results of Pearson's χ^2 and Fisher's exact test usually suggest that it would not be wise to collapse the data set over the factor variable department without serious distortion to the association between the two variables been considered (Lawal, 2003).

In general, we are interested in collecting information for each of several 2 x 2 tables across the levels of the subpopulations (which may be determined by various configurations of factor variables or covariates).

In many cases, the primary question involves the relationship between an independent variable (factor) that is either present or absent and a dependent (response) variable that is either present or absent in the presence of several covariates. This could give rise to frequency data that may be summarized as a set of 2 x 2 tables.

	Response		**Total**
Factor	n_{h11}	n_{h12}	$n_{h1.}$
	n_{h21}	n_{h22}	$n_{h2.}$
Total	$n_{h.1}$	$n_{h.2}$	$n_{h...}$

Cochran (1954) introduced a test statistic to extend the chi-square test of independence in a 2 X 2 table to multiple 2 X 2 tables where each table corresponds to a different level of an intervening variable. Cochran proposed a test of conditional independence of the variables forming the rows and columns of the tables, conditional on the levels of a third variable.

To establish notation, let n_{hij} represents the number of responses observed at the i^{th} level of the row variable, the j^{th} level of the column variable, and the h^{th} level of the intervening variable. Assuming H levels of the intervening variable $i = 2$ rows, and $j = 2$ columns, we have data that may be summarized as in the table a below for $h = 1, \ldots, H$.

	Factor B		**Total**
Factor A	n_{h11}	n_{h12}	$n_{h1.}$
	n_{h21}	n_{h22}	$n_{h2.}$
Total	$n_{h.1}$	$n_{h.2}$	$n_{h...}$

For this situation Cochran conditioned on the row totals, considering each 2 X 2 table to consist of independent binomials. He based his statistic on a weighted sum of the table-specific differences in proportions; $d_w = \sum_{h=1}^{H} w_h(\hat{p}_{h1} - \hat{p}_{h2})$ where $w_h = \frac{n_{h1.} * n_{h2.}}{n_{h..}}$ and $\hat{p}_{hi} = \frac{n_{hi1}}{n_{hi.}}, i = 1, 2$

Using the asymptotic normality of d_w, he justified that;

$\chi_C^2 = \frac{d_w^2}{\widehat{var}(d_w)} = \frac{d_w^2}{\sum_{h=1}^{H} \frac{(n_{h1.} * n_{h2.} * n_{h.1} * n_{h.2})}{(n_{h..}^3)}}$ as an appropriate test statistic, having an approximate $\chi^2(1)$ distribution under the *null hypothesis of conditional independence.*

Mantel and Haenszel (1959) proposed a similar test statistic using a hypergeometric assumption. Conditional on the row and column totals, the cell counts in each table have a hypergeometric distribution. This fact suggests a test statistic based on the difference between the observed and expected frequencies in each 2 X 2 table. As with the classical chi-square test of independence in a single 2 X 2 table, it suffices to compare the observed and expected count in one cell per table.

In particular, let n_{h11} be the "pivot" cell frequency of subjects in the i^{-th} table who have both factor and response present. Under the assumption that the marginal totals are fixed, the overall null hypothesis of ***no partial association*** against the alternative hypothesis that ***on the average across the h sub-tables, there is a consistent relationship between the row and column variables*** is conducted by obtaining the Cochran-Mantel-Haenszel (CMH) test statistic χ_{MH}^2, which is computed as follows:

For table i, for instance, $n_{h..} = (n_{h11}, n_{h12}, n_{h21}, n_{h22})$ follows the hypergeometric distribution and therefore; $P\{n_{h..}|H_0\} = \left(\frac{n_{h1.}! * n_{h2.}! * n_{h.1}! * n_{h.2}!}{n_{h..}! * n_{h11}! * n_{h12}! * n_{h21}! * n_{h22}!}\right)$

Hence, it follows that the expected value for the pivot cell in the *i-th* sub-table is given by:

ISSN 2348 – 7968

$$E(n_{h11}|H_0) = \left(\frac{n_{h1.*}n_{h.1}}{n_{h..}}\right)$$

$$Var(n_{h11}|H_0) = \left(\frac{n_{h1.*}n_{h.1} * n_{h2.*}n_{h.2}}{n_{h..}^2(n_{h..} - 1)}\right)$$

The Mantel-Haenszel test is, therefore;

$$\chi^2_{MH} = \frac{\left\{\left|\sum_{h=1}^{H} n_{h11} - \left(\frac{n_{h1.} * n_{h.1}}{n_{h..}}\right)\right| - 0.5\right\}^2}{\sum_{h=1}^{H} \frac{(n_{h1.} * n_{h.1} * n_{h2.} * n_{n.2})}{n_{h..}^3(n_{h..}^2 - 1)}} = \frac{\{|d_w| - 0.5\}^2}{\sum_{h=1}^{H} \frac{(n_{h1.} * n_{h.1} * n_{h2.} * n_{n.2})}{n_{h..}^3(n_{h..}^2 - 1)}}$$

Aside from the continuity correction in χ^2_{MH}, the two test statistics _differ by a factor of_ $\left(\frac{n_{h..}}{n_{h..} - 1}\right)$ in each table. For moderate to large sample sizes per table, the difference between the two statistics is typically negligible.

In general, though χ^2_{MH} offers advantages. Both statistics are asymptotically $\chi^2(1)$, but the quality of this approximation depends upon the table-specific sample sizes only for χ^2_C. In the extreme, Mantel and Haenszel's statistic will perform adequately in matched pair studies in which $n_{h..} = 2$ for all h, for sufficient total sample size. Cochran's statistic is not be used for such a situation (Daniel et al, 2000).

In addition, the Mantel-Haenszel test has been shown to be optimal under the assumption of a constant odds ratio across tables (Birch, 1964) and it is asymptotically equivalent to likelihood ratio tests from unconditional and conditional logistic regression models for large strata and sparse data situations, respectively (Breslow and Day, 1980).

In addition to the test statistic, χ^2_{MH}, Mantel and Haenszel proposed an odds ratio estimator in their original 1959 paper. Their estimator is a weighted average of the table-specific observed odds ratios:

$$\hat{\psi}_{MH} = \frac{\sum_{h=1}^{H} v_h \hat{\psi}_h}{\sum_{h=1}^{H} v_h} = \frac{\sum_{h=1}^{H} R_h}{\sum_{h=1}^{H} S_h}$$

where $v_h = \frac{n_{h12*}n_{h21}}{n_{h..}}$; $\hat{\psi}_h = \frac{n_{h11*}n_{h22}}{n_{h12*}n_{h21}}$; $R_h = \frac{n_{h11*}n_{h22}}{n_{h..}}$ and $S_h = \frac{n_{h12*}n_{h21}}{n_{h..}}$

Since the introduction of χ^2_C, χ^2_{MH}, and $\hat{\psi}_{MH}$, a large literature has developed. Mantel-Haenszel-type estimators have been developed for the rate ratio (Rothman and Boice, 1979), rate difference (Greenland, 1982), risk ratio (Rothman and Boice, 1979; Tarone, 1981; Nurminen, 1981; and Kleinbaum et al., 1982), and risk difference (Greenland, 1982).

Hypothesis Statement

H_0: Controlling for (or within) departments, there is no relationship between factor (sex) and response (final grade).

H_a: Controlling for (or within) departments, there is relationship between factor (sex) and response (final grade).

Odds Ratio: The Mantel-Haenszel Method

The odds of an event (or condition) is defined by $\frac{\pi}{1-\pi}$, where π is the probability of the event. The odds ratio Ψ is the ratio of the odds of an event occurring in one group to the odds of that event in another group. These groups might be sex (male and female), students' grade (High or Low), or any other dichotomous classification. The odds ratio is used to test whether the probability of a certain event is the same for two groups. We note that the odds ratio takes values in (0,∞). An odds ratio of 1 indicates that the event under study is equally likely in both groups. If Ψ > 1, then the event is more likely in the first group, whereas Ψ < 1 indicates that it is less likely. The 2 × 2 table shows observations for two such groups and events A and A¯, the complement of A.

ISSN 2348 – 7968

	A	$\bar{A}$	Totals
Group 1 Group 2	X_1 X_2	$n_1 - X_1$ $n_2 - X_2$	n_1 n_2
Totals	$X_1 + X_2$	$n_1 + n_2 - X_1 - X_2$	$n_1 + n_2$

The odds ratio $\Psi = \frac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)}$ is estimated by $\frac{X_1(n_2-X_2)}{X_2(n_1-X_1)}$, which is invariant if rows or columns (or both simultaneously) are interchanged.

In clinical studies there are often only a few subjects. Multicentre trials increase the sample size, but populations differ for different centres and one cannot assume that probabilities for different centres are equal. However, one can assume that the odds ratios for each of the ***K*** centres are identical, that is, assuming a *common odds ratio* Ψ with $\Psi = \Psi_1 = \ldots = \Psi_k$. Under this *common odds ratio assumption*, the MH (Mantel-Haenszel, 1959*)* estimator $\widehat{\Psi}$ of the common odds ratio is widely used by practising statisticians and epidemiologists. The MH estimator is a ratio of two sums C_{12} and C_{21}, where each summand of C_{ij} has the form $\frac{X_{ik}(n_{jk}-X_{jk})}{(n_{ik}+n_{jk})}$ with index ***k*** referring to the quantities of the ***k***th table or ***k***th centre. The factor $\frac{1}{(n_{ik}+n_{jk})}$ is a weight accounting for the sample size of the ***k***th table. The MH estimator is also often applied for other stratified data for which the common odds ratio assumption is reasonable.

Even if the assumption of a common odds ratio is slightly violated, the MH estimator is still a useful tool to summarise the association across tables. Despite the MH estimator's simplicity, it has some useful properties. First, it applies to very sparse data. More precisely, it is defined when only one summand of C_{12} and of C_{21} is non-zero (Suesse, 2009).

It is also *dually consistent*, that is, consistent under two types of asymptotic models:

(i) when the sample size of each stratum increases and the number of strata is fixed, and

(ii) when the number of observations becomes large with the number of strata, while the sample size of each stratum remains fixed. (i) is referred to as a *large-stratum* limiting model, or model (i), and to (ii) as a *sparse data* limiting model, or model (ii).

In practice, model (i) represents large $n_{1k} + n_{2k}$ for each stratum and model (ii) represents large ***K***. The MH estimator is robust under any such extreme data. The consistency of the MH estimator for model (i) was shown by Gart (1962) and for model (ii) by Breslow (1981). Hauck (1979) derived the limiting variance of the MH estimator under model (i), whereas Breslow (1981) derived two asymptotic variances under model (ii): one based on the conditional distribution of the observations for each table given the marginal totals, and the other on the empirical variance. Applying either of the variance estimators depending on the given data, whether the data resembles the sparse data or large stratum case, is very unsatisfactory. Breslow and Liang (1982) proposed a weighted average of the two variance estimators to account for the two different limiting models. Robins et al. (1986) proposed a variance estimator which is dually consistent under models (i) and (ii) based on the unconditional distribution of the data.

An alternative way to estimate the common odds ratio for $K \times 2 \times 2$ tables is to fit an ordinary logit model with main effects and no interaction, where the ***K*** strata and one binary classification are treated as factors and the other binary classification as a response. The corresponding loglinear model is a model with no three-way interaction among rows, columns and strata. However, the unconditional maximum likelihood (ML) estimator is a poor estimator, because under model (ii) the nuisance parameters grow as the sample size grows. For instance when each table consists of a single matched pair, then the unconditional ML estimator of the common odds ratio converges to the square of the true common odds ratio (Anderson 1980, p.244). The nuisance parameters can be eliminated by conditioning on the margins of the 2×2 contingency table. The ML estimator based on the conditional distribution, which is non-central hypergeometric in each stratum, is also dually consistent. As a by-product, the ML fitting yields a variance estimator of the odds ratio estimator.

If the assumption of a common odds ratio fails, we can still use the MH estimate as a summary of the odds ratios among the strata. Without the common odds ratio assumption, the MH estimator is consistent under model (i) only; and appropriate standard errors were suggested by Guilbaud (1983), since the dually consistent variance estimator of Robins et al. (1986) fails.

A simple way to test the homogeneity of the odds ratio across strata is to apply a goodness-of-fit test to a logit model with only main effects and no interaction. The goodness-of-fit test statistic has K − 1 degrees of freedom (df) if the model holds.

Breslow-Day Test for Homogeneity of the Odds Ratios

Breslow and Day (1980) developed a test statistic which does not require model fitting and focuses directly on the potential lack of homogeneity. The Breslow-Day test statistic sums the squared deviations of observed and fitted values each standardised by its variance. According to Breslow and Day (1980) the test is used for stratified analysis of 2 ×2 tables to test the null hypothesis that the odds ratios for the *k*-strata are all equal. When the null hypothesis is true, the statistic has an asymptotic chi-square distribution with *k*-1 degrees of freedom.

Tarone (1985) proved that it is stochastically larger under the homogeneity assumption, and developed a modified score test statistic that is indeed asymptotically $\chi^2_{(K-1)}$. Liang and Self (1985) proposed a score test assuming the log odds ratios across strata are independent and identically distributed, which is valid also when the sample size increases with the number of strata. Paul and Donner (1989) conducted a simulation study generally recommending Tarone's modified test statistic. Liu and Pierce (1993) used a different approach by assuming that the log odds ratios across the strata are a sample from a population with unknown mean and variance. They investigated the conditional likelihood functions for the mean and the variance. A test of homogeneity of the odds ratios can be conducted by testing whether the variance of the log odds ratio equals zero. Liu and Pierce (1993)'s approach is more general than that of Liang and Self (1985), because it describes the heterogeneity of the log odds ratios across the strata.

The estimation of the common odds ratio assumes that the strength of association as measured by the odds ratios in each sub-table is the same. This assumption is tested by the test of homogeneity of the odds ratio. To test this hypothesis, the Breslow-Day test is often employed. This statistic is compared to a standard X^2 distribution with $(H - 1)$ degrees of freedom. The null hypothesis of homogeneity of odds ratio across the sub-tables is rejected if P-value is < α. The Breslow-Day statistic is computed as:

$$\hat{\theta}_{BD} = \frac{\sum_{h=1}^{H}\left(n_{h11} - E\left(n_{h11}|\hat{\theta}_{MH}\right)\right)^2}{var\left(n_{h11}|\hat{\theta}_{MH}\right)}$$

where *E* and *var* denote expected value and variance, respectively. *The summation does not include any table with a zero row or column.*

It is advisable to test the homogeneity of the odds ratios in the different repeats, and if different repeats show significantly different odds ratios, the Cochran–Mantel–Haenszel test is not essential.

NOTE: *Unlike the Cochran-Mantel-Haenszel statistics, the Breslow-Day test requires a large sample size within each stratum, and this limits its usefulness. In addition, the validity of the CMH tests does not depend on any assumption of homogeneity of the odds ratios; therefore, the Breslow-Day test is not to be used as such an indicator of validity.*

Estimating the Common Odds Ratio

While the Cochran-Mantel-Haenszel test provides the significance of the relationship between two variables (sex and final grade) across the sub-tables (department), it does not tell us the strength of this association. An estimator (The Mantel-Haenszel estimate of the common odds ratio in case-control studies) of the common odds ratio is given *by:*

$$\hat{\theta}_{MH} = \frac{\sum_{h=1}^{H}\left(\frac{n_{h11} * n_{h22}}{n_{h..}}\right)}{\sum_{h=1}^{H}\left(\frac{n_{h12} * n_{h21}}{n_{h..}}\right)}$$

If the confidence interval of the common odds does not include 1; then, we conclude that there is dependence on the two variables of interest across the sub-tables else.

The estimate of the common odds ratio is based on the assumption that the strength of the association is the same in each department. If this is not the case, then we would have believed that there is *interaction* or *effect modification* between department and final grade. The factor variable (department) is often referred to as the *effect modifier*.

Using the estimated variance for $\log(\hat{\theta}_{MH})$ given by Robins et al (1986), we can compute the corresponding 100(1 – α)% confidence limits for the odds ratio as:

$$\left(\hat{\theta}_{MH} * \exp(-z\hat{\sigma}), \hat{\theta}_{MH} * \exp(z\hat{\sigma})\right)$$

where

$$\hat{\sigma}^2 = var\left[\ln\hat{\theta}_{MH}\right] = \frac{\frac{\sum_{h=1}^{H}(n_{h11}+n_{h22})(n_{h11}n_{h22})}{n_{h..}^2}}{2\left(\frac{\sum_{h=1}^{H} n_{h11}n_{h22}}{n_{h..}}\right)^2} + \frac{\frac{\sum_{h=1}^{H}(n_{h12}+n_{h21})(n_{h12}n_{h21})}{n_{h..}^2}}{2\left(\frac{\sum_{h=1}^{H} n_{h12}n_{h21}}{n_{h..}}\right)^2}$$

$$+ \frac{\frac{\sum_{h=1}^{H}(n_{h11}+n_{h22})(n_{h12}n_{h21}) + (n_{h12}+n_{h21})(n_{h11}n_{h22})}{n_{h..}^2}}{2\left(\frac{\sum_{h=1}^{H} n_{h11}n_{h22}}{n_{h..}}\right)\left(\frac{\sum_{h=1}^{H} n_{h12}n_{h21}}{n_{h..}}\right)}$$

MATERIALS

The data used for this research work is obtained from academic records of graduated students from ten randomly selected departments of The Federal Polytechnic, Ado-Ekiti, Nigeria for 2012/2013 academic session.

DATA ANALYSIS AND RESULTS

Sex * Final Grade CMH * Department Analysis

Table 1: Tests of Homogeneity of the Odds Ratio

	Chi-Squared	df	Asymp. Sig. (2-sided)
Breslow-Day	6.767	9	.661
Tarone's	6.767	9	.661

Both Breslow-Day and Tarone's statistics show that the null hypothesis of homogeneity of odds ratio across the departments is not rejected since P-value (0.661) > α (0.05). This implies that the odds ratio across the ten departments (relating to sex and final grade) are all equal.

Table 2: Tests of Conditional Independence (sex)

	Chi-Squared	df	Asymp. Sig. (2-sided)
Cochran's	4.671	1	.031
Mantel-Haenszel	4.214	1	.040

Hypothesis Statement

H_0: Controlling for (or within departments), there is no relationship between gender and final grade (Low Grade or High Grade).
H_a: Controlling for (or within departments), there is relationship between gender and final grade (Low Grade or High Grade).
The P-values (0.031 and 0.040) for both Cochran's and Mantel-Haenszel statistics reveals that the null hypothesis of no association of final grade and sex of students among the randomly selected ten departments of Federal Polytechnic, Ado-Ekiti is not rejected. Hence, we conclude that the final grade of students (High Grade or Low Grade) is not associated with sex of students.

Table 3: Mantel-Haenszel Common Odds Ratio Estimate (sex)

Estimate			.670
ln(Estimate)			-.401
Std. Error of ln(Estimate)			.187
Asymp. Sig. (2-sided)			.032
Asymp. 95% Confidence Interval	Common Odds Ratio	Lower Bound	.465
		Upper Bound	.966
	ln(Common Odds Ratio)	Lower Bound	-.767
		Upper Bound	-.035

The odd in favour of male students graduating with Low Grade (Pass and Lower Credit) is 0.670. The 95% confidence interval for this common odds ratio is (0.465, 0.966). This interval does not include 1; therefore, there is dependence on sex of the students and final grade in two categories (High Grade or Low Grade).

CONCLUSION

From the result of graduate of randomly selected ten departments of The Federal Polytechnic, Ado-Ekiti, Chi-Square test of independence reveals that: (i) Sex is dependent on the final grade and (ii) Department is dependent on the final grade
Breslow-Day and Tarone's statistics show that the null hypothesis of homogeneity of odds ratio across the departments is not rejected for sex. This implies that the odds ratio across the ten departments (relating to sex & final grade) are all equal. Cochran's and Mantel-Haenszel statistics reveals the final grade of students (Low Grade or High Grade) is not associated with sex

POLICY IMPLICATION

This research has revealed that there is equal chance for both male and female students to graduate from the Polytechnic with High or Low Grade, hence, governments' policy on education should be focused on both genders instead of special attention usually given to female students.

REFERENCES

Daniel, A. P. and Yu, X. (1999): "*Statistical methods for categorical data analysis*", Academic press, Inc

Lawal Bayo (2003): *Categorical Data Analysis with SAS and SPSS Applications*, Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersy. p. 138

Cochran, W. G. (1954): *Some methods for strengthening the common 2x2 tests*, Biometrics 10(4), 417–451.

Mantel, N. and Haenszel, W. (1959): *Statistical aspects of the analysis of data from retrospective studies of disease*, Journal of the National Cancer Institute 22, 719–748.

Birch, M. W. (1964): *The detection of partial association: the 2 X 2 case,* Journal of the Royal Statistical Society, B 26, 313 – 324

Breslow, N. and Day, N. (1980): *Statistical Methods in Cancer Research I: The Analysis of Case-Control Studies*, International Agency for Research on Cancer, Lyon.

Rothman, K. J. and Boice, J. D. (1979): Epidemiologic Analysis with a Programmable Calculator, U.S. Government Printing Office, Washington, DC.

Tarone, R. E. (1981): On summary estimators of relative risk, Journal of Chronic Disease (34), 463 – 468

Nurminen, M. (1981): Asymptotic efficiency of general non-iterative estimators of common relative risk, Biometrika (68), 525 – 530

Kleinbaum, D. G., Kupper. L. L. and Morgenstern. H. (1982): *Epidemiologic Research: Principles and Quantitative Methods*, Lifetime Learning Publications, Belmont, CA.

Greenland, S. (1982): Interpretation and estimation of summary ratios under heterogeneity, Statistics in Medicine, (1), 217 – 227

Suesse, Thomas Falk (2009): *Analysis and Diagnostics of Categorical Variables with Multiple Outcomes,* Unpublished PhD thesis, Victoria University of Wellington

Gart, J. J. (1962): *On combination of relative risks*, Biometrics 18(4), 594–600.

Breslow, N. E. (1981): *Odds ratio estimators when the data are sparse*, Biometrika 68(1), 73–84.

Hauck, W. W. (1979): *Large sample variance of the Mantel-Haenszel estimator of a common odds ratio*, Biometrics 35(4), 817–819.

Breslow, N. E. and Liang, K. Y. (1982): *The variance of the Mantel-Haenszel estimator*, Biometrics 38(4), 943–952.

Robins, J., Breslow, N. and Greenland, S. (1986): *Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models*, Biometrics 42(2), 311–323.

Anderson, E. B. (1980): *Discrete Statistical Models with Social Science Applications*, North-Holland, New York.

Guilbaud, O. (1983): *On the large-sample distribution of the Mantel-Haenszel odds-ratio estimator,* Biometrics 39(2), 523–525.

Tarone, R. E. (1985): *On heterogeneity tests based on efficient scores*, Biometrika 72(1), 91–95.

Liang, K. Y. and Self, S. G. (1985): *Tests for homogeneity of odds ratio when the data are sparse*, Biometrika 72(2), 353–358.

Paul, S. R. and Donner, A. (1989): *A comparison of tests of homogeneity of odds ratios in Kx2x2 tables*, Statistics in Medicine 8(12), 1455–1468.

Liu, Q. and Pierce, D. A. (1993): *Heterogeneity in Mantel-Haenszel-type models,* Biometrika 80(3), 543–556.