DOCUMENT RESUME

ED 049 280 TM 000 452

AUTHOR	Timm, Neil H.
TITLE	Neyman's Restricted Chi-Square Tests.
PUB DATE	Feb 71
NOTE	14p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Behavioral Science Research, Data Analysis, *Hypothesis Testing, Mathematical Applications, *Mathematical Models, Probability, Research Methodology, *Research Tools, *Statistical Analysis
IDENTIFIERS	*Chi Square Test

ABSTRACT

Pearson's unrestricted chi-square procedure is reviewed, and an historical presentation of Neyman's restricted chi-square test is introduced with a discussion of its theory and applicability to education. An example of the Neyman procedure is discussed in detail to familiarize researchers with this useful technique for analyzing contingency tables. The analysis also displays the need for researchers to check model assumptions and power in order to produce constructive analysis. This presentation of a statistical procedure developed by mathematical statisticians allows researchers in the behavioral sciences a facility with the method for application in their particular research. (Author/PR)

Neyman's Restricted Chi-Square Tests*

Abstract

Neil H. Timm
University of Pittsburgh

*Presented at the American Educational Research Association Annual Meeting
New York City 1971

Neyman's Restricted Chi-Square Tests

Neil H. Timm

University of Pittsburgh

Abstract

A historical presentation of Neyman's restricted chi-square tests is introduced with a discussion of its theory and applicability to education included. This presentation of a statistical procedure developed by mathematical statisticians allows researchers in the behavioral sciences a facility with the method for application in their particular research.

Introduction

Karl Pearson's (1900) chi-square test criterion for contingency tables has been employed in many areas of educational research whenever mutually exclusive and exhaustive qualitative events occur. Neyman's (1949) classical paper, which has remained unnoticed by educational methodologists along with the Fix, Hodges and Lehmann (1959) article, extends the analysis of categorized data by restricting the class of admissible hypotheses. Before considering the consequences of Neyman's modification of the chi-square test, a review of Pearson's (unrestricted) chi-square procedure will be made.

The Unrestricted Chi-Square Test

For illustration, consider a two-way rxc contingency table. Suppose n observations can be classified according to two characteristics, A and B, with $A_1, A_2, \ldots, A_r$ and $B_1, B_2, \ldots, B_c$ mutually exclusive and exhaustive categories. The observed frequency in the cell A_iB_j is denoted by n_{ij} and the probability represented by P_{ij}. Also define

$$n_{i.} = \sum_{j=1}^{c} n_{ij} \qquad\qquad n_{.j} = \sum_{i=1}^{r} n_{ij}$$

$$p_{i.} = \sum_{j=1}^{c} p_{ij} \qquad\qquad p_{.j} = \sum_{i=1}^{r} p_{ij}$$

Since each of the n observations have to be classified into one of the rc cells, it follows that

$$\sum_i \sum_j n_{ij} = \sum_i n_{i.} = \sum_j n_{.j} = n$$

$$\sum_i \sum_j p_{ij} = \sum_i p_{i.} = \sum_j p_{.j} = 1$$

The class of admissible hypotheses Ω is represented by

$$\Omega = \{ p_{ij} > 0 \text{ with } \sum_i \sum_j p_{ij} = 1 \qquad i=1, \ldots, r,\ j=1, \ldots, c \}$$

where parameters are $p_{11}, p_{12}, \ldots, p_{rc}$ and the sample size is $n_{11}, n_{12}, \ldots, n_{rc}$.

Although there are many hypotheses which may be tested under the above model, only the test of independence is considered. To test the hypothesis H of independence

$$H: \quad p_{ij} = p_{i.} p_{.j} \qquad i=1, \ldots, r,\ j=1, \ldots, c$$

that the parameters belong to ω, a subset of Ω, where

$$\omega = \{ p_{ij} \text{ such that } p_{ij} \in \Omega \text{ and } p_{ij} = p_{i.} p_{.j} \qquad i=1, \ldots, r,\ j=1, \ldots, c \}$$

it is necessary to find estimates of p_{ij} that minimize

$$X_H^2 = \sum_i \sum_j (n_{ij} - np_{ij})^2 / np_{ij} = \sum_{i,j} \frac{(O-E)^2}{E}$$

under ω such that $\sum_i \sum_j p_{ij} = 1$.

In general, the distribution of X_H^2 depends on the estimation procedure and on the number of unknown parameters. R.A. Fisher (1924) was the first to find the limiting distribution of X_H^2 for an important class of methods of estimation under rather general conditions. Fisher illustrated that as n tends to ∞ and the number of cells remain fixed, the distribution of X_H^2 tended to a $\chi^2_{f(H)}$ distribution on f(H) degrees of freedom; f(H) being the number of cells minus the number of independent parameters estimated minus 1. Cramer (1946, page 424) and Neyman (1949) extended Fisher's results. Neyman showed his outcome true for any best asympototically normal, BAN, estimates which include maximum likelihood, minimum chi-square, minimum modified chi-square and Neyman's minimum "linearized" chi-square estimates.

The distribution of the random variables $N_{11}, N_{12}, \ldots, N_{rc}$ under the above model is by definition, multinomial,

$$P\{ (N_{11}=n_{11}), \ldots, (N_{rs}=n_{rs}) \mid \Omega \} = n! \prod_i \prod_j p_{ij}^{n_{ij}} / \prod_i \prod_j n_{ij}!$$

with $\sum_i \sum_j n_{ij}$ and $\sum_i \sum_j p_{ij} = 1$. Under H, the maximum likelihood estimate of p_{ij} is $\hat{p}_{ij} = n_{i.}n_{.j}/n^2$ so that the observed value of X_H^2 is

$$X_H^2 = \sum_i \sum_j (n_{ij}-n_{i.}n_{.j}/n)^2 / n_{i.}n_{.j}/n.$$

If n is large and not too many of the $\hat{p}_{ij}$'s are small (all $n\hat{p}_{ij}$'s ≥ 3), then X_H^2 is distributed approximately under H as a $\chi^2_{f(H)}$ with $f(H) = rc-(r+c-2) -1 = (r-1)(c-1)$ degrees of freedom. This procedure may be used with caution even if some of the expected cell frequencies are less than 1 provided n is large and rc is moderate (greater than 6).

One often finds that researchers regroup their data to remove low expected values. This procedure effects the power of the chi-square test and if estimates of the parameters are based on the ungrouped data the limiting distribution of X_H^2 is not chi-squared (as it is when the estimates are based on the grouped data) according to Chernoff and Lehmann (1954).

The unrestricted chi-square test may be used to test nearly any hypothesis in which observations are grouped into distinct cells (a weak restriction) for which BAN estimates of the expected cell frequencies can be found. The test is easy to apply, consistent against all alternatives to the hypothesis tested and sensitive in all directions. However, in any particular problem it may be less desirable than a test designed to test particular alternatives.

Restricted Chi-Square Tests

Under the class Ω of admissible hypotheses, Neyman imposes restrictions on the cell probabilities $p_1, p_2, \ldots, p_m$. Given these restrictions, the hypothesis H further constrains the relations among the p_i's. For example, the probabilities under the general model Ω may depend in a particular manner on some unknown parameters $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_s)$ so that we may write $p_k(\underline{\theta})$ under Ω. The hypothesis H can, for example, specify that $\theta_2 = 0$. Under this restriction, the unrestricted chi-square test would seem to be undesirable since this statistic does not consider what happens under Ω. Intuitively, the chi-square criterion should test the hypothesis H against Ω - H and not against the most general possible alternative in Ω. Neyman's restricted chi-square test does exactly this; the restricted chi-square criterion is the difference

$$X_R^2 = X_H^2 - X_\Omega^2.$$

This difference measures the increase in chi-square imposed by the additional restrictions in H, over those in Ω.

More formally, suppose n observations $X_1, X_2, \ldots, X_n$ are classified into m mutually exclusive and exhaustive cells, and that the number of observations recorded in the k^{th} cell is x_k. Under Ω assume that the probability of any X occurring in the k^{th} cell is $p_k(\underline{\theta}) = p_k(\theta_1, \ldots, \theta_s)$ involving $s < k$ unknown parameters and that under H assume that the probability is $\pi_k(\underline{\theta}) = \pi_k(\theta_1, \ldots, \theta_s)$ where $H \subset \Omega$ and $\sum_k p_k(\underline{\theta}) = 1$ and $p_k(\underline{\theta}) > c^2 > 0$ for all k. Further, let $\underline{\hat{\theta}}$ be a BAN estimate of $\underline{\theta}$ under H and $\underline{\hat{\hat{\theta}}}$ under Ω. Define the restricted chi-square criterion as

$$X_R^{\,2} = X_H^{\,2} - X_\Omega^{\,2} = \sum_{i=1}^{m} \frac{(x_i - n\pi_i(\hat{\theta}))^2}{n\,\pi_i(\hat{\theta})} - \sum_{i=1}^{m} \frac{(x_i - np_i(\hat{\hat{\theta}}))^2}{np_i(\hat{\hat{\theta}})}.$$

with

$$f(R) = f(H) - f(\Omega)$$

degrees of freedom. $f(\Omega)$ is the total number of <u>independent</u> <u>cells</u> minus the number of <u>independent</u> <u>parameters</u> estimated from the data under Ω; and $f(H)$ is the same under the hypothesis H.

It should be observed that $X_H^{\,2}$ is the unrestricted chi-square described by Pearson and thus has the same number of degrees of freedom as before.

Neyman (1949) shows that the restricted chi-square criterion $X_R^{\,2}$ (given Ω) has, asymptotically as $n \to \infty$ and m remains fixed, a chi-square distribution on $f(R)$ degrees of freedom. He also shows the asympotic equivalence of his criterion with the Wilks λ-criterion, a result which is in agreement with the unrestricted case.

The test statistic $X_\Omega^{\,2}$ may be employed to test Ω against more general hypotheses with only trivial restrictions (such as $\sum_k p_k = 1$). Thus, $X_\Omega^{\,2}$ may be used to test the model. Upon rejection of the model, one can either relax the conditions under Ω to some wider class Ω^* or abandon Ω altogether and use the most general class of admissible hypotheses. In this case, X_H^2 would become the test criterion rather than X_R^2.

In utilizing the restricted chi-square test in practice, the model must be tested. The procedure allows one to compute

X_H^2 which tests H against all admissible hypotheses
X_Ω^2 which tests Ω against all admissible hypotheses
X_R^2 which tests H against Ω (admissible) hypotheses

Asymptotic Power Calculations

A general idea of the power of the test under consideration is required in educational research. Basic to the approximation of power in contingency tables is the noncentral chi-square distribution with non-centrality parameter ϕ_H and one or more alternative hypotheses $k \in K$. Wald (1943) has shown how one may estimate the non-centrality parameter for large samples. Let H be a particular hypothesis and $T(X_1, \ldots, X_n)$ under an alternative $k \in K$; just as under H, the central chi-square distribution is supposed to approximate the discrete distribution of $T(X_1, \ldots, X_n)$. Further, the approximate non-centrality parameter ϕ_H under an alternative $k \in K$ is the value of the statistic $T(X_1, \ldots, X_n)$ with the expected value of X_i under K substituted everywhere for X_i. Hence $\phi_H = T(E_k X_1, \ldots, E_k X_n)$.

Tables prepared by Fix (1949) and Fix, Hodges and Lehmann (1959) are available to allow easy power calculations by entering with ϕ_H(their λ) and degrees of freedom f.

For Neyman's restricted chi-square test, the same rule applies; however, ϕ_H and ϕ_Ω is required.

Example

An educational researcher elects to study the relationship between students having to satisfy the statistics methodology requirement and the advent of those students being placed on scholastic probation. A (hypothetical) random sample of students who enrolled in the College of Education has been collected.

Let j denote the number of quarters of statistics completed (j=0, 1, 2, 3, 4 or more) and i denote whether a student has ever been on probation (i=1, yes; i=2, no). Rather than merely testing whether going on probation is independent of the number of quarters of statistics completed, the question might be whether the probability of being on probation decreases as the number of quarters of statistics completed increases.

Under Ω, the model becomes

Ω: Multinomial (p_{ij}, n) where

$$n = \sum_{i,j} n_{ij} \; ; \quad \sum_i \sum_j p_{ij} = 1$$

and

$$p_{1j} = (\alpha + \beta\ (j-2))p_{.j} \qquad j=0,\ 1,\ 2,\ 3,\ 4$$

or since $p_{1j} + p_{2j} = p_{.j}$

$$p_{2j} = (1 - \alpha - \beta\ (j-2))p_{.j}$$

Some type of relationship would be expected to exist among the p_{ij}'s as j increases, the probability of running into academic difficulty would be less likely. That the model is linear will have to be checked. The parameter β measures the increased difficulty due to the statistics requirement, and α is a sort of average $p_{i.}$.

Under the hypothesis H, that no linear relationship exists, it is desirable to test

$$H: \ \beta = 0$$

$$K: \ \beta \neq 0$$

The data for this study follow

j = number of quarters

		0	1	2	3	4 or more	
i = on probation at any time	yes	16	[illegible]	3	2	20	50
	no	11	17	7	4	115	154
		27	26	10	6	135	204

To apply the Neyman restricted chi-square test to this data, estimates of α and β under Ω are necessary. From this model

$$p_{1j} = (\alpha + \beta\,(j-2))p_{.j}$$

or

$$\frac{p_{1j}}{p_{.j}} = \alpha + \beta\,(j-2)$$

Since a multinomial distribution exists, the maximum likelihood estimate of

$$\frac{p_{1j}}{p_{.j}} = \frac{n_{1j}/n}{n_{.j}/n} = n_{1j}/n_{.j}.$$

Letting

$$G(\alpha,\beta) = \sum_{j=0}^{4} ((n_{1j}/n_{.j}) - \alpha-\beta(j-2))^2 n_{.j}.$$

Taking partial derivatives with respect to α and β, and equating to zero, the following equations result

$$\sum_j n_{1j} = \alpha \sum_j n_{.j} + \beta \sum_j n_{.j}(j-2)$$

$$\sum_j n_{1j}(j-2) = \alpha \sum_j n_{.j}(j-2) + \beta \sum_j n_{.j}(j-2)^2.$$

Employing the data,

$$50 = 204\alpha + 196\beta$$

$$1 = 196\alpha + 680\beta$$

solving this system yields the estimates $\hat{\alpha}_\Omega = .33702$, $\hat{\beta}_\Omega = -.09567$.

Estimating the parameters under H: $\beta = 0$, only $\hat{\alpha}_H$ needs to be found. Minimizing

$$G(\alpha) = \sum_j \left((n_{1j}/n_{.j}) - \alpha\right)^2 n_{.j}$$

with respect to α yields

$$\hat{\alpha}_H = \frac{n_{1.}}{n} = \frac{50}{204} = .24510$$

Thus, under Ω

$$\hat{\hat{p}}_{1j} = (.33702 - .09567\,(j-2))\,\hat{p}_{.j} \qquad (1)$$

and under H

$$\hat{p}_{1j} = (.24510)\,p_{.j} \qquad (2)$$

where $\hat{p}_{.0} = .1323$, $\hat{p}_{.1} = .1275$, $\hat{p}_{.2} = .0490$, $\hat{p}_{.3} = .0294$, and $\hat{p}_{.4} = .6618$.

In terms of equations (1) and (2), the expression for X^2_R becomes

$$X^2_R = X^2_H - X^2_\Omega$$

$$= \sum_{i,j} \frac{(n_{ij} - n\hat{p}_{ij})^2}{n\hat{p}_{ij}} - \sum_{i,j} \frac{(n_{ij} - n\hat{\hat{p}}_{ij})^2}{n\hat{\hat{p}}_{ij}}$$

The expected tables under H and Ω are respectively:

j = number of quarters

H:

	6.615	6.375	2.450	1.470	33.090	50
i	20.385	19.625	7.550	4.530	101.910	154
	27	26	10	6	135	204

j = number of quarters

Ω:	14.261	11.254	3.369	1.448	19.668	50
i	12.739	14.746	6.631	4.552	115.332	154
	27	26	10	6	135	204

so that

$$X^2_H = 26.343 \text{ where } f(H) = 4$$

$$X^2_\Omega = 1.590 \text{ where } f(\Omega) = 3$$

$$X^2_R = 24.753 \text{ where } f(R) = 1$$

The statistic X^2_Ω is employed to test the model assumptions. The degrees of freedom $f(\Omega)$ is obtained by subtracting the number of independent parameters estimated (6) from the number of independent cells (9). Since $X^2_\Omega < \chi^2_3(.95) = 7.81$, the model assumptions are tenable.

Given that the model assumptions are satisfied, testing the hypothesis H: $\beta = 0$ against the alternative K: $\beta \neq 0$ may proceed. Since $X^2_R > \chi^2_1(.95) = 3.84$ the hypothesis of independence is rejected (the same conclusion as reached by the unrestricted test).

More information has been gathered by use of the restricted chi-square procedure because of a prediction equation involving the cell probabilities has been obtained

$$\hat{p}_{1j} = .33702 - .09567\ (i-2)\ \hat{p}_{.j}$$

$$\hat{p}_{1j} = .33702 - .09567\ (i-2)\ (n_{.j}/n) \text{ for } j=0,\ 1,\ 2,\ 3,\ 4.$$

This type of relationship could not have been procured by an unrestricted chi-square procedure. At most, a phi coefficient might have been calculated.

Having a regression equation for the probabilities, association of a(1-α) confidence interval for the coefficient β is sought. Employing the formula,

$$\hat{\beta} - \hat{\sigma}_{\hat{\beta}} \sqrt{\chi_1^2(.95)} \leq \beta \leq \hat{\beta} + \hat{\sigma}_{\hat{\beta}} \sqrt{\chi_1^2(.95)}$$

the following approximate $1 - \alpha$ confidence interval for β is created.

$$-.11 \leq \beta \leq .08.$$

As expected, zero is not included in the interval, again verifying that β is significantly different from zero.

Power Calculations

The power of a restricted chi-square test is always greater than an unrestricted chi-square test given that the alternative of interest reasonably satisfies the model restrictions. Deciding on whether to use restricted or unrestricted chi-square tests reduces to a choice between excellent power for a limited class of alternative or weak power for every other (Fix, Hodges, Lehmann, 1959).

To compute power in the example considered, all values of p_{ij} for $p_{ij} \in k$ must be specified and ϕ_R by the Wald (1954) procedure needs to be evaluated.

$$\phi_R = \sum_{i,j} (E(N_{ij}|k) - \hat{E}(N_{ij}|k|H))^2 / \hat{E}(N_{ij}|k|H)$$

$$- \sum_{i,j} (E(N_{ij}|k) - \hat{E}(N_{ij}|k|\Omega))^2 / \hat{E}(N_{ij}|k|\Omega)$$

$$= \phi_H - \phi_\Omega$$

But since $k \in \Omega$, it is observed that $\phi_\Omega = 0$ so that $\phi_R = \phi_H$ with degrees of freedom equal to f(R). By contrast, the power of an unrestricted chi-square test is obtained by use of ϕ_H with degrees of freedom f(H).

The above relationship indicates the necessity of always testing the model; when the model assumptions are not satisfied the unrestricted chi-square procedure should be employed since power will be larger.

Conclusion

The purpose of the preceeding example and discussion of the Neyman procedure was to familiarize researchers with a useful technique for analyzing contingency tables. However, the analysis also displays the need for researchers to check model assumptions and power in order to produce constructive analysis.

References

Chernoff, H. and E.L. Lehmann (1954). "The use of maximum likelihood estimates in χ^2 tests for goodness of fit," Annals of Mathematical Statistics, Volume 25, pages 579-586.

Cochran, William (1955). "A test of a linear function of the deviations between observed and expected numbers," Journal of the American Statistical Association, Volume 50, pages 377-397.

Cramer, Harold (1946). Mathematical Methods of Statistics. Princeton University Press.

Fisher, R.A. (1924). "The conditions under which χ^2 measures the discrepancy between observation and hypothesis," Journal of the Royal Statistical Society, Volume 87, pages 442-449.

Fix, Evelyn (1949). "Tables of noncentral χ^2," University of California Publications in Statistics, Volume 1, pages 15-19.

Fix, E., Hodges, J.L. and Lehmann, E.L. (1959). "The restricted chi-square test," Probability and Statistics, The Harold Cramer Volume. John Wiley & Sons, pages 92-107.

Neyman, J. (1949). "Contribution to the theory of the χ^2 test," Berkeley Symposium on Mathematics, Statistics and Probability, pages 239-273.

Pearson, Karl (1900). "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have risen from random sampling," Philosophical Magazine, Volume 50, pages 157-172.

Wald, A. (1943). "Tests of statistical hypotheses concerning several parameters when the number of observations is large," Transactions of the American Mathematical Society, (3), Volume 54, pages 426-482.