Statistics Canada Statistique Canada

Ottawa

APPROXIMATIONS TO THE DISTRIBUTION OF A SUM OF WEIGHTED CHI-SQUARE VARIABLES

WORKING PAPER NO. BSMD 85-007E.

M.A. HIDIROGLOU
JANUARY 1983

TABLE OF CONTENTS

APPROXIMATIONS TO THE DISTRIBUTION OF A SUM OF WEIGHTED CHI-SQUARE VARIABLES

1. Introduction

There are many areas in statistical analysis that need probability computations of quadratic forms in normal variables which we denote by X^2. The quadratic form X^2 is of the type

$$X^2 = \sum_{j=1}^{k} \lambda_j (S_j + a_j)^2$$

where S_j are independent and identically distributed standard normal variables (i.e., mean zero and variance one), and where λ_j and a_j are nonnegative constants. This form of linear combination of quadratic forms with $a_j = 0$, $j = 1, \ldots, k$ is one obtained by Rao and Scott (1981) in studying chi-square tests to data obtained from complex surveys. As shown by Rao and Scott (1981), the usual Pearson goodness-of-fit chi-square test,

$$X^2 = n \sum_{i=1}^{k+1} (\hat{p}_i - p_{oi})^2 / p_{oi}$$

where $\hat{p}_i$'s are design-consistent estimator of $\hat{p}_i$'s (the population proportions) under p(s) (the sample design under consideration), with $\Sigma\, \hat{p}_i = 1$, n being the sample size and p_{oi} is the expected proportion under the null hypothesis, may be written for large n as $\sum_{i=1}^{k} \lambda_{oi} S_i^2$. Here S_i's are asymptotically independent N(0,1) random variables and λ_{oi}'s are the eigenvalues of $\underset{\sim}{D}_o = n\, \underset{\sim}{P}_o^{-1} \underset{\sim}{\Sigma}_o$ $(\lambda_{o1} \geq \lambda_{o2} \geq \ldots \geq \lambda_{o,k} \geq o)$, where $\underset{\sim}{\Sigma}_o$ denotes the covariance matrix of the $\hat{p}_i$'s under the design and $\underset{\sim}{P}_o$ is the covariance matrix of the estimates of p_i's under multinomial sampling, when $p_i = p_{oi}$, $i = 1, \ldots k$.

Significance levels for X^2 are required in order to compare them to a nominal size (α), i.e.

$$SL(X^2) = Pr[X^2 \geq \chi^2_k(\alpha)],$$

where $\chi^2_k(\alpha)$ is the upper α-point of chi-square random variable with k degrees of freedom. These significance levels may be obtained exactly using a procedure given by Imhof (1961). Imhof's procedure is based on numerical inversion of the characteristic function of X^2. This method, though accurate for all practical purposes, is relatively expensive in computer time. Various approximations to $SL(X^2)$ have been proposed: they fall into two types, Gaussian and chi-square. Jensen and Solomon (1972) proposed the Gaussian approximation $Z = (X^2/\theta)^h$, where θ is the mean of X^2, and approximates Z by a normal distribution, where the mean and variance of the approximating distribution depend on the first three cumulants of X^2. Chi-square approximations have been proposed by Satterthwaite (1946) and Solomon and Stephens (1977). Solomon and Stephens' study suggests that their approximation is more accurate than the Jensen-Solomon approximation in both tails of the distribution for X^2. The behaviour of the Solomon-Stephens approximation will be compared to the Satterthwaite approximation in the upper tail.

2. Solomon-Stephens Approximation ($a_j = 0$, $j = 1, \ldots, k$)

The distribution of $X^2 = \Sigma_i \lambda_i S_i^2$ is fitted by $\hat{X}^2 = A\,Y^r$ where Y has the χ^2_t distribution, and the constants A, t and r are found by matching the first three moments of $\hat{X}^2$ to those of X^2. The first three moments of X^2 when all the a_j's are zero are

$$\mu_1' = \sum_{i=1}^{k} \lambda_i,$$

$$\mu_2' = 2\sum_{i=1}^{k} \lambda_i^2 + \Big(\sum_{i=1}^{k} \lambda_i\Big)^2 \qquad (2.1)$$

and

$$\mu_3' = 8\sum_{i=1}^{k} \lambda_i^3 + 6\Big(\sum_{i=1}^{k} \lambda_i\Big)\Big(\sum_{j=1}^{k} \lambda_j^2\Big) + \Big(\sum_{i=1}^{k} \lambda_i\Big)^3$$

while those for $\hat{X}^2$ are

$$\mu_1' = A\, 2^r\{\Gamma(r+v)\}/C,$$

$$\mu_2' = A^2 4^r\{\Gamma(2\,r+v)\}/C, \tag{2.2}$$

and

$$\mu_3' = A^3 8^r\{\Gamma(3\,r+v)\}/C,$$

where $v = t/2$ and $C = \Gamma(v)$.

The system of equations (2.1) and (2.2) yields

$$R_2 = \mu_2'/\mu_1'^2 = C\ \Gamma(2r+v)/\{\Gamma(r+v)\}^2 \tag{2.3}$$

$$R_3 = \mu_3'/\mu_1'^3 = C^2\ \Gamma(3r+v)/\{\Gamma(r+v)\}^3 \tag{2.4}$$

where R_2 and R_3 are obtained from (2.1). Equations (2.3) and (2.4) can now be solved using Newton-Raphson or secant procedures.

3. Satterthwaite's Approximation

Satterthwaite (1946) fitted X^2 by $\hat{X}^2 = A\, w$ where w has the χ_r^2 chi-squared distribution by matching only the first two moments of X^2 to those of $\hat{X}^2$. This approximation provides direct estimates of A and r as

$$A = \left(\sum_{i=1}^{k} \lambda_i^2\right)/\left(\sum_{i=1}^{k} \lambda_i\right)$$

and

$$r = \left(\sum_{i=1}^{k} \lambda_i\right)^2/\left(\sum_{i=1}^{k} \lambda_i^2\right).$$

A and r may be re - expressed as

$$A = \bar{\lambda}\ \{1 + CV^2(\lambda_i)\}$$

and

$$r = k/\{1 + CV^2(\lambda_i)\}$$

where $CV^2(\lambda_i) = \Sigma_i(\lambda_i - \bar{\lambda})^2/k\bar{\lambda}^2$ is the squared coefficient of variation of the λ_i's. For our purposes, the advantage of this approximation over the Solomon-Stephens approximation is that the λ_i's need not be known individually. Rather, only the covariance matrix $\hat{\underset{\sim}{\Sigma}}_o$ need be known as will be seen later.

4. Uses of the Approximations

Analysis of survey data is frequently done by summarizing the data in the form of multi-way tables. The summarized data is then analysed using Pearson chi-squared statistics in order to test various hypotheses put forth by social or economic data researchers. The significance levels of such tests are known to be substantially higher than the nominal levels (see Rao and Scott 1981, Holt, Scott and Ewings 1980, Fellegi 1980) if the effect of the sample design is pronounced: that is the combined stratification and cluster effect make the sample design depart significantly from a simple random sample. The effect of design in such cases must be taken into account by either computing a Wald chi-square statistic or modifying the Pearson chi-square statistic. Rao and Scott (1981) have proposed procedures for the latter treatment in the case of goodness-of-fit, homogeneity and independence in two-way tables. The two approximations will be used to study significance levels for the test of independence in particular.

For the test of independence in a two-way table $(I+1) \times (J+1)$, the null hypothesis of interest is given by

$$H_o: \; p_{ij} = p_{i+} \, p_{+j}, \; i=1, 2, \ldots, I+1; \; j=1, 2, \ldots, J+1$$

where p_{ij} denotes the population proportion in the $(i,j)^{th}$ cell, $p_{i+} = \sum_{j=1}^{J+1} p_{ij}$ and $p_{+j} = \sum_{i=1}^{I+1} p_{ij}$. The customary Pearson statistic for testing H_o is

$$X_P^2 = n \sum_{i=1}^{I+1} \sum_{j=1}^{J+1} (\hat{p}_{ij} - \hat{p}_{i+} \hat{p}_{+j})^2 / (\hat{p}_{i+} \hat{p}_{+j})$$

where $\hat{p}_{ij}$, $\hat{p}_{i+}$ and $\hat{p}_{+j}$ are the respective estimates of p_{ij}, p_{i+} and p_{+j} under the sample design p(s). This test may be written in matrix form as

$$X_P^2 = n \, \hat{\underset{\sim}{h}}^T (\hat{P}_{I+}^{-1} \otimes \hat{P}_{+J}^{-1}) \, \hat{\underset{\sim}{h}}$$

where $\hat{\underset{\sim}{h}} = (\hat{h}_{11}, \ldots, \hat{h}_{1J}; \ldots; \hat{h}_{I1}, \ldots, \hat{h}_{IJ})^T$ with

$$\hat{h}_{ij} = \hat{p}_{ij} - \hat{p}_{i+}\,\hat{p}_{+j},\ \hat{\underset{\sim}{P}}_{I+} = \text{diag}\,(\hat{\underset{\sim}{p}}_{I+}) - \hat{\underset{\sim}{p}}_{I+}\,\hat{\underset{\sim}{p}}_{I+}^T,$$

$$\hat{\underset{\sim}{P}}_{+J} = \text{diag}\,(\hat{\underset{\sim}{p}}_{+J}) - \hat{\underset{\sim}{p}}_{+J}\,\hat{\underset{\sim}{p}}_{+J}^T,\ \hat{\underset{\sim}{p}}_{I+} = (\hat{p}_{1+}, \ldots, \hat{p}_{I+})^T,$$

$$\hat{\underset{\sim}{p}}_{+J} = (\hat{p}_{+1}, \hat{p}_{+2}, \ldots, \hat{p}_{+J})^T,$$

and $\otimes$ denotes the direct product.

Rao and Scott (1981) have shown that X_p^2 is asymptotically distributed as ($\approx$) a weighted sum, $\sum_{k=1}^{IJ} \delta_{ok}\, w_k$, of independent χ_1^2 random variables w_k under H_o. Here the δ_{ok}'s are the eigenvalues of the "design-effects matrix" $\underset{\sim}{D}_o = n(\underset{\sim}{P}_{I+}^{-1} \otimes \underset{\sim}{P}_{+J}^{-1})\,\underset{\sim}{\Gamma}$ where $\underset{\sim}{\Gamma}$ is the asymptotic covariance matrix of $\hat{\underset{\sim}{h}}$. Rao and Scott (1981) proposed a simple modification to X_p^2 requiring only the deffs of the $\hat{h}_{ij}$'s, $d_{ij}(\hat{\underset{\sim}{h}}) = \text{var}(\hat{h}_{ij})/[\hat{p}_{i+}(1-\hat{p}_{i+})\,\hat{p}_{+j}(1-\hat{p}_{+j})/n]$ $i=1, 2, \ldots, I+1$; $j=1, 2, \ldots, J+1$, where $\text{var}(\hat{h}_{ij})$ is the estimated variance of $\hat{h}_{ij}$ under H_o. The form of $\text{var}(\hat{h}_{ij})$ has been given by Hidiroglou and Rao (1981) for the case of the estimated proportions being derived as the ratio of unadjusted blown-up counts and as the ratio of adjusted (post-stratified on age-sex) blown-up counts. The modified statistic is

$$X_P^2\,(\hat{\delta}_.) = X_P^2 \,/\, \hat{\delta}_. \,,$$

where $(IJ)\,\hat{\delta}_. = \sum_{i=1}^{I+1} \sum_{j=1}^{J+1} d_{ij}(\hat{\underset{\sim}{h}})(1-\hat{p}_{i+})(1-\hat{p}_{+j})$.

An improved approximation, but requiring full knowledge of $\hat{\underset{\sim}{\Gamma}}$, is obtained by treating $X_P^2(S) = X_P^2(\hat{\delta}_.)/(1+\hat{C}^2)$ as χ_ν^2 where $\nu = IJ/(1+\hat{C}^2)$, $\hat{C}^2 = \sum_{k=1}^{IJ} \hat{\delta}_k^2/(\hat{\delta}_.^2\, IJ) - 1$. $\hat{C}^2$ may be obtained as a function of the estimated proportions and the estimated covariance by using $\sum_{k=1}^{IJ} \hat{\delta}_k^2 = \sum_{i=1}^{I+1} \sum_{j=1}^{J+1} \sum_{m=1}^{I+1} \sum_{\ell=1}^{J+1} \hat{\sigma}_{ij,m\ell}(\hat{\underset{\sim}{h}})/(p_{i+}\, p_{+j}\, p_{m+}\, p_{+\ell})$

without evaluating individual $\hat{\delta}_i$'s, where $\hat{\sigma}_{ij,m\ell}(\hat{\underset{\sim}{h}})$ is the estimated covariance between $\hat{h}_{ij}$ and $\hat{h}_{m\ell}$.

The Wald statistic takes the sample design into account and is of the form:

$$X_W^2 = \hat{\underset{\sim}{h}}^T \hat{\underset{\sim}{\Gamma}}^{-1} \hat{\underset{\sim}{h}}$$

where $\hat{\underset{\sim}{\Gamma}}$ is the estimated covariance matrix of $\hat{\underset{\sim}{h}}$ under the particular sample design used.

The asymptotic significance level of X_P^2 under H_o may be obtained using the Satterthwaite approximation as

$$\begin{aligned} SL\,(X_P^2) &= Pr[X_P^2 \geq \chi_{IJ}^2(\alpha)] \\ &\doteq Pr[\chi_\nu^2 \geq \chi_{IJ}^2(\alpha)/\{\hat{\delta}_.(1+\hat{C}^2)\}] \end{aligned}$$

where $\chi_{IJ}^2(\alpha)$ is the upper α - percentage point of a χ^2 random variable with IJ degrees of freedom. Using the Solomon-Stephens approximation, the estimated significance level is

$$\begin{aligned} SL(X_P^2) &= Pr[X_P^2 \geq \chi_{IJ}^2(\alpha)] \\ &\doteq Pr[\chi_t^2 \geq \{\chi_{IJ}^2(\alpha)/A\}^{1/r}] \end{aligned}$$

where X_p^2 has been fitted to the distribution of $A\,Y^r$ with Y distributed as χ_t^2.

Modified chi-square statistics of the form $X_p^2(\hat{b}_.) = X_p^2/\hat{b}_.$ have been proposed in the literature where $\hat{b}_.$ is a suitable linear combination of the estimated cell design effects $\hat{d}_{ij}$. Using Satterthwaite's approximation, the estimated significance level of $X_p^2(\hat{b}_.)$ is

$$SL[X_P^2(\hat{b}_.)] = Pr[X_P^2(\hat{b}_.) \geq \chi_{IJ}^2(\alpha)]$$

$$\doteq Pr[\chi_\nu^2 \geq \hat{b}_. \ \chi_{IJ}^2(\alpha)/\{\hat{\delta}_. (1+\hat{C}^2)\}] .$$

Using the Solomon-Stephens approximation, the estimated significance level is

$$SL[X_P^2(\hat{b}_.)] = Pr[X_P^2(\hat{b}_.) \geq \chi_{IJ}^2(\alpha)]$$

$$\doteq Pr[\chi_t^2 \geq \{\hat{b}_. \ \chi_{IJ}^2(\alpha)/A\}^{1/r}] .$$

Two particular cases of $X_p^2(\hat{b}_.)$ are:

$$X_P^2(\hat{d}_.) = X_P^2/\hat{d}_.$$

$$X_P^2(\hat{\lambda}_.) = X_P^2/\hat{\lambda}_.$$

where

$$\hat{d}_. = \sum_{i=1}^{I+1} \sum_{j=1}^{J+1} \hat{d}_{oij}/[(I+1)(J+1)]$$

and

$$\hat{\lambda}_. = \sum_{i=1}^{I+1} \sum_{j=1}^{J+1} (1 - \hat{p}_{i+} \hat{p}_{+j}) \ \hat{d}_{oij}/[(I+1)(J+1) - 1]$$

with $\hat{d}_{oij}$ being the estimated design effect of $\hat{p}_{ij}$ under H_o, that is: $\hat{d}_{oij} = var(\hat{p}_{ij})/[\hat{p}_{i+} \hat{p}_{+j}(1 - \hat{p}_{i+} \hat{p}_{+j})/n]$.

A modification based on Satterthwaite's approximation is given by

$$X_P^2(S,\alpha) = X_P^2[\frac{\chi_{IJ}^2(\alpha)}{\hat{\delta}_. (1+\hat{C}^2) \ \chi_\nu^2(\alpha)}]$$

where $SL[X_P^2(S,\alpha)] = P[X_P^2(S,\alpha) \geq \chi_{IJ}^2(\alpha)]$

$\doteq \alpha$.

Using the Solomon-Stephens approximation,

$$X_P^2(ST) = (X_P^2/A)^{1/r}$$

and

$$X_P^2(ST,\alpha) = X_P^2(ST)[\chi_{IJ}^2(\alpha)/\chi_t^2(\alpha)]$$

is such that $SL[X_P^2(ST,\alpha)] \doteq \alpha$.

5. Empirical Results

We now investigate the large-sample performance of X_P^2 for the test of independence and its given modifications, using some data from the Canada Health Survey. This survey is a stratified multi-stage cluster sample design (see The Health of Canadians pp. 229-233). All the computations have been done for the case of proportions of age-sex adjusted totals: further details of the estimation procedures used can be found in Hidiroglou and Rao (1981).

Example 1: (Independence in a 4×2 table)

Consider the estimated counts cross-classified by drug use (four categories: 0, 1, 2, 3+ drug classes in a 2-day period) and sex (male, female). Here, n = 31,688, I+1 = 4 and J+1 = 2 (IJ = 3) Drug use includes information on the use of medicines, pills or ointments within the past two days. The estimated counts (in thousands) of the population reporting in each of the drug by sex categories are given in the table provided below.

Table 1: Population by Variety of Drugs Taken, by sex (in thousands)

Sex	Drugs			
	0	1	2	3+
Male	6759	3081	1100	476
	(0.011)a	(0.026)	(0.027)	(0.045)
	(1.56)b	(3.37)	(1.15)	(1.38)
Female	5243	3659	1669	1035
	(0.020)	(0.023)	(0.034)	(0.036)
	(3.59)	(3.13)	(2.85)	(1.96)

Note: "a" stands for coefficient of variation of the cell
"b" stands for the cell DEFF

FOOTNOTE: A star above the numbers indicates that these numbers were suppressed in the Health of Canadians (1981) report.

Table 2: Chi-Square Statistics and Estimated Levels: Nominal Size 0.05

Statistic	Estimated Level Using	
	Satterthwaite	Solomon-Stephens
$X_p^2 = 774$	0.226	0.216
$X_p^2(\hat{\delta}_.) = 437$	0.062	0.060
$X_p^2(\hat{\lambda}_.) = 331$	0.022	0.024
$X_p^2(\hat{d}_.) = 327$	0.023	0.023
$X_p^2(S,0.05) = 408$	0.050	0.049
$X_p^2(ST) = 189$	-	-
$X_p^2(ST,0.05) = 138$	-	0.050
$X_W^2 = 538$	-	-

Note: Dashes indicate that no computation was done.

Table 2 gives the estimated significance levels using the Satterwaite approximation and the Solomon-Stephens approximation. For this example the eigenvalues were 0.9982, 1.3728 and 2.9386 with an estimated coefficient of variation equal to 0.47. The constants in the Solomon-Stephens approximation are: $r = 1.38$, $A = 0.558$ and $t = 4.766$. In this example, the Pearson statistic is quite high (774). The Wald statistic which is distributed approximately as a X_3^2 random variable under H_o is quite high (538) relative to the modified Pearson statistics. The $\hat{\delta}_.$ modification provides an estimated level fairly close to the required nominal level (0.05). Note, however, that both the $\hat{\lambda}_.$ and $\hat{d}_.$ modifications are quite conservative. The estimated significance levels using either the Solomon-Stephens or Satterthwaite approximations are quite close to each other.

Note on the other hand that the chi-square statistic obtained using the Solomon-Stephens approximation is quite a bit lower

(189) than all the other provided statistics. In fact, $X_p^2(ST) = 189$ is smaller that $X_p^2/(\text{max eigenvalue}) = 263.4$. This is something not to be expected since $X_p^2 = \sum_{k=1}^{IJ} \delta_{ok} W_k \le \delta_{max}(\sum_{k=1}^{IJ} W_k)$ where W_k's are independent $\chi^2_{(1)}$ random variables and $\delta_{max} = \max\{\delta_{ok}: k = 1, 2, \ldots, IJ\}$. Now, $(X_p^2/\delta_{max}) \le \sum_{k=1}^{IJ} W_k \approx \chi^2_{(IJ)}$; hence, $P[X_p^2/(\delta_{max}) \ge \chi^2_{IJ}(\alpha)] \le \alpha$. The explanation for this phenomenon is that the Solomon-Stephens approximation is not accurate in the extreme tails and the value $X_p^2 = 774$ is quite far out in the tail of the distribution with a minutely small α.

An approximation based on three moments (the first, second and third) is adequate over most of the range but not necessarily in the extreme tails. Box's (1954) investigation, however, has shown that the Satterthwaite approximation is fairly accurate in the tail of interest. In our example, $X_p^2(S,.05)$ would have been equal to $X_p^2(ST,.05)$ if X_p^2 had been equal to 15.1. This is obtained by equating $X_p^2(S,0.05)$ to $X_p^2(ST,.05)$ and solving for X_p^2:

$$\text{i.e. solving } X_p^2 = \exp\{\frac{r}{r-1}[\ln d - \frac{1}{r}\ln A - \ln g]\}$$

where $d = \chi^2_{IJ}(\alpha)/\chi^2_t(\alpha)$ and $g = \dfrac{\chi^2_{IJ}(\alpha)}{\hat{\delta}_{\cdot}(1+\hat{c}^2)\,\chi^2_\nu(\alpha)}$ with $\alpha = 0.05$, $\nu = 2.45$, $IJ = 3$, $\hat{\delta}_{\cdot} = 1.77$ and $\hat{c} = 0.47$.

• Example 2: <u>(Independence in a 7×2 table)</u>

Consider the estimated counts cross-classified by diastolic blood pressure (seven categories: less than 55, 55-64, 65-74, 75-84, 85-94, 95-104, 105+ in units of mmHg) and sex (male, female). Here n = 5760, I+1 = 7 and J+1 = 2. Blood pressure was measured on respondents five years of age and older. The estimated counts (in thousands) for the above cross-classification are provided in the table provided below.

Table 3: Population 5 years and over by diastolic blood pressure by sex

Sex	Blood Pressure						
	<55	55-64	65-74	75-84	85-94	95-104	105+
Male	390 (0.124)a (1.69)b	1307 (0.072) (1.98)	2589 (0.065) (3.35)	3524 (0.046) (2.31)	1905 (0.070) (2.87)	544 (0.154) (3.67)	184 (0.203) (2.11)
Female	484 (0.134) (2.56)	1900 (0.060) (1.94)	3490 (0.039) (1.76)	2981 (0.063) (3.70)	1270 (0.080) (2.33)	478 (0.173) (4.03)	44* (0.231) (0.64)

Note: "a" stands for coefficient of variation of the cell
"b" stands for the cell DEFF

Table 4: Chi-Square Statistics and Estimated Levels: Nominal Size 0.05

Statistic	Estimated Level Using	
	Satterthwaite	Solomon-Stephens
$X_p^2 = 141$	0.307	0.290
$X_p^2(\hat{\delta}_.) = 81$	0.081	0.077
$X_p^2(\hat{\lambda}_.) = 56$	0.019	0.021
$X_p^2(\hat{d}_.) = 56$	0.018	0.021
$X_p^2(S,0.05) = 70$	0.050	0.049
$X_p^2(ST) = 59$	-	-
$X_p^2(ST,0.05) = 36$	-	0.050
$X_W^2 = 107$	-	-

Note: Dashes indicate that no computation was done

Table 4 provides the estimated significance levels using the Satterthwaite approximation and Solomon-Stephens approximation for the usual Pearson statistic X_p^2 and its modifications. Summary conclusions for the above table will be provided after all examples have been produced.

• Example 3: (Independence in a 3×3 table)

Physical fitness was measured from pulse readings obtained after three minutes of a stepping exercise. Respondents were classified into three categories: "Recommended Level" (pulse below specified rate after six minutes), "Minimum Acceptable" (pulse rate below three-minute criterion but above six-minute criterion) and "Unacceptable" (pulse rate above criterion at three minutes). Physical fitness was classified by smoking level. The analysis was done for the domain where the fitness level was considered to be above a required minimum. Three levels of smoking were defined: current smoker, occasional smoker, non-smoker. Here $n = 1731$, $I+1 = 3$ and $J+1 = 3$. The estimated counts (in thousands) are provided below.

Table 5: Population 15-64 years by Fitness Level and Type of Cigarette Smoker

Type of Smoker	Fitness Level		
	Recommended	Minimum	Below
Current	2083 (0.083)a (4.28)b	1415 (0.091) (3.26)	138* (0.230) (2.76)
Occasional	218* (0.246) (4.06)	91* (0.182) (0.86)	8* (0.801) (1.68)
Never	1924 (0.081) (1.86)	940 (0.071) (1.26)	32* (0.314) (0.77)

NOTE: "a" stands for coefficient of variation of the cell
"b" stands for the cell DEFF

Table 6: Chi-Square Statistics and Estimated Levels: Nominal Size 0.05

Statistic	Estimated Level Using: Satterthwaite	Estimated Level Using: Solomon-Stephens
$X_p^2 = 24$	0.138	0.130
$X_p^2(\hat{\delta}_.) = 18$	0.071	0.068
$X_p^2(\hat{\lambda}_.) = 11$	0.008	0.009
$X_p^2(\hat{d}_.) = 10$	0.006	0.007
$X_p^2(S,0.05) = 16$	0.050	0.049
$X_p^2(ST) = 20$	-	-
$X_p^2(ST,0.05) = 14$	-	0.050
$X_W^2 = 24$	-	-

NOTE: Dashes indicate that no computation was done.

• Example 4: (Independence in a 5×3 table)

The table to be studied is cross-classified by diastolic blood pressure (five categories: 55-64, 65-74, 75-84, 85-94, 95-104 in units of mmHg) and type of smoking habit (three categories: current smoker, past smoker, never smoke). Here $n = 4007$, $I+1 = 5$ and $J+1 = 3$. The estimated counts (in thousands) for the above cross-classification are provided in the table provided below.

Table 7: Population 15 and over by diastolic blood pressure and type of cigarette smoker

Type of Smoker	Blood Pressure: 55-64	65-74	75-84	85-94	95-104
Current	824 (0.096)a (2.19)b	1876 (0.055) (1.92)	2087 (0.054) (2.03)	943 (0.091) (2.38)	364 (0.195) (3.72)
Past	384 (0.100) (1.12)	1073 (0.108) (3.99)	1574 (0.116) (5.53)	834 (0.098) (1.98)	256 (0.177) (2.07)
Never	475 (0.136) (2.41)	1417 (0.066) (1.78)	1817 (0.072) (2.36)	885 (0.064) (1.11)	293 (0.161) (2.10)

Table 8: Chi-Square Statistics and Estimated Levels: Nominal Size 0.05

Statistic	Estimated Level Using	
	Satterthwaite	Solomon-Stephens
$X_p^2 = 31$	0.308	0.289
$X_p^2(\hat{\delta}_.) = 19$	0.107	0.099
$X_p^2(\hat{\lambda}_.) = 13$	0.023	0.024
$X_p^2(\hat{d}_.) = 13$	0.022	0.024
$X_p^2(S,0.05) = 15$	0.050	0.049
$X_p^2(ST) = 24$	-	-
$X_p^2(ST,0.05) = 15$	-	0.050
$X_W^2 = 36$	-	

• Example 5: <u>(Independence in a 3×3 table)</u>

The final example provides a cross-classification between type of beverage (beer, wine, liquor) and number of drinks consumed per week (1-5, 6-13, 14+ drinks). Here $n = 6966$, $I+1 = 3$ and $J+1 = 3$. The estimated counts (in thousands) are provided below.

Table 9: Population 15 years (current drinkers only) and over by type of beverage (usually consumed) and number of drinkgs consumed per week

Number of drinks	Beverage		
	Beer	Wine	Liquor
1-5	223 (0.082)a (1.28)b	1124 (0.038) (1.15)	4585 (0.026) (0.62)
6-13	327 (0.060) (1.00)	806 (0.049) (1.04)	777 (0.048) (1.23)
14+	13* (0.258) (0.61)	1162 (0.045) (1.94)	294 (0.063) (1.01)

Table 10: Chi-Square Statistics and Estimated Levels: Nominal Size 0.05

Statistic	Estimated Level Using	
	Satterthwaite	Solomon-Stephens
$X_p^2 = 1974$	0.255	0.258
$X_p^2(\hat{\delta}_.) = 1102$	0.072	0.072
$X_p^2(\hat{\lambda}_.) = 1763$	0.210	0.210
$X_p^2(\hat{d}_.) = 1765$	0.218	0.218
$X_p^2(S,0.05) = 970$	0.050	0.049
$X_p^2(ST) = 336$	-	-
$X_p^2(ST,0.05) = 282$	-	0.050
$X_W^2 = 660$	-	-

6. Summary

The values for the X_p^2 chi-square statistic and selected modified versions for the five data sets are summarized in table 11. Several points may be noted from this table. For all the examples chosen, all the chi-square statistics are significant to at least the 5% level. The Wald statistic appears to be consistently higher than the given modified versions of the Pearson chi-square statistics and smaller than the Pearson chi-square statistic (see examples 1-4).

Table 11: Summary Chi-square Statistics for Examples 1-5

Example	Chi-square Statistic				
	X_p^2	$X_p^2(\hat{\delta}_{.})$	$X_p^2(S,0.05)$	$X_p^2(ST,0.05)$	X_W^2
1. DRUG*SEX	774***	437***	408***	138***	538***
2. PRESS*SEX	141***	81***	70***	36***	107***
3. FIT*SMOK	24***	18***	16***	14***	24***
4. PRESS*SMOK	31***	19**	15**	15**	36***
5. BEV*DRINK	1974***	1102***	970***	282***	660***

** - 0.05 level of significance
*** - 0.01 level of significance

The modified Solomon-Stephens chi-square statistics, $X_p^2(ST,0.05)$ is consistently lower than the other chi-square statistics. In fact, when the Pearson chi-square value is very high (see examples 1, 2 or 5), the modified Solomon-Stephens chi-square statistics is much lower than the other statistics. When the Pearson chi-square statistic is not high (as in examples 3 or 4), the modified Solomon-Stephens chi-square statistic is much closer to the modified Satterthwaite statistic $X_p^2(S,0.05)$. Note also that the estimated nominal levels obtained using either the Satterthwaite or Solomon-Stephens approximations are quite close to each other.

$X_p^2(S,0.05)$, $X_p^2(ST,0.05)$ or X_W^2 require the computation of the covariance matrix of the cell proportion estimates. For small contingency tables, this need not be a problem. For larger tables, as the number of variances and covariances increases with the square of the number of cells, which itself increases with the product of the number of categories in each classification variable, this matrix quickly gets out of hand. Furthermore, X_W^2 requires the inversion of the estimated covariance matrix which may itself be instable. $X_P^2(ST,0.05)$ moreover requires the computation of eigenvalues of the design effect matrix $\underset{\sim}{D}_o$. Note that $X_P^2(\hat{\delta}_.)$ is close to $X_P^2(S,0.05)$ in terms of value and estimated nominal level and requires far fewer computations to be obtained.

For all the examples provided in this report, inferences based on all approaches yield the same result: the null hypothesis is rejected. Acceptance or rejection of the null hypothesis using $X_P^2(S,0.05)$ depends on two factors, $\hat{\delta}_.$ and X_P^2. There are probably instances where the modified statistics do not yield the same inferences (for borderline cases). It would be interesting to compare the behaviour of all proposed chi-square statistics using a Monte-Carlo study. Such a study would provide information on the actual significance levels as opposed to the nominal ones for all the proposed statistics under various conditions. Some of these conditions are as follows: How suceptible are the chi-square statistics to variability amongst the individual cell designs effects in terms of inference? What is the effect of the magnitude of large or small X_P^2 on the inference? Small-sample power considerations would also be important.

7. Acknowledgements

The author is most thankful to Philip Giles for his assistance in preparing summary tapes. Thanks are also due to Nanjamma Chinnappa and Geoffrey Hole for their support. The author is most indebted to Professor J.N.K. Rao for his suggestions and comments.

APPENDIX

Table 12: Summary Statistics for Examples 1-5

Statistic	Example				
	1 DRUG*SEX	2 PRESS*SEX	3 FIT*SMOK	4 PRESS*SMOK	5 BEV*DRINK
A	0.560	0.075	0.233	0.062	0.693
r	1.38	1.85	1.54	1.94	1.37
t	4.77	13.70	6.98	14.63	5.17
$\hat{d}_.$	2.37	2.50	2.31	2.45	1.09
$\hat{\lambda}_.$	2.34	2.49	2.18	2.43	1.12
$\hat{\delta}_.$	1.77	1.74	1.28	1.61	1.79
$cv(\hat{\lambda}_i)$	0.47	0.75	0.63	1.08	0.69
ν	2.45	3.84	2.86	3.67	2.71
IJ	3.00	6.00	4.00	8.00	4.00
n	31688	5760	1731	4007	6966

Table 13: Eigenvalues for Examples 1-5

Example	Eigenvalues ($\hat{\delta}_i$)							
	1	2	3	4	5	6	7	8
1. DRUG*SEX	1.00	1.37	2.94					
2. PRESS*SEX	0.51	0.73	0.82	1.89	2.19	4.32		
3. FIT*SMOK	0.45	0.75	1.36	2.56				
4. PRESS*SMOK	0.20	0.42	0.47	0.69	0.92	1.20	3.65	5.36
5. BEV*DRINK	0.26	1.15	2.17	3.58				

Table 14: Frequency Distribution of the Cell Coefficients of Variation for Examples 1-5

Example	Coefficient of Variation Ranges				Mean (cv)	Sample Size n
	[0-5]	(5-10]	(10-20]	(20+		
1. DRUG*SEX	100%	-	-	-	3%	31688
2. PRESS*SEX	14%	43%	29%	14%	11%	5760
3. FIT*SMOK	-	57%	11%	44%	23%	1731
4. PRESS*SMOK	-	60%	40%	-	10%	4007
5. BEV*DRINK	56%	33%	-	11%	7%	6966

References

BOX, G.E.P., Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification, Ann.Math.Stat., Vol. 25 (1954), pp. 290-302.

FELLEGI, I.P. (1980), Approximate Tests of Independence and Goodness of Fit Based on Stratified Multistage Samples, Journal of the American Statistical Association, 75, 261-268.

HIDIROGLOU, M.A. and RAO, J.N.K. (1981), Chisquare Tests for the Analysis of Categorical Data from the Canada Health Survey, presented at the 43rd ISI Conference in Buenos Aires.

HOLT, D., SCOTT, A.J., and EWINGS, P.O. (1980), Chi-Squared Tests with Survey Data, Journal of the Royal Statistical Society, Sec. A, 143, 302-320.

IMHOF, J.P., (1961), Computing the Distribution of Quadratic Forms in Normal Variables, Biometrika, 417-426.

JENSEN, Donald R., and SOLOMON, Herbert (1972), A Gaussian Approximation to the Distribution of a Definite Quadratic Form, Journal of the American Statistical Association, 67, 898-902.

RAO, J.N.K. and SCOTT, A.J. (1981), The Analysis of Categorical Data from Complex Surveys: Chi-Squared Tests for Goodness of Fit and Independence in Two-Way Tables, Journal of the American Statistical Association, 76, 221-230.

SATTERTHWAITE, F.E. (1946), An Approximate Distribution of Estimates of Variance Components, Biometrics, 2, 110-114.

SOLOMON, H. and STEPHENS, M.A. (1977), Distribution of a Sum of Weighted Chi-Square Variables, Journal of the American Statistical Association, 72, 881-885.

The Health of Canadians, (1981), Report of the Canada Health Survey, Health and Welfare, and Statistics Canada publication, Catalogue 82-538E.