Power and Sample Size for Approximate Chi-Square Tests

WILLIAM C. GUENTHER*

Abstract

Approximate chi-square tests for hypotheses concerning multinomial probabilities are considered in many textbooks. In this article power calculations and sample size based on power are discussed and illustrated for the three most frequently used tests of this type. Available noncentrality parameters and existing tables permit a relatively easy solution of these kinds of problems.

KEY WORDS: Approximate chi-square tests; Hypotheses about multinomials; Power of the test; Sample size based on power; Noncentral chi-square; Noncentrality parameter.

1. Introduction

Approximate chi-square tests for hypotheses concerning multinomial probabilities are among the most frequently used statistical procedures. Rarely is any thought given to the power of the test or to sample-size determination based on the power, even though all the tools are available to handle these problems with relative ease.

As is well-known, statistics are available which are approximately distributed as chi-square, provided the multinomial conditions are satisfied and the hypothesis is true. Cases considered in numerous textbooks are: (1) the hypothesis which specifies all the multinomial probabilities, (2) the hypothesis of independence, and (3) the hypothesis of homogeneity. It is also known that if these hypotheses are false, then the appropriate statistic has approximately a noncentral chi-square distribution with the same degrees of freedom ν and a noncentrality parameter λ which depends on the alternative considered. For the cases mentioned above, formulas for λ are known for specific kinds of alternatives.

Easy solution to power problems also depends upon the availability of good tables of the noncentral chi-square distribution. Haynam, Govindarajulu, and Leone [2] have prepared two tables specifically designed for these problems. Their Table I gives power to four decimal places for:

$$\alpha = .001, .005, .01, .025, .05, .10,$$

$$\lambda = 0(.1)1(.2)3(.5)5(1)40(2)50(5)100,$$

$$\nu = 1(1)30(2)50(5)100,$$

while their Table II gives λ to three decimals for the same α and ν for which

$$\text{Power} = .1(.02).7(.01).99.$$

The latter table is particularly useful for sample size problems.

* William C. Guenther is with the Department of Statistics, University of Wyoming, Laramie, WY 82071.

2. Multinomial p's Specified by the Hypothesis

Let $(X_1, \ldots, X_k)$ be a multinomial random variable with parameters $n, p_1, p_2, \ldots, p_k$. Suppose we wish to test

$$H_0: \quad p_i = p_{i0}, \qquad i = 1, 2, \ldots, k \tag{2.1}$$

against

$$H_1: \quad \text{not all } p\text{'s are as given by } H_0,$$

where the p_{i0} are given numbers. The approximate chi-square procedure rejects H_0, if

$$y'_{k-1} > \chi^2_{k-1;1-\alpha}, \tag{2.2}$$

where

$$y'_{k-1} = \sum_{i=1}^{k} \frac{(x_i - np_{i0})^2}{np_{i0}}, \tag{2.3}$$

α is the significance level, and $\chi^2_{k-1;1-\alpha}$ is the quantile of order $1 - \alpha$ of the chi-square distribution with $k - 1$ degrees of freedom. To evaluate

$$\text{Power} = \Pr(Y'_{k-1} > \chi^2_{k-1;1-\alpha}), \tag{2.4}$$

we need the noncentrality parameter,

$$\lambda = n \sum_{i=1}^{k} \frac{(p_{i1} - p_{i0})^2}{p_{i0}}, \tag{2.5}$$

where p_{i1}, $i = 1, 2, \ldots, k$ are the p's for a specified alternative. Although Haynam, Govindarajulu, and Leone give one example, we will present another.

Example 1: Suppose that we wish to test a six-sided die for unbiasedness so that in (2.1) all six p_{i0}'s are $\frac{1}{6}$. If side six is loaded so that actually $p_{61} = \frac{1}{4}$ and the other five sides retain equal probability of occurrence, find the power for this alternative if the significance level is .05 and $n = 120$. Then find the minimum n needed to make the power at the alternative be at least .90.

Solution: Let p be the probability associated with sides one to five under the alternative. Then $5p + \frac{1}{4} = 1$, $p = p_{i1} = \frac{3}{20}$, $i = 1, 2, \ldots, 5$, $p_{61} = \frac{5}{20}$, $k = 6$, and

$$\lambda = n\left[5\frac{(\frac{3}{20} - \frac{1}{6})^2}{\frac{1}{6}} + \frac{(\frac{5}{20} - \frac{1}{6})^2}{\frac{1}{6}}\right] = \frac{n}{20} = 6.$$

From Table I of Haynam, Govindarajulu, and Leone [2], we read that the power is approximately .4329.

To achieve power of .90 we read from Table II that $\lambda = 16.469$. Hence we need $n/20 \geq 16.469$, $n \geq 329.4$ so that the minimum sample size is approximately 330.

3. Test of Independence

Let $\{X_{ij}, i = 1, 2, \ldots, r, j = 1, 2, \ldots, c\}$ be a multinomial random variable with parameters n, p_{ij}, i

$= 1, 2, \ldots, r, j = 1, 2, \ldots, c, \Sigma_{j=1}^{c} \Sigma_{i=1}^{r} p_{ij} = 1$. Suppose that we wish to test (independence)

$$H_0: \quad p_{ij} = p_{i.}p_{.j}, i = 1, 2, \ldots, r, j = 1, 2, \ldots, c$$

against (3.1)

H_1: not all the equations given under H_0 are satisfied,

where $p_{i.} = \Sigma_{j=1}^{c} p_{ij}, p_{.j} = \Sigma_{i=1}^{r} p_{ij}$. The approximate chi-square procedure rejects if

$$y'_{(r-1)(c-1)} > \chi^2_{(r-1)(c-1);1-\alpha}, \tag{3.2}$$

where

$$y'_{(r-1)(c-1)} = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(x_{ij} - t_{i.}t_{.j}/n)^2}{t_{i.}t_{.j}/n}, \tag{3.3}$$

and $t_{i.} = \Sigma_{j=1}^{c} x_{ij}, t_{.j} = \Sigma_{i=1}^{r} x_{ij}$. To evaluate power when H_0 is not true we need a noncentrality parameter λ. Chapman and Meng [1] have shown that for alternatives of the type

$$p_{ij} = p_{i.}p_{.j} + \frac{c_{ij}}{\sqrt{n}}, \quad \sum_{i=1}^{r} \sum_{j=1}^{c} c_{ij} = 0, \tag{3.4}$$

The noncentrality parameter is

$$\lambda = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{c_{ij}^2}{p_{i.}p_{.j}} - \sum_{i=1}^{r} \frac{c_{i.}^2}{p_{i.}} - \sum_{j=1}^{c} \frac{c_{.j}^2}{p_{.j}}, \tag{3.5}$$

where $c_{i.} = \Sigma_{j=1}^{c} c_{ij}, c_{.j} = \Sigma_{i=1}^{r} c_{ij}$. Unfortunately λ depends upon unknown parameters $p_{i.}, p_{.j}$. We could guess at these or use rough estimates. Another possibility is to choose the set which minimizes λ and power so that we have a lower bound on power. At any rate, we can evaluate power for any choices of these parameters.

Example 2: Suppose that we wish to test for independence of smoking and lung ailments on the basis of the results in the following table.

TABLE
Smoking Versus Lung Ailments

	Lung ailment	*No lung ailment*	*Totals*
Smokers	X_{11}	x_{12}	$t_{1.}$
Nonsmokers	x_{21}	x_{22}	$t_{2.}$
Total	$t_{.1}$	$t_{.2}$	n

For our alternative we select

$$p_{11} = p_{1.}p_{.1} + \Delta, \qquad p_{12} = p_{1.}p_{.2} - \Delta,$$

$$p_{21} = p_{2.}p_{.1} - \Delta, \qquad p_{22} = p_{2.}p_{.2} + \Delta.$$

Suppose that results suggest $p_{.1} = .25, p_{.2} = .75, p_{1.} = .6, p_{2.} = .4$. If we use a sample of size 200 and take $\Delta = .025$, find the power when $\alpha = .05$. Then, find the minimum n required to make the power at the alternative at least .80.

Solution: We identify $c_{11} = \Delta\sqrt{n} = c_{22}, c_{21} = -\Delta\sqrt{n} = c_{12}$, and see that $c_{1.} = c_{2.} = c_{.1} = c_{.2} = 0$, so that

$$\begin{aligned}\lambda &= n\Delta^2[1/p_{1.}p_{.1} + 1/p_{2.}p_{.1} + 1/p_{1.}p_{.2} + 1/p_{2.}p_{.2}]\\ &= n\Delta^2/p_{1.}p_{2.}p_{.1}p_{.2}\\ &= n(.025)^2/(.6)(.4)(.25)(.75)\\ &= n/72\\ &= .013889n\\ &= 2.78.\end{aligned}$$

With one degree of freedom and $\alpha = .05$, Table I in Haynam, Govindarajulu, and Leone [2] yields an approximate power of .385.

To achieve power .80 we read from Table II in [2] that $\lambda = 7.849$. Hence we need $.013889n \geq 7.849, n \geq 565.1$.

Had we used $p_{.1} = p_{.2} = p_{1.} = p_{2.} = .5$, the worst values, then $\lambda = 2$ when $n = 200$ and power is approximately .293. To make power $\geq .80$ requires $n \geq 784.9$.

4. Test of Homogeneity

Let $\{X_{ij}, j = 1, 2, \ldots, c\}$ be a multinomial random variable with parameters n_i, p_{ij} for $i = 1, 2, \ldots, r$. Suppose that we wish to test (homogeneity)

$$H_0: \quad p_{1j} = p_{2j} = \ldots = p_{rj} = p_j, j = 1, 2, \ldots, c$$

against (4.1)

H_1: not all the equations given under H_0 are satisfied.

The approximate chi-square procedure rejects if (3.2) is satisfied where $y'_{(r-1)(c-1)}$ is given by (3.3) with $t_{i.} = n_i, n = \Sigma_{i=1}^{r} n_i$. To evaluate power when H_0 is not true, we need a noncentrality parameter λ. Chapman and Meng [1] have shown that for alternatives of the type

$$p_{ij} = p_j + c_{ij}/\sqrt{n}, \quad \sum_{j=1}^{c} c_{ij} = 0, \tag{4.2}$$

the noncentrality parameter is

$$\lambda = \sum_{j=1}^{c} \frac{1}{p_j} \left[\sum_{i=1}^{r} c_{ij}^2 \frac{n_i}{n} - \left(\sum_{i=1}^{r} c_{ij} \frac{n_i}{n} \right)^2 \right]. \tag{4.3}$$

Again λ depends upon unknown parameters, and the comments of Section 3 can be repeated.

Example 3. Suppose that we wish to test the hypothesis of homogeneity with a three by three contingency table for which $n_1 = 300, n_2 = 200, n_3 = 100$. Consider the alternative that has all the diagonal p's increased by an amount Δ over their hypothesized values while the off diagonal p's are all decreased by the same amount, $\Delta/2$. Take $p_1 = .45, p_2 = .20, p_3 = .35$, and find the power if $\alpha = .05$ and $\Delta = .05$. Then, find the minimum n required to make the power at the alternative at least .75 if the n_i are to be in the same ratio.

Solution. We see that

$$c_{11}=\Delta\sqrt{n}, \quad c_{12} = -(\Delta\sqrt{n}/2), \; c_{13}=-(\Delta\sqrt{n}/2),$$
$$c_{21}=-(\Delta\sqrt{n}/2), c_{22}=\Delta\sqrt{n}, \quad c_{23}=-(\Delta\sqrt{n}/2),$$
$$c_{31}=-(\Delta\sqrt{n}/2), c_{32}=-(\Delta\sqrt{n}/2), \quad c_{33}=\Delta\sqrt{n},$$

and observe that $n_1/n = \frac{1}{2}$, $n_2/n = \frac{1}{3}$, $n_3/n = \frac{1}{6}$. Then

$$\begin{aligned}\lambda &= n\Delta^2/p_1\{1(\tfrac{1}{2}) + \tfrac{1}{4}(\tfrac{1}{3}) + \tfrac{1}{4}(\tfrac{1}{6}) \\ &\quad - [1(\tfrac{1}{2}) - \tfrac{1}{2}(\tfrac{1}{3}) - \tfrac{1}{2}(\tfrac{1}{6})]^2\} \\ &\quad + n\Delta^2/p_2\{\tfrac{1}{4}(\tfrac{1}{2}) + 1(\tfrac{1}{3}) + \tfrac{1}{4}(\tfrac{1}{6}) \\ &\quad - [-\tfrac{1}{2}(\tfrac{1}{2}) + 1(\tfrac{1}{3}) - \tfrac{1}{2}(\tfrac{1}{6})]^2\} \\ &\quad + n\Delta^2/p_3\{\tfrac{1}{4}(\tfrac{1}{2}) + \tfrac{1}{4}(\tfrac{1}{3}) + 1(\tfrac{1}{6}) \\ &\quad - [-\tfrac{1}{2}(\tfrac{1}{2}) - \tfrac{1}{2}(\tfrac{1}{3}) + 1(\tfrac{1}{6})]^2\} \\ &= n\Delta^2(.5625/p_1 + .5/p_2 + .3125/p_3) \\ &= 13n/1120 = .011607n = 6.96.\end{aligned}$$

With four degrees of freedom and $\alpha = .05$, Table I [2] yields an approximate power of .537.

To achieve power .75 we read from Table II that $\lambda = 10.722$. Hence we need $.011607n \geq 10.722$, $n \geq 923.7$. Taking $n = 924$ we need $n_1 = 462$, $n_2 = 308$, $n_3 = 154$.

Had we used $p_1 = 3/D$, $p_2 = 2\sqrt{2}/D$, $p_3 = \sqrt{5}/D$, where $D = 3 + 2\sqrt{2} + \sqrt{5}$, the worst values, then $\lambda = 6.10$ when $n = 600$, and power is approximately .477. To make power $\geq .75$ requires $n \geq 1055.1$.

[*Received August 5, 1976. Revised December 16, 1976.*]

References

[1] Chapman, Douglas G. and Meng, Rosa C., "The Power of Chi-Square Tests for Contingency Tables," *Journal of the American Statistical Association*, 61 (1966), 965–75.

[2] Haynam, G. E., Govindarajulu, Z., and Leone, F. C., "Tables of the Cumulative Non-Central Chi-Square Distribution," in *Selected Tables in Mathematical Statistics*, Vol. I, eds. H. L. Harter and D. B. Owen, Chicago: Markham Publishing Co., 1970.

On the Use of Limit Theorem Arguments in Economic Statistics

R. BARTELS*

Abstract

The heavy reliance on the normal distribution in the teaching of economic statistics is usually justified by appealing to the central limit theorem. Limit theorem arguments can, however, also lead to the nonnormal stable distributions. It is argued that the conditions required for nonnormal stable limits are, in fact, frequently satisfied by economic variables and in particular by disturbance terms in regressions. Normal theory results should, therefore, be used with caution, and it might be wise for economic statisticians to devote more attention to robust statistical methods.

KEY WORDS: Central limit theorem; Economic statistics; Economic variables; Normal distribution; Pareto's law; Regression disturbances; Stable distributions.

1. Normal Distributions in Economic Statistics

In teaching economic statistics, the normal distribution and normal sampling theory are relied on heavily. This approach is usually justified on intuitive grounds or by pointing to the role of the normal distribution in the central limit theorem. In economic statistics, perhaps more than in other areas of statistical application, recourse must be had to heuristic or theoretical arguments since many economic variables are not directly observable. In particular, this is true for the disturbance terms in the regression model. Furthermore, where direct observation is possible, the sample sizes are often too small to make distribution fitting a worthwhile exercise.

Following the work of Mandelbrot [9, 10, 11] it is now becoming apparent, however, that the arguments commonly put forward to obtain the normal distribution can, with slight modification, also lead to any member of a large class of distributions known as the stable distributions, of which the normal is a special case. For instance, the sum of independent, identically distributed random variables will tend to a nonnormal stable distribution if the summands are chosen from a suitable distribution.

Unfortunately, the nonnormal stable distributions have several awkward properties. For example, elementary expressions for the density functions are known in only a few special cases. The distributions are, therefore, usually described by their characteristic functions, making them accessible only to advanced students of statistics, but more important is the fact that the normal distribution is the only stable distribution with a finite variance. A course incorporating the whole class of stable distributions can, therefore, no longer present the variance as the universal measure of dispersion. Furthermore, common techniques in areas such as regression analysis and spectral analysis may no longer be formally valid.

The dilemma of having distributions whose use is easily justified on theoretical grounds, yet which present difficult problems in application, is causing considerable concern to research workers (see, e.g.,

* R. Bartels is with the Department of Economic Statistics, University of Sydney, Australia 2006.