



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Econometrics 123 (2004) 189–193

JOURNAL OF
Econometrics

www.elsevier.com/locate/econbase

Decompositions of Pearson's chi-squared test

Gianna Boero, Jeremy Smith, Kenneth F. Wallis*

Department of Economics, University of Warwick, Coventry CV4 7AL, UK

Accepted 29 October 2003

Abstract

We study the properties of Pearson's goodness-of-fit test, and show that the components-of-chi-squared or "Pearson analog" tests of Anderson (*J. Econometrics* 62 (1994) 265) are less generally applicable than was originally claimed.

© 2004 Elsevier B.V. All rights reserved.

JEL classification: C12; C14

Keywords: Goodness-of-fit; Tests of distributional assumptions; Pearson's chi-squared test

1. Introduction

Many areas of economics require the comparison of distributions. Analysis of the distribution of income has a history that extends over more than a century, while a new area of application is the evaluation of density forecasts. The statistical problem in all such applications is to assess the degree of correspondence or goodness of fit between observed data and a hypothesised distribution. In many applications a nonparametric procedure is appropriate, the best known being Pearson's classical chi-squared test.

Anderson (1994) presents a rearrangement of the chi-squared goodness-of-fit statistic to provide additional information about the nature of departures from the hypothesised distribution, in respect of specific features of the empirical distribution such as its location, scale and skewness. An application of this components-of-chi-squared or "Pearson analog" test to the comparison of income distributions is given by Anderson (1996), and it is also used in density forecast evaluation by Wallis (2003).

A formal derivation of the components test is presented in this article, and it is shown that some of Anderson's claims for the generality of the test are not correct.

* Corresponding author. Tel.: +44-24-7652-3055; fax: +44-24-7652-3032.

E-mail address: k.f.wallis@warwick.ac.uk (K.F. Wallis).

This fact explains some of the unexpected simulation results reported by Anderson (2001).

2. The chi-squared goodness-of-fit test and its components

Pearson’s classical test proceeds by dividing the range of the variable into k mutually exclusive classes and comparing the observed frequencies with which outcomes fall in these classes with the expected frequencies under the hypothesised distribution. With observed class frequencies $n_i, i = 1, \dots, k, \sum n_i = n$ and class probabilities under the null $p_i > 0, i = 1, \dots, k, \sum p_i = 1$, the test statistic is

$$X^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}.$$

This statistic has a limiting χ^2 distribution with $k - 1$ degrees of freedom (denoted χ^2_{k-1}) if the hypothesised distribution is correct.

The asymptotic distribution of the test statistic rests on the asymptotic k -variate normality of the multinomial distribution of the observed frequencies. Placing these in the $k \times 1$ vector \mathbf{x} , under the null hypothesis this has mean vector $\boldsymbol{\mu} = (np_1, np_2, \dots, np_k)'$ and covariance matrix

$$\mathbf{V} = n \begin{bmatrix} p_1(1 - p_1) & -p_1 p_2 & \dots & -p_1 p_k \\ -p_2 p_1 & p_2(1 - p_2) & \dots & -p_2 p_k \\ \vdots & \vdots & \ddots & \vdots \\ -p_k p_1 & -p_k p_2 & \dots & p_k(1 - p_k) \end{bmatrix}.$$

The covariance matrix is singular, with rank $k - 1$: note that each column (row) sums to zero. Defining its generalised inverse \mathbf{V}^- , the limiting distribution of the quadratic form $(\mathbf{x} - \boldsymbol{\mu})' \mathbf{V}^- (\mathbf{x} - \boldsymbol{\mu})$ is then χ^2_{k-1} (Pringle and Rayner, 1971, p.78).

In his derivation and application of the components test Wallis (2003) considers the case of equiprobable classes, as is often recommended to improve the power properties of the overall test. In this case $p_i = 1/k, i = 1, \dots, k$ and

$$\mathbf{V} = (n/k)[\mathbf{I} - \mathbf{e}\mathbf{e}'/k],$$

where \mathbf{e} is a $k \times 1$ vector of ones. Since the matrix in square brackets is symmetric and idempotent it coincides with its generalised inverse, and the chi-squared statistic is equivalently written

$$\begin{aligned} X^2 &= \sum_{i=1}^k \frac{(n_i - n/k)^2}{(n/k)} \\ &= (\mathbf{x} - \boldsymbol{\mu})' [\mathbf{I} - \mathbf{e}\mathbf{e}'/k] (\mathbf{x} - \boldsymbol{\mu}) / (n/k) \end{aligned}$$

(note that $e'(x - \mu) = 0$). There exists a $(k - 1) \times k$ transformation matrix A , such that

$$AA' = I, \quad A'A = [I - ee'/k]$$

(Rao and Rao, 1998, p.252). Hence defining $y = A(x - \mu)$ the statistic can be written as an alternative sum of squares

$$X^2 = y'y/(n/k),$$

where the $k - 1$ components $y_i^2/(n/k)$ are independently distributed as χ^2_1 under the null hypothesis.

To construct the matrix A we consider Hadamard matrices, which are square matrices whose elements are 1 or -1 and whose columns are orthogonal: $H'H = kI$. For k equal to a power of 2, we begin with the basic Hadamard matrix

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

and form Kronecker products

$$H_4 = H_2 \otimes H_2, \quad H_8 = H_4 \otimes H_2$$

and so on. Deleting the first row of 1s and dividing by \sqrt{k} then gives the required matrix A , that is, H is partitioned as

$$H = \begin{bmatrix} e' \\ \sqrt{k}A \end{bmatrix}.$$

With $k = 4$, and reordering the rows, we have

$$A = \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

and Anderson (1994) suggests that the three components of the test relate in turn to departures from the null distribution with respect to location, scale and skewness. Taking $k = 8$ allows a fourth component related to kurtosis to appear, although the remaining three components are difficult to relate to characteristics of the distribution:

$$A = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}.$$

Note that the first three components when $k = 8$ coincide with the (only) three components when $k = 4$.

Returning to the general case of unequal class probabilities we write the covariance matrix as

$$V = n[\mathbf{P} - \mathbf{p}\mathbf{p}'],$$

where $\mathbf{P} = \text{diag}(p_1, p_2, \dots, p_k)$ and $\mathbf{p} = (p_1, p_2, \dots, p_k)'$. Tanabe and Sagae (1992) give the result that

$$[\mathbf{P} - \mathbf{p}\mathbf{p}']^{-1} = (\mathbf{I} - \mathbf{e}\mathbf{e}'/k)\mathbf{P}^{-1}(\mathbf{I} - \mathbf{e}\mathbf{e}'/k)$$

hence the test statistic can again be written in terms of the transformed variables $\mathbf{y} = \mathbf{A}(\mathbf{x} - \boldsymbol{\mu})$, as

$$X^2 = \mathbf{y}'\mathbf{A}\mathbf{P}^{-1}\mathbf{A}'\mathbf{y}/n.$$

The covariance matrix of \mathbf{y} is $E(\mathbf{y}\mathbf{y}') = \mathbf{A}\mathbf{V}\mathbf{A}'$, whose inverse $\mathbf{A}\mathbf{P}^{-1}\mathbf{A}'/n$ appears in the above quadratic form: this can be checked by multiplying out, noting that $\mathbf{e}'[\mathbf{P} - \mathbf{p}\mathbf{p}'] = \mathbf{0}$ and $\mathbf{A}\mathbf{P}^{-1}\mathbf{p} = \mathbf{A}\mathbf{e} = \mathbf{0}$. The diagonal elements give the variance of y_i , $i = 1, \dots, k - 1$, as $\sigma_i^2 = n(1 - k(\mathbf{a}_i'\mathbf{p})^2)/k$ where \mathbf{a}_i' is the i th row of \mathbf{A} , corresponding to expressions given by Anderson (1994, p.267) for the first four elements when $k = 8$. However, the covariance matrix is not diagonal, hence the test statistic does not reduce to a simple sum of squares in the case of unequal p_i . Equivalently, terms of the form y_i^2/σ_i^2 are not independently distributed as χ^2 in this general case, contrary to Anderson's claim. It remains the case, however, that the marginal distribution of an individual y_i^2/σ_i^2 is χ_1^2 .

Anderson (1994) goes on to argue that the power of the component tests to detect departures of the alternative distribution from the null distribution depends on the closeness of the intersection points of the two distributions to the location of the class boundaries. In the equiprobable case the class boundaries are the appropriate quantiles, and moving them to improve test performance clearly alters the class probabilities. In practical examples this is often done in such a way that the resulting class probabilities are symmetric, that is, $p_1 = p_k$, $p_2 = p_{k-1}, \dots$, which is an interesting special case. Now the odd-numbered and even-numbered components of the chi-squared statistic are orthogonal. Hence focusing only on location and scale, for example, a joint test can be based on $y_1^2/\sigma_1^2 + y_2^2/\sigma_2^2$, distributed as χ_2^2 under the null; similarly, focusing only on skewness and kurtosis, a joint test can be based on $y_3^2/\sigma_3^2 + y_4^2/\sigma_4^2$. As in the more general case, however, these two test statistics are not independent, equivalently their sum is not distributed as χ_4^2 .

This observation helps to explain some of the simulation results of Anderson (2001), who includes components-of-chi-squared tests in a comparison of a range of tests for location and scale problems. With four non-equiprobable classes whose boundaries are placed symmetrically around the mean, the class probabilities are symmetric in the sense of the previous paragraph if the null distribution is symmetric, but not otherwise. The components considered are the individual location and scale components, each of which is χ_1^2 in general; a "joint" test based on their sum, as in the above example, which is χ_2^2 in the symmetric case but not in general; and a "general" test defined as

the sum of the first three components, which is no longer χ_3^2 even in the symmetric but non-equiprobable case. Anderson (2001, p.25) reports that the power and consistency properties of the various components tests are good, “with the exception of the general and joint tests under the asymmetric distribution.” The above analysis shows that in these circumstances the null distributions are not χ^2 , contrary to what is assumed in the simulation study, hence the problems noted are the result of comparing test statistics to inappropriate critical values.

Our own simulations (Boero et al., 2002) show that the use of unequal class probabilities can indeed improve the power of an individual component test, although once attention has been focused in this way many researchers might alternatively choose to use a moment-based test.

Acknowledgements

The helpful comments of Gordon Anderson and two anonymous referees are gratefully acknowledged.

References

- Anderson, G., 1994. Simple tests of distributional form. *Journal of Econometrics* 62, 265–276.
- Anderson, G., 1996. Nonparametric tests of stochastic dominance in income distributions. *Econometrica* 64, 1183–1193.
- Anderson, G., 2001. The power and size of nonparametric tests for common distributional characteristics. *Econometric Reviews* 20, 1–30.
- Boero, G., Smith, J., Wallis, K.F., 2002. The properties of some goodness-of-fit tests. Economic Research Paper No 653, University of Warwick.
- Pringle, R.M., Rayner, A.A., 1971. *Generalized Inverse Matrices with Applications to Statistics*. Charles Griffin, London.
- Rao, C.R., Rao, M.B., 1998. *Matrix Algebra and its Applications to Statistics and Econometrics*. World Scientific, Singapore.
- Tanabe, K., Sagae, M., 1992. An exact Cholesky decomposition and the generalized inverse of the variance–covariance matrix of the multinomial distribution, with applications. *Journal of the Royal Statistical Society B* 54, 211–219.
- Wallis, K.F., 2003. Chi-squared tests of interval and density forecasts, and the Bank of England’s fan charts. *International Journal of Forecasting* 19, 165–175.