Available online at www.sciencedirect.com
SCIENCE DIRECT®

Computational Statistics & Data Analysis 48 (2005) 803–808

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

A normal approximation for the chi-square distribution

Luisa Canal*

Dipartimento di Scienze della Cognizione e della Formazione, University of Trento, Via Matteo del Ben, 38068 Rovereto (TN), Italy

Received 22 December 2003; received in revised form 1 April 2004; accepted 1 April 2004

Abstract

An accurate normal approximation for the cumulative distribution function of the chi-square distribution with n degrees of freedom is proposed. This considers a linear combination of appropriate fractional powers of chi-square. Numerical results show that the maximum absolute error associated with the new transformation is substantially lower than that found for other power transformations of a chi-square random variable for all the degrees of freedom considered ($1 \leqslant n \leqslant 1000$).

Keywords: Chi-square distribution; Maximum absolute error; Normal approximation; Power transformation

1. Introduction

Power transformations of the chi-square random variable can be employed to improve its approximate normality. Among these, the best known are the Fisher's square root transformation $(2n\chi^2)^{1/2}$ (Fisher, 1922) and the third root transformation $(\chi^2/n)^{1/3}$ (Wilson and Hilferty, 1931), where n stands for the number of degrees of freedom (d.f.). Also the fourth root transformation $(\chi^2/n)^{1/4}$ has a distribution which is close to normality for all degrees of freedom (Cressie and Hawkins, 1980; Hawkins and Wixley, 1986).

The third root transformation produces a closer approximation to normality than the square root transformation (Merrington, 1941; Goldberg and Levine, 1946; Zar, 1974).

* Tel.: +39461882111; fax: +39464483554.
E-mail address: luisa.canal@unitn.it (L. Canal).

0167-9473/$ - see front matter
doi:10.1016/j.csda.2004.04.001

For 1 and 2 degrees of freedom the fourth root is superior to the third root; however, for larger numbers of degrees of freedom the third root is better.

Goria (1992) proposed a linear combination of the square and of the fourth root transformations, which are the only that, within the family of power transformations, have the property that the Pearson's kurtosis index is zero to an appropriate order. However, the results were substantially similar to those obtained using the third root transformation, with clear benefits only for high values of the degrees of freedom. In this paper a more accurate linear combination of power transformations of the chi-square random variable will be proposed, that fits well regardless of the number of degrees of freedom.

2. Power approximation

The basic idea underlying the transformation was to combine different powers of χ^2 in order to have the Pearson's kurtosis index equal to zero to an appropriate order, and the symmetry near to zero also for a small number of degrees of freedom. A preliminary search suggested that a linear combination of powers $\frac{1}{6}, \frac{1}{3}, \frac{1}{2}$ would be suitable. So, the following linear combination was considered:

$$L = a\left[\left(\frac{\chi^2}{n}\right)^{1/6} + b\left(\frac{\chi^2}{n}\right)^{1/3} + c\left(\frac{\chi^2}{n}\right)^{1/2}\right],$$

where the constants a, b, c must be determined in such a way that both the skewness and the kurtosis of L tend rapidly to those of the normal distribution. It is clear from the above equation that the constant a affects only the variance, but neither skewness nor kurtosis, so it was set equal to one. An admissible solution was found for $b = -\frac{1}{2}$ and $c = \frac{1}{3}$ (see the appendix for details); therefore the linear combination proposed as normal approximation for the chi-square distribution is the following:

$$L = \left(\frac{\chi^2}{n}\right)^{1/6} - \frac{1}{2}\left(\frac{\chi^2}{n}\right)^{1/3} + \frac{1}{3}\left(\frac{\chi^2}{n}\right)^{1/2}.$$

The resulting linear combination does have an appealing feature as it can be expressed as

$$6L = 6\left(\frac{\chi^2}{n}\right)^{1/6} - 3\left(\frac{\chi^2}{n}\right)^{1/3} + 2\left(\frac{\chi^2}{n}\right)^{1/2}.$$

The skewness and kurtosis coefficients of L are, respectively, $\frac{1}{27\sqrt{2}n^{3/2}} + O(n^{-2})$ and $O(n^{-2})$, while its expected value is

$$E[L] = \frac{5}{6} - \frac{1}{9n} - \frac{7}{648n^2} + \frac{25}{2187n^3} + O(n^{-4}) \tag{1}$$

and the variance is

$$\mathrm{Var}[L] = \frac{1}{18n} + \frac{1}{162n^2} - \frac{37}{11664n^3} + O(n^{-4}). \tag{2}$$