Computational Statistics and Data Analysis 53 (2009) 853–856

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Short communication

A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables

Huan Liu [a], Yongqiang Tang [b,*], Hao Helen Zhang [c]

[a] Center for Biostatistics in AIDS Research, Harvard School of Public Health, Boston, MA, United States

[b] SUNY Downstate Medical Center, Brooklyn, NY, United States

[c] Department of Statistics, North Carolina State University, Raleigh, NC, United States

ARTICLE INFO

Article history:
Received 8 June 2008
Received in revised form 27 November 2008
Accepted 28 November 2008
Available online 6 December 2008

ABSTRACT

This note proposes a new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables. The unknown parameters are determined by the first four cumulants of the quadratic forms. The proposed method is compared with Pearson's three-moment central χ^2 approximation approach, by means of numerical examples. Our method yields a better approximation to the distribution of the non-central quadratic forms than Pearson's method, particularly in the upper tail of the quadratic form, the tail most often needed in practical work.

1. Introduction

Consider the quadratic form $Q(X) = X'AX$, where X follows an n-dimensional multivariate normal distribution with mean vector μ_x and non-singular variance matrix Σ, and A is an $n \times n$ symmetric and non-negative definite matrix. The problem of estimating the tail probability of $Q(X)$

$$\Pr(Q(X) > t) \tag{1}$$

arises in many statistical applications, such as the power analysis of a test procedure if the (asymptotic) distribution of the test statistic (e.g., Pearson's chi-square statistic) takes the form of $Q(X)$.

Computing (1) is usually not straightforward except in some special cases. An approximation which uses numerical integration to invert the characteristic function of $Q(X)$ is derived by Imhof (1961); this can be made very accurate, and bounds on accuracy can be found. A similar formula based on numerical inversion of the characteristic function was given in Davies (1980). Alternatively, Kotz et al. (1967) expressed (1) as an infinite series in central chi-square distribution functions, which was programmed in Fortran by Sheil and O'Muircheartaigh (1977); see also Section 4 of this note. Kuonen (1999) proposed a saddle point approximation to (1). Other majority approximation approaches are based on the moments of $Q(X)$; for a summary, see Solomon and Stephens (1977, 1978) and the references therein.

This note proposes approximating the distribution of $Q(X)$ using a $\chi_l^2(\delta)$ distribution, where the degrees of freedom l and the non-central parameter δ are determined by the first four cumulants of $Q(X)$. Its relationship with Pearson's three-moment central χ^2 approximation approach (Imhof, 1961) will be discussed. Like other moment-based approaches, our method does not involve inverting a matrix or calculating the eigenvalues of a matrix. It is easy to implement since the chi-square distribution function is available in nearly all statistical packages.

* Corresponding author.
E-mail address: Yongqiang_ytang@yahoo.com (Y. Tang).

0167-9473/$ – see front matter
doi:10.1016/j.csda.2008.11.025

2. A new non-central chi-square approximation for $Q(X)$

2.1. Cumulants of the quadratic form

Let P be a $n \times n$ orthonormal matrix which converts $B = \Sigma^{1/2} A \Sigma^{1/2}$ to the diagonal form $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = PBP'$, where $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. Let $m = \mathrm{rank}(A)$. If $m < n$, then $\lambda_{m+1} = \cdots = \lambda_n = 0$. Since $Y = P\Sigma^{-1/2} X$ is normally distributed with mean $\mu_y = P\Sigma^{-1/2}\mu_x$ and variance I_n, $Q(X)$ can be expressed as a weighted sum of chi-square variables:

$$Q(X) = X'AX = Y'\Lambda Y = \sum_{i=1}^{n} \lambda_i \chi^2_{h_i}(\delta_i) = \sum_{i=1}^{m} \lambda_i \chi^2_{h_i}(\delta_i), \tag{2}$$

where $h_i = 1$, $\delta_i = \mu_{yi}^2$, and μ_{yi} is the ith component of μ_y for $i = 1, \ldots, m$. The cumulant generating function of $Q(X)$ is given by (Imhof, 1961)

$$K(t) = \frac{1}{2} \sum_{i=1}^{m} h_i \log(1 - 2t\lambda_i) + \sum_{i=1}^{m} \frac{\delta_i \lambda_i t}{1 - 2\lambda_i t}.$$

The formula for the kth cumulant of $Q(X)$ is

$$\kappa_k = 2^{k-1}(k-1)! \left(\sum_{i=1}^{m} \lambda_i^k h_i + k \sum_{i=1}^{m} \lambda_i^k \delta_i \right). \tag{3}$$

It is unnecessary to calculate the λ_i's and δ_i's explicitly since

$$\sum_{i=1}^{m} \lambda_i^k h_i = \mathrm{trace}(\Lambda^k) = \mathrm{trace}((PBP')^k) = \mathrm{trace}(B^k) = \mathrm{trace}((A\Sigma)^k)$$

and

$$\sum_{i=1}^{m} \lambda_i^k \delta_i = \mu_y' \Lambda^k \mu_y = \mu_x' \Sigma^{-1/2} P'(PBP')^k P\Sigma^{-1/2} \mu_x = \mu_x'(A\Sigma)^{k-1} A\mu_x.$$

The mean, standard deviation, skewness and kurtosis of $Q(X)$ are given by

$$\mu_Q = \kappa_1 = c_1, \qquad \sigma_Q = \sqrt{\kappa_2} = \sqrt{2c_2}, \qquad \beta_1 = \frac{\kappa_3}{\kappa_2^{3/2}} = \sqrt{8}s_1, \qquad \beta_2 = \frac{\kappa_4}{\kappa_2^2} = 12s_2,$$

where $c_k = \sum_{i=1}^{m} \lambda_i^k h_i + k \sum_{i=1}^{m} \lambda_i^k \delta_i$, $s_1 = c_3/c_2^{3/2}$ and $s_2 = c_4/c_2^2$.

2.2. The new approach

We propose using a non-central $\chi_l^2(\delta)$ distribution to approximate the distribution of

$$Q(X) = \sum_{i=1}^{m} \lambda_i \chi^2_{h_i}(\delta_i).$$

We shall not impose the restriction $h_i = 1$. The tail probability (1) is approximated by

$$\begin{aligned} \Pr(Q(X) > t) &= \Pr\left(\frac{Q(X) - \mu_Q}{\sigma_Q} > t^* \right) \\ &\approx \Pr\left(\frac{\chi_l^2(\delta) - \mu_\chi}{\sigma_\chi} > t^* \right) = \Pr\left(\chi_l^2(\delta) > t^*\sigma_\chi + \mu_\chi \right) \end{aligned} \tag{4}$$

where $t^* = (t - \mu_Q)/\sigma_Q$, $\mu_\chi = \mathrm{E}(\chi_l^2(\delta)) = l + \delta$, $\sigma_\chi = \sqrt{\mathrm{var}(\chi_l^2(\delta))} = \sqrt{2}\,a$, and $a = \sqrt{l + 2\delta}$. The parameters δ and l are determined so that the skewnesses of $Q(X)$ and $\chi_l^2(\delta)$ are equal and the difference between the kurtoses of $Q(X)$ and $\chi_l^2(\delta)$ is minimized. The solution to (4) will be given in (5) and (6) below.

The skewness of $\chi_l^2(\delta)$ is $\sqrt{8}(a^2 + \delta)/a^3$, and that of $Q(X)$ is $\sqrt{8}s_1$. Since the two skewnesses are equal, we obtain $\delta = s_1 a^3 - a^2$. Note that $\delta \geq 0$. Thus $a \geq 1/s_1$. The difference in kurtosis of $Q(X)$ and $\chi_l^2(\delta)$ is

$$\Delta_K = \left| \frac{12(l + 4\delta)}{(l + 2\delta)^2} - 12s_2 \right| = 12 \left| \left(\frac{1}{a} - s_1 \right)^2 + s_2 - s_1^2 \right|.$$

If $s_1^2 > s_2$, the minimum value of Δ_K is 0 when $s_1^2 - s_2 = (1/a - s_1)^2$. Thus

$$a = 1/\left(s_1 - \sqrt{s_1^2 - s_2}\right), \qquad \delta = s_1 a^3 - a^2 \quad \text{and} \quad l = a^2 - 2\delta. \tag{5}$$

If $s_1^2 \le s_2$, Δ_K is minimized when $a = 1/s_1$. The final solution is given by

$$\delta = s_1 a^3 - a^2 = 0 \quad \text{and} \quad l = a^2 - 2\delta = 1/s_1^2 = c_2^3/c_3^2. \tag{6}$$

We can show that $l > 0$ and $\delta > 0$ in (5) and that $l > 0$ and $\delta = 0$ in (6).

2.3. Relation to Pearson's central χ^2 approximation

Our method is closely related to Pearson's three-moment central χ^2 approximation approach. The Pearson (1959) approach was originally developed for approximating a non-central $\chi_l^2(\delta)$ distribution by a central $\chi_{l^*}^2$ distribution:

$$\Pr\left(\frac{\chi_l^2(\delta) - \mu_\chi}{\sigma_\chi} > t^*\right) \approx \Pr\left(\frac{\chi_{l^*}^2 - l^*}{\sqrt{2l^*}} > t^*\right), \tag{7}$$

where $\chi_l^2(\delta)$ and $\chi_{l^*}^2$ have equal skewnesses. Imhof (1961) extended Pearson's method to the non-negative quadratic form $Q(X)$:

$$\begin{aligned}\Pr(Q(X) > t) &= \Pr\left(\frac{Q(X) - \mu_Q}{\sigma_Q} > t^*\right)\\ &\approx \Pr\left(\frac{\chi_{l^*}^2 - l^*}{\sqrt{2l^*}} > t^*\right) = \Pr\left(\chi_{l^*}^2 > l^* + t^*\sqrt{2l^*}\right),\end{aligned} \tag{8}$$

where $l^* = 1/s_1^2$ is determined so that $Q(X)$ and $\chi_{l^*}^2$ have equal skewnesses. Although it is difficult to assess in a mathematical way, the accuracy of Pearson's approximation has been demonstrated in many empirical studies (Johnson, 1959; Imhof, 1961; Solomon and Stephens, 1977; Kuonen, 1999).

The accuracy of our approach is guaranteed since the approximation error of (4) is bounded by the sum of the approximation errors of (7) and (8):

$$\begin{aligned}&\left|\Pr\left(\frac{Q(X) - \mu_Q}{\sigma_Q} > t^*\right) - \Pr\left(\frac{\chi_l^2(\delta) - \mu_\chi}{\sigma_\chi} > t^*\right)\right|\\ &\le \left|\Pr\left(\frac{Q(X) - \mu_Q}{\sigma_Q} > t^*\right) - \Pr\left(\frac{\chi_{l^*}^2 - l^*}{\sqrt{2l^*}} > t^*\right)\right| + \left|\Pr\left(\frac{\chi_{l^*}^2 - l^*}{\sqrt{2l^*}} > t^*\right) - \Pr\left(\frac{\chi_l^2(\delta) - \mu_\chi}{\sigma_\chi} > t^*\right)\right|.\end{aligned}$$

Note that this bound is quite conservative. In the next section, we will use a simulation study to show that the actual approximation error of our method is usually much smaller than that of Pearson's method.

Pearson's method essentially uses the central χ^2 approximation by requiring the match of the third-order moments. The proposed method considers a broader class of non-central $\chi^2(\delta)$ distributions, and searches for the distribution which not only matches the third-order moments but also has the best match of the fourth-order moments. The better high-order moment approximation often improves the tail probability approximation in practice. Our method is equivalent to Pearson's approach when $s_1^2 \le s_2$. For central quadratic forms ($\delta_i = 0$ for $i = 1, \ldots, m$), the Cauchy–Schwarz inequality implies that $s_1^2 \le s_2$. Our method produces the exact tail probability when $\lambda_1 = \cdots = \lambda_m$; under this assumption, $Q(X)/\lambda_1$ follows a non-central chi-square distribution.

3. Numerical examples

This section numerically compares our method with Pearson's three-moment χ^2 approximation method. The tail probabilities of chi-square distributions are evaluated using the CDF function in SAS. Since $Q(X)$ has the same distribution as $Q(z) = \sum_{i=1}^m \sum_{j=1}^{h_i} \lambda_i \left(z_{ij} + \sqrt{\delta_i/h_i}\right)^2$, where the z_{ij}'s are independent standard normal variables, an application of Eq. (144) of Kotz et al. (1967) yields an exact formula for computing (1):

$$\Pr(Q(X) > t) = \sum_{k=0}^{\infty} c_k \Pr\left(\chi_{2k+\tilde{h}}^2 > \frac{t}{\beta}\right), \tag{9}$$

for any $0 < \beta \le \min(\lambda_1, \ldots, \lambda_m)$, where $\tilde{h} = \sum_{i=1}^m h_i$, $\gamma_i = 1 - \beta/\lambda_i$, $g_k = \left[\sum_{i=1}^m h_i \gamma_i^k + k \sum_{i=1}^m \delta_i \gamma_i^{k-1}(1 - \gamma_i)\right]/2$, $d = \prod_{i=1}^m (\beta/\lambda_i)^{h_i/2}$, $e = \sum_{i=1}^m \delta_i$, $c_0 = d\exp(-e/2)$, and $c_k = k^{-1}\sum_{r=0}^{k-1} g_{k-r} c_r$ for $k \ge 1$. If the series is truncated after N terms, the truncation error is bounded by

$$\sum_{k=N+1}^{\infty} c_k \Pr\left(\chi_{2k+\tilde{h}}^2 > \frac{t}{\beta}\right) \le \sum_{k=N+1}^{\infty} c_k = 1 - \sum_{k=1}^{N} c_k.$$

Table 1
Probability that the quadratic form exceeds t. P_1: exact values with accuracy 0.000001; P_2: the proposed non-central χ^2 approximation; E_2: the absolute error for the proposed method; P_3: Pearson's central χ^2 approximation; E_3: the absolute error for Pearson's approximation.

Quadratic form	t	P_1	P_2	E_2	P_3	E_3	E_3/E_2
$Q_1 = .5\chi_1^2(1) + .4\chi_2^2(.6)$	2	.457461	.457753	.000292	.458967	.001507	5.2
$+ .1\chi_1^2(.8)$	6	.031109	.031079	.000030	.030929	.000180	6.0
	8	.006885	.006883	.000002	.006908	.000023	8.1
$Q_2 = .7\chi_1^2(6) + .3\chi_1^2(2)$	1	.954873	.955046	.000173	.951516	.003357	19.0
	6	.407565	.407587	.000022	.408359	.000794	36.2
	15	.022343	.022340	.000003	.022294	.000049	14.2
$Q_3 = .995\chi_1^2(1)$	2	.347939	.347946	.000007	.357398	.009459	1342.3
$+ .005\chi_2^2(1)$	8	.033475	.033475	.000000	.032348	.001132	3677.2
	12	.006748	.006748	.000000	.006807	.000059	2132.3
$Q_4 = .35\chi_1^2(6) + .15\chi_1^2(2)$	3.5	.956318	.956315	.000003	.955961	.000357	100.3
$+ .35\chi_6^2(6) + .15\chi_2^2(2)$	8	.415239	.415248	.000009	.415273	.000034	4.0
	13	.046231	.046228	.000002	.046085	.000146	59.8

In Table 1, all quadratic forms satisfy $s_1^2 > s_2$. The exact tail probability is calculated using (9). As expected, our method yields better approximation to the distribution of $Q(X)$ than Pearson's method, especially in the upper tail of $Q(X)$. In all examples given in Table 1, the absolute errors for Pearson's approximation are at least three times as large as those for the proposed method.

References

Davies, R.B., 1980. Algorithm as 155: The distribution of a linear combination of χ^2 random variables. Applied Statistics 29, 323–333.
Imhof, J.P., 1961. Computing the distribution of quadratic forms in normal variables. Biometrika 48, 419–426.
Johnson, N.L., 1959. On an extension of the connection between Poisson and χ^2 distributions. Biometrika 46, 352–363.
Kotz, S., Johnson, N.L., Boyd, D.W., 1967. Series representations of distributions of quadratic forms in normal variables ii. non-central case. The Annals of Mathematical Statistics 38, 838–848.
Kuonen, D., 1999. Saddlepoint approximations for distributions of quadratic forms in normal variables. Biometrika 86, 929–935.
Pearson, E.S., 1959. Note on an approximation to the distribution of non-central χ^2. Biometrika 46, 364.
Sheil, J., O'Muircheartaigh, I., 1977. Algorithm as 106: The distribution of non-negative quadratic forms in normal variables. Applied Statistics 26, 92–98.
Solomon, H., Stephens, M.A., 1977. Distribution of a sum of weighted chi-square variables. Journal of the American Statistical Association 72, 881–885.
Solomon, H., Stephens, M.A., 1978. Approximations to density functions using Pearson curves. Journal of the American Statistical Association 73, 153–160.