Computing the distribution of quadratic forms: further comparisons between the Liu-Tang-Zhang approximation and exact methods*

Pierre DUCHESNE†

Département de mathématiques et statistique

Université de Montréal

Pierre LAFAYE DE MICHEAUX

Département de mathématiques et statistique

Université de Montréal

November 19, 2009

Abstract

Liu, Tang and Zhang (*Computational Statistics & Data Analysis*, 853-856, 53, 2009) proposed a chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables. To approximate the distribution of interest, they used a non-central chi-square distribution, where the degrees of freedom and the non-centrality parameter were calculated using the first four cumulants of the quadratic form. Numerical examples were encouraging, suggesting that the approximation was particularly accurate in the upper tail of the distribution. We present here additional empirical evidence, comparing the Liu-Tang-Zhang's four-moment chi-square approximation with exact methods. While the moment-based method is interesting because of its simplicity, we demonstrate that it should be used with care in practical work, since numerical examples suggest that significant differences may occur between that method and exact methods, even in the upper tail of the distribution.

Key words and phrases: Quadratic forms; Imhof's method; Farebrother's algorithm; Non-central chi-square variables; Pearson's three-moment chi-square approximation.

---

*Abbreviated title: "Computing the distribution of quadratic forms".

Corresponding author: Pierre Duchesne, Université de Montréal; Département de mathématiques et de statistique; C.P. 6128 Succursale Centre-Ville; Montréal, Québec H3C 3J7; Canada. tel: (514) 343-7267; fax: (514) 343-5700; e-mail: duchesne@dms.umontreal.ca.

†This work was supported by a grant from the Natural Science and Engineering Research Council of Canada.

1. INTRODUCTION

Let $\mathbf{X} = (X_1, \ldots, X_p)^\top$ be a multivariate normal random vector $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with mean $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)^\top$ and positive definite covariance matrix $\boldsymbol{\Sigma}$. Consider the quadratic form $\mathcal{Q} = \mathbf{X}^\top \mathbf{A} \mathbf{X}$, where $\mathbf{A}$ represents a $p \times p$ symmetric and non-negative definite matrix. A problem of interest is to evaluate the probability

$$Pr(\mathcal{Q} > q), \tag{1}$$

where q is a scalar.

In the simplest case $\boldsymbol{\Sigma} = \mathbf{A} = \mathbf{I}_p$, where $\mathbf{I}_p$ denotes the $p \times p$ identity matrix, $\mathcal{Q}$ represents a non-central chi-square random variable with p degrees of freedom and non-centrality parameter $\delta = \boldsymbol{\mu}^\top \boldsymbol{\mu}$. In the general case, let $\mathbf{P}$ be such that $\mathbf{P}\mathbf{P}^\top = \mathbf{I}_p$ and that diagonalizes $\mathbf{C}\mathbf{A}\mathbf{C}^\top$, that is $\mathbf{P}\mathbf{C}\mathbf{A}\mathbf{C}^\top\mathbf{P}^\top = \mathbf{D} = \text{diag}(\lambda_1, \ldots, \lambda_p)$. The matrix $\mathbf{C}$ corresponds to the Cholesky decomposition of $\boldsymbol{\Sigma}$ and satisfies the relation $\mathbf{C}^\top\mathbf{C} = \boldsymbol{\Sigma}$. We assume that $\lambda_1 \geq \ldots \geq \lambda_r > 0$ and $\lambda_{r+1} = \ldots \lambda_p = 0$; thus the rank of $\mathbf{A}$ is $r = \text{rank}(\mathbf{A})$. Let $\mathbf{Y} = \mathbf{P}(\mathbf{C}^\top)^{-1}\mathbf{X}$ and $\boldsymbol{\nu} = \mathbf{P}(\mathbf{C}^\top)^{-1}\boldsymbol{\mu}$. Thus the distribution of $\mathbf{Y}$ is $\mathcal{N}_p(\boldsymbol{\nu}, \mathbf{I}_p)$, and it follows that the quadratic form $\mathcal{Q}$ can be expressed as a weighted sum of chi-square random variables:

$$\mathcal{Q} = \mathbf{X}^\top\mathbf{A}\mathbf{X} = \mathbf{Y}^\top\mathbf{D}\mathbf{Y} = \sum_{i=1}^{r} \lambda_i \chi^2_{h_i}(\delta_i),$$

where $h_i = 1$, $\delta_i = \nu_i^2$, with ν_i^2 the ith component of the vector $\boldsymbol{\nu}$, $i = 1, \ldots, r$.

Many test statistics converge in distribution toward finite weighted sum of chi-square random variables. A famous example is the Chernoff-Lehmann test statistic for goodness-of-fit to a fixed distribution, which converges to a finite weighted sum of central (non-central) chi-square under the null (alternative) hypothesis (see, e.g., Moore and Spruill (1975), Spruill (1976)). In time series analysis, a popular procedure often encountered in applied work is the Box-Pierce-Ljung portmanteau test statistic for lack of fit in autoregressive-moving-average (ARMA) time series models. The approximate critical values of that test procedure are often taken from a chi-square distribution, but in general the valid asymptotic distribution under the null hypothesis is given by a finite weighted sum of central chi-square random variables, see Ljung (1986), Francq, Roy and Zakoïan (2005) and Duchesne and Francq (2008), among others. Under local alternatives (in the Pitman's sense), it may be seen that the Box-Pierce-Ljung portmanteau test statistic converges in distribution to a finite weighted sum of non-central chi-square random variables. Another example is the portmanteau test statistic of Peña and Rodriguez (2002) for model checking in linear and non-linear time series models, which under certain conditions display more power than the Box-Pierce-Ljung test statistics. See also Lin and Mcleod (2006). Our literature review is far from being exhaustive, and in fact is very selective. However, it suggests the importance of determining (1) for quadratic forms in central and non-central normal variables in level and power studies.

The computation of (1) for quadratic forms in non-central normal variables will typically arise in power analysis (note in passing that level studies are typically performed under the null hypothesis, and thus (1) should also be computed accurately for quadratic forms in central normal variables). Many methods have been proposed for that problem, including methods relying on numerical inversion of the characteristic function (see, e.g., Imhof (1961), Davies (1973, 1980)). These methods are not limited to non-negative quadratic forms and they are found to perform better than Pearson's three-moment chi-square approximation in these situations (Imhof (1961)). Farebrother (1984) and Sheil and O'Muircheartaigh (1977), based on results of Ruben (1962), exploit the fact that (1) can be written as an infinite series of central chi-square distributions. Farebrother (1990) proposed a method which expresses a quadratic form in an alternative form, using the so-called tridiagonal form. Another reference is Kuonen (1999), who uses saddlepoint approximations.

Recently, Liu, Tang and Zhang (2009) proposed a new moment-based approach. Their method relies on a chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables. When the normal variables have zero mean, their method reduces to the Pearson's three-moment chi-square approach. It should be noted that Pearson's three-moment chi-square approximation may be inaccurate to determine probabilities in certain regions of the domain, but it is generally accurate in the upper tail of the distribution (see, e.g., Imhof (1961) or Kuonen (1999), among others). Liu, Tang and Zhang (2009) presented interesting and encouraging numerical examples: in the upper tail of the distribution of the quadratic form in non-central normal variables, their results suggested that their method provided a better approximation of the distribution than Pearson's method. From their numerical results, the probabilities obtained from their approximation were also very close to the exact values (in the upper tail, the absolute errors were no more that 3×10^{-6}).

The principal objective of this note is to provide additional empirical evidence. The Liu-Tang-Zhang approach is compared to exact methods in Section 2, notably Imhof's (1961) method and Farebrother's (1984) algorithm (note that these methods are called exact in the sense that it is possible to bound the approximation error, which can be made arbitrarily small). The comparisons are made at various points of the distribution support in order to appreciate when the moment-based approximation is satisfactory. While the Liu-Tang-Zhang's four-moment chi-square approximation is interesting because of its inherent simplicity, our numerical findings suggest that the method should be used with care, since significant differences may occur between that moment-based approach and exact methods, even in the upper tail of the distribution. Section 3 offers concluding remarks.

2. EMPIRICAL COMPARISONS AND DISCUSSION

This section compares the Liu-Tang-Zhang's four-moment chi-square approximation with two exact methods. In order to approximate accurately the true probabilities, we proceed as in Liu, Tang and Zhang (2009) and we use Farebrother's (1984) algorithm, using a theoretical result of Ruben (1962), as reported in Kotz, Johnson and Boyd (1967, p. 843). See Liu, Tang and Zhang (2009, p. 855) and Farebrother (1984, pp. 333-334). We also include Imhof's (1961) method in the numerical comparisons, which is popular in several studies. We would like to point out that all the computations were performed using the **R** package `CompQuadForm` that we developed recently. Note that in addition to Imhof's (1961) method and Farebrother's (1984) algorithm, our **R** package also includes an algorithm provided by Davies (1980). In order to achieve computational efficiency, the algorithms included in our package have been implemented in the C language and interfaced with the **R** software. The package `CompQuadForm` is freely available from the authors. All the results using Farebrother's (1984) algorithm in our **R** package have been verified with the NAG routine G01JCF (Mark 18) of the FORTRAN 77 language; the results were identical with a precision of 10^{-6}.

We considered eight quadratic forms, which are defined in Table 1. The quadratic forms $\mathcal{Q}_1$ and $\mathcal{Q}_2$ represent modifications of the first and second cases of Liu, Tang and Zhang (2009). For $\mathcal{Q}_1$, the weights 0.4 and 0.1 are replaced by the values 4.0 and 1.0, respectively, and the non-centrality parameters 0.6 and 0.8 are divided by two. For the quadratic form $\mathcal{Q}_2$, the weight 0.3 is multiplied by 10, and the non-centrality parameter 2 is divided by 10. In $\mathcal{Q}_3$, we consider a situation where the weight of the first $\chi_1^2(\delta)$ random variable is large but its non-centrality parameter is small, while the weight of the second $\chi_1^2(\delta)$ random variable is small, but the associated non-centrality parameter is large. The quadratic forms $\mathcal{Q}_4$ is inspired from Davies (1980, Table 3), who considered a similar quadratic form expressed as a finite sum of centered chi-square random variables; we simply introduced non-centrality parameters.

The quadratic forms $\mathcal{Q}_5$-$\mathcal{Q}_8$ are inspired from the asymptotic distribution of the Box-Pierce-Ljung portmanteau test statistic, which is widely used in time series analysis. If a time series follows a first-order autoregressive or moving-average model with parameter α and is correctly fitted, then the Box-Pierce-Ljung portmanteau test statistic follows asymptotically the quadratic form $\chi_{m-1}^2 + \alpha^{2m}\chi_1^2$. For $\mathcal{Q}_5$ and $\mathcal{Q}_6$, we considered $m = 2$, $\alpha = 0.6$, and we introduced non-centrality parameters 1.0 and 7.0 in $\mathcal{Q}_5$, while the non-centrality parameters were 0.1 and 10.0 in $\mathcal{Q}_6$. For the quadratic form $\mathcal{Q}_7$, we examined $m = 3$, $\alpha = 0.7$, with non-centrality parameters 0.2 and 10.0. For $\mathcal{Q}_8$, we considered $m = 3$, $\alpha = 0.8$, with non-centrality parameters 0.0 and 8.0. We do not claim that these choices for the non-centrality parameters have a precise statistical interpretation: they have been chosen in order to appreciate the quality of the Liu-Tang-Zhang approximation compared with Imhof's (1961) method and Farebrother's (1984) algorithm.

Table 1. Definitions of the quadratic forms

$$
\begin{aligned}
\mathcal{Q}_1 &= 0.5\chi_1^2(1.0) + 4.0\chi_2^2(0.3) + 1.0\chi_1^2(0.4),\\
\mathcal{Q}_2 &= 0.7\chi_1^2(6.0) + 3.0\chi_1^2(0.2),\\
\mathcal{Q}_3 &= 10.0\chi_1^2(0.1) + \chi_1^2(10.0),\\
\mathcal{Q}_4 &= 6.0\chi_2^2(0.2) + 3.0\chi_2^2(0.8) + 1.0\chi_2^2(12.0),\\
\mathcal{Q}_5 &= \chi_1^2(1.0) + (0.6)^4\chi_1^2(7.0),\\
\mathcal{Q}_6 &= \chi_1^2(0.1) + (0.6)^4\chi_1^2(10.0),\\
\mathcal{Q}_7 &= \chi_2^2(0.2) + (0.7)^6\chi_1^2(10.0),\\
\mathcal{Q}_8 &= \chi_2^2 + (0.8)^6\chi_1^2(8.0).
\end{aligned}
$$

The results are presented in Table 2. Our numerical examples demonstrate that Farebrother's (1984) and Imhof's (1961) methods differ very little, using both the absolute or relative errors. When the probabilities are compared with the exact values, the relative errors of the Liu-Tang-Zhang four-order moment chi-square approximation are more important than those presented in Liu, Tang and Zhang (2009). More importantly, the numerical examples in Liu, Tang and Zhang (2009) seem not representative of all situations of interest in practical applications, since the probabilities obtained from the Liu-Tang-Zhang method may be relatively far from the exact probabilities.

We now discuss the results presented in Table 2 in more details. We first study $\mathcal{Q}_1$. The difference with the first example of Liu, Tang and Zhang (2009) is related to more important weights in the quadratic forms for the non-central chi-square variables with small non-centrality parameters. It may be noted that the absolute errors are larger than those displayed in Liu, Tang and Zhang (2009). For example, in their numerical example, when the true probability was 3.1% they obtained an absolute error of 3×10^{-5}, while in our example the absolute error is larger by a factor 10 when the true probability is around 5%. We did a similar experiment with $\mathcal{Q}_2$, where the non-central chi-square variable with a small weight was associated with a large non-centrality parameter, and the other non-central chi-square variable offered a larger weight but a smaller non-centrality parameter. In that case, the Liu-Tang-Zhang four-order moment chi-square approximation gave large errors, as large as 7% when the true probability was about 5%. The quadratic form $\mathcal{Q}_3$ is similar to $\mathcal{Q}_2$ but it offers more extreme differences between the weights are the non-centrality parameters; again, the Liu-Tang-Zhang approximation performed poorly. The quadratic form $\mathcal{Q}_4$ also suggests that the moment-based approximation can be relatively unsatisfactory. Under $\mathcal{Q}_5$, the non-centrality parameters were 1.0 and 7.0, and the relative error in the upper tail of the distribution is about 1.2%, when the true probability is approximately 3.6%. The quadratic form $\mathcal{Q}_6$ is similar to $\mathcal{Q}_5$, with the differences that the first non-centrality parameter is divided by 10, and the second non-centrality parameter is increased from 7.0 to 10.0. In that case, the relative error in the upper tail of the distribution was about 8.6% when the true probability was 4.7%, about 2.6% when the true probability was 1.6%, and 6.8% when the true probability was 0.6%. Note that Imhof's (1961) method and Farebrother's (1984) algorithm displayed some differences using the relative error criterion when the true probability was very small. However, the absolute error was 10^{-6} and was considerably smaller than the one of the Liu-Tang-Zhang approximation. The quadratic forms $\mathcal{Q}_7$ and $\mathcal{Q}_8$ also suggested that the moment-based approximation can be unsatisfactory in the upper tail of the distribution, while the exact methods were generally in close agreement.

3. CONCLUSION

Our numerical results suggest that the moment-based chi-square approximations may provide poor approximations to the distributions of quadratic forms in non-central normal variables, when compared to the exact values. While that kind of techniques are nevertheless interesting and appealing giving their inherent simplicity, it should be noted that modern computer resources allow us to implement exact methods such as Imhof's (1961) method very efficiently. In the numerical examples presented in this note, the cost in computer time was too small to be a determinant factor. The objective here was not to explain under which conditions the Liu-Tang-Zhang chi-square approximation is appropriate (however, our numerical examples may suggest when the approximation is not satisfactory), but to point out that it should be used with care, even in the upper tail of the distribution. In conclusion, it is possible that the Liu-Tang-Zhang approximation performs better that Pearson's method under certain conditions, but the probabilities found using that method may be relatively far from the exact probabilities.

REFERENCES

Davies, R. B. (1973), 'Numerical inversion of a characteristic function', *Biometrika* **60**, 415–417.

Davies, R. B. (1980), 'Algorithm AS 155: the distribution of a linear combination of χ^2 random variables', *Journal of the Royal Statistical Society, Series C (Applied Statistics)* **29**, 323-333.

Duchesne, P. and Francq, C. (2008), 'On diagnostic checking time series models with portmanteau test statistics based on generalized inverses and { 2 }-inverses', *COMPSTAT 2008, Proceedings in Computational Statistics*, 143–154.

Farebrother, R. W. (1984), 'Algorithm AS 204: the distribution of a positive linear combination of χ^2 random variables', *Journal of the Royal Statistical Society, Series C (Applied Statistics)* **33**, 332-339.

Farebrother, R. W. (1990), 'Algorithm AS 256: the distribution of a quadratic form in normal variables', *Journal of the Royal Statistical Society, Series C (Applied Statistics)* **39**, 294-309.

Francq, C., Roy, R. and Zakoïan, J.-M. (2005), 'Diagnostic checking in ARMA models with uncorrelated errors', *Journal of the American Statistical Association* **100**, 532–544.

Imhof, J. P. (1961), 'Computing the distribution of quadratic forms in normal variables', *Biometrika* **48**, 419–426.

Kotz, S., Johnson, N. L. and Boyd, D. W. (1967), 'Series representations of distributions of quadratic forms in normal variables II. Non-central case', *Annals of Mathematical Statistics* **38**, 838-848.

Kuonen, D. (1999), 'Saddlepoint approximations for distributions of quadratic forms in normal variables', *Biometrika* **86**, 929–935.

Lin, J.-W. and McLeod, A. I. (2006), 'Improved Peña-Rodriguez portmanteau test', *Computational Statistics & Data Analysis* **51**, 1731–1738.

Liu, H., Tang, Y. and Zhang, H. H. (2009), 'A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables', *Computational Statistics & Data Analysis* **53**, 853–856.

Ljung, G. M. (1986), 'Diagnostic testing of univariate time series models', *Biometrika* **73**, 725–730.

Peña, D. and Rodriguez, J. (2002), 'A powerful portmanteau test of lack of fit for time series', *Journal of the American Statistical Association* **97**, 601–610.

Moore, D. S. and Spruill, M. C. (1975), 'Unified large-sample theory of general chi-squared statistics for tests of fit', *The Annals of Statistics* **3**, 599-616.

Ruben, H. (1962), 'Probability content of regions under spherical normal distributions. IV The distribution of homogeneous and non-homogeneous quadratic functions of normal variables', *Annals of Mathematical Statistics* **33**, 542-570.

Sheil, J. and O'Muircheartaigh, I. (1977), 'Algorithm AS 106: the distribution of non-negative quadratic forms in normal variables', *Journal of the Royal Statistical Society, Series C (Applied Statistics)* **26**, 92-98.

Spruill, M. C. (1976), 'A comparison of chi-square goodness-of-fit tests based on approximate Bahadur slope', *The Annals of Statistics* **4**, 409-412.

TABLE 2. Probability that the quadratic form exceeds q. The column F gives the exact values with accuracy 10^{-6} using Farebrother's (1984) algorithm. The column I provides the values using Imhof's (1961) method. The column AE_I ($RE_I(\%)$) gives the absolute error (relative error in percentage) of Imhof's (1961) method compared to Farebrother's (1984) algorithm. The column LTZ presents the values using the Liu, Tang and Zhang's (2009) method. The column AE_{LTZ} ($RE_{LTZ}(\%)$) gives the absolute error (relative error in percentage) of the Liu-Tang-Zhang method compared to Farebrother's (1984) algorithm.

Quadratic form	q	F	I	AE_I	$RE_I(\%)$	LTZ	AE_{LTZ}	$RE_{LTZ}(\%)$
$\mathcal{Q}_1$	6	0.679440	0.679440	0.000000	0.000003	0.669712	0.009728	1.431733
	8	0.556520	0.556520	0.000000	0.000004	0.549650	0.006870	1.234521
	20	0.152962	0.152962	0.000000	0.000008	0.154503	0.001541	1.007563
	30	0.050874	0.050874	0.000000	0.000059	0.051294	0.000419	0.824287
$\mathcal{Q}_2$	6	0.591269	0.591269	0.000000	0.000008	0.567547	0.023722	4.012093
	15	0.127068	0.127068	0.000000	0.000012	0.132639	0.005570	4.383763
	20	0.052153	0.052153	0.000000	0.000113	0.056008	0.003855	7.391635
	25	0.022099	0.022100	0.000001	0.004046	0.023224	0.001126	5.093545
$\mathcal{Q}_3$	40	0.114930	0.114930	0.000000	0.000022	0.122072	0.007142	6.214041
	50	0.064546	0.064546	0.000000	0.000140	0.069020	0.004474	6.931217
	60	0.037203	0.037203	0.000000	0.000392	0.039296	0.002093	5.625822
	70	0.021772	0.021772	0.000000	0.001076	0.022482	0.000710	3.260405
$\mathcal{Q}_4$	30	0.570073	0.570073	0.000000	0.000000	0.561196	0.008877	1.557103
	40	0.330482	0.330482	0.000000	0.000010	0.329954	0.000528	0.159705
	60	0.086659	0.086659	0.000000	0.000031	0.089430	0.002771	3.197466
	70	0.041871	0.041871	0.000000	0.000314	0.043121	0.001250	2.985924
$\mathcal{Q}_5$	4	0.249843	0.249843	0.000000	0.000000	0.254194	0.004351	1.741461
	5	0.169313	0.169313	0.000000	0.000003	0.171716	0.002402	1.418880
	6	0.115043	0.115043	0.000000	0.000051	0.115845	0.000802	0.696923
	9	0.035784	0.035785	0.000001	0.000289	0.035350	0.000434	1.214135
$\mathcal{Q}_6$	4	0.151668	0.151668	0.000000	0.000014	0.159370	0.007702	5.078090
	6	0.047146	0.047146	0.000000	0.000073	0.051219	0.004073	8.638854
	8	0.016046	0.016046	0.000000	0.000471	0.016459	0.000413	2.576934
	10	0.005676	0.005675	0.000001	0.010978	0.005289	0.000387	6.818694
$\mathcal{Q}_7$	3	0.484953	0.484953	0.000000	0.000018	0.474929	0.010024	2.067013
	6	0.125851	0.125851	0.000000	0.000014	0.129265	0.003414	2.712853
	7	0.079686	0.079686	0.000000	0.000116	0.081951	0.002264	2.841448
	8	0.050413	0.050413	0.000000	0.000084	0.051630	0.001217	2.414282
$\mathcal{Q}_8$	6	0.215605	0.215605	0.000000	0.000004	0.217333	0.001728	0.801447
	8	0.085359	0.085359	0.000000	0.000028	0.088209	0.002850	3.338825
	10	0.032178	0.032178	0.000000	0.000875	0.033241	0.001063	3.303265
	14	0.004394	0.004395	0.000001	0.010090	0.004111	0.000284	6.453687