

## A COMPARISON OF APPROXIMATIONS TO PERCENTILES OF THE NONCENTRAL $\chi^2$ -DISTRIBUTION

HARDEO SAHAI\*

MARIO MIGUEL OJEDA†

*Recibido: 26 Jun 2002*

---

### Abstract

Various approximations to percentiles of the noncentral  $\chi^2$ -distribution are examined for their accuracy over a wide range of values of the parameters of the distribution.

**Keywords:** Distribución de probabilidad, percentiles, aproximación numérica.

### Resumen

Se examinan varias aproximaciones de los percentiles de la distribución  $\chi^2$  no centrada, debido a su precisión en un amplio rango de valores de los parámetros de la distribución.

**Palabras clave:** Probability distribution, percentiles, numerical approximation.

**Mathematics Subject Classification:** 62E17.

## 1 Introduction and summary

It is widely recognized that the noncentral  $\chi^2$ -distribution is of considerable theoretical and practical importance in many mathematical and statistical applications. For instance, noncentral  $\chi^2$  has applications in deriving expected values of quadratic forms in analysis of variance (Graybill, 1976, pp. 139-140), in approximating the nonnull distribution and power of the  $\chi^2$ -test of goodness of fit (Tiku, 1985), as well as certain other nonparametric tests (Andrews, 1954; Lehmann, 1975, p. 247). In addition, noncentral  $\chi^2$  also appears in the derivation of asymptotic ( $n$  tends to infinity) nonnull distribution of the Hotelling's

---

\*Department of Biostatistics and Epidemiology, University of Puerto Rico, San Juan, Puerto Rico 00936, USA. E-Mail:

†Facultad de Estadística e Informática, Universidad Veracruzana, Av. Xalapa s/n esquina Ávila Camacho, Xalapa, Veracruz, México. E-Mail: [mojeda@uv.mx](mailto:mojeda@uv.mx)

$T^2$  (Anderson, 1984, pp.163-164; Tiku, 1985), and that of likelihood ratio-statistics for tests of multivariate linear hypotheses (Wilks, 1962, p. 419; Sugiura, 1968; Graybill, 1976, pp. 189-190).

In many applications involving the noncentral  $\chi^2$  one has to compute its percentiles involving the evaluation of the inverse probability functions (see, *e.g.*, Bagui, 1996). However, the evaluation of such inverse functions is extremely tedious involving slow and expensive techniques of numerical iteration such as the Newton-Raphson procedure (see, *e.g.*, Ralston and Wilf, 1967; Carnahan *et al.*, 1969). There are a number of approximations for computing the percentage points of the noncentral  $\chi^2$ -distribution, at arbitrary probability levels, available in the literature. The applicability of several of these approximations is further enhanced by ease of their computational simplicity. The purpose of this paper is to compare many of these approximations to determine their accuracy. Some of these approximations were previously investigated by Patnaik (1949), Pearson (1959), Sankaran (1959, 1963), and Cox and Reid (1987). A brief description of each procedure is given and appropriate tables comparing their accuracy, calculated for each procedure, are presented. A more comprehensive set of tables is given in Sahai and Ojeda (1998).

## 2 Approximations

The noncentral  $\chi^2$ -distribution was obtained by Fisher (1928, p. 663), as a limiting case of the distribution of the multiple correlation coefficient, who also gave upper 5% points of the distribution for certain selected values of the degrees of freedom and the noncentrality parameter. There are many approximations to the noncentral  $\chi^2$ -distribution discussed in the literature which can be used to compute the percentiles of the distribution. Some of the important ones are considered here.

In this paper,  $\chi_\nu'^2(\lambda)$  will be used to denote a noncentral  $\chi^2$ -variate with  $\nu$  degrees of freedom and the noncentrality parameter  $\lambda$ . In addition,  $\chi_{\nu,\alpha}'^2(\lambda)$  will denote its  $100\alpha$ -th percentile defined by

$$\Pr \left[ \chi_\nu'^2(\lambda) \leq \chi_{\nu,\alpha}'^2(\lambda) \right] = \alpha. \quad (1)$$

Patnaik (1949) suggested an approximation of  $\chi_\nu'^2(\lambda)$  by  $c\chi_f^2$  where  $c$  and  $f$ , obtained by equating the first two moments of the two variables, are

$$c = (\nu + 2\lambda) / (\nu + \lambda) \text{ and } f = (\nu + \lambda)^2 / (\nu + 2\lambda). \quad (2)$$

Patnaik (1949) also proposed a normal approximation of  $\chi_\nu'^2(\lambda)$  which consists in first approximating  $\chi_\nu'^2(\lambda)$  by  $c\chi_f^2$  and then approximating  $\sqrt{2c\chi_f^2}$  by a normal variate with mean  $\sqrt{2f-1}$  and variance 1. The resulting approximation is:  $\left\{ [2(\nu + \lambda) / (\nu + 2\lambda)] \chi_\nu'^2 \right\}^{1/2}$  has a normal distribution with mean  $\left\{ [2(\nu + \lambda)^2 / (\nu + 2\lambda)] - 1 \right\}^{1/2}$  and variance 1. We will refer the two Patnaik's approximations as the Patnaik's 1st and 2nd approximations, respectively.

Pearson (1959) suggested an improvement to the approximation (2) which consists in approximating  $\chi_\nu'^2(\lambda)$  by  $b + c\chi_f^2$  where  $b$ ,  $c$  and  $f$ , obtained by equating the first three

moments of the two variables, are

$$b = -\lambda^2/(\nu + 3\lambda), c = (\nu + 3\lambda)/(\nu + 2\lambda) \text{ and } f = (\nu + 2\lambda)^3/(\nu + 3\lambda)^2 \quad (3)$$

Abdel-Aty (1954) also considered a normal approximation which consists in first approximating  $\chi_\nu'^2(\lambda)$  by  $c\chi_f^2$  and then applying the Wilson-Hilferty (1931) approximation to the central  $\chi_f^2$ . The resulting approximation is:  $\left\{(\nu + \lambda)^{-1} \chi_\nu'^2(\lambda)\right\}^{1/3}$  has a normal distribution with mean  $1 - 2(\nu + 2\lambda)/9(\nu + \lambda)^2$  and variance  $2(\nu + 2\lambda)/9(\nu + \lambda)^2$ .

Sankaran (1959, 1963) discussed among others the following normal approximations of  $\chi_\nu'^2(\lambda)$ :

(i)  $\left[\chi_\nu'^2(\lambda) - (\nu - 1)/2\right]^{1/2}$  has a normal distribution with mean  $[\lambda + (\nu - 1)/2]^{1/2}$  and variance 1.

(ii)  $\left\{\frac{\chi_\nu'^2(\lambda) - (\nu - 1)/3}{\nu + \lambda}\right\}^{1/2}$  has a normal distribution with mean

$$1 - \frac{\nu + 2}{6r} - \frac{\nu^2 - 2\nu + 10}{72r^2} - \frac{\nu^3 - 12\nu^2 - 6\nu + 44}{432r^3} - 5\frac{\nu^4 - 28\nu^3 + 24\nu^2 + 1112\nu - 1028}{10368r^4}$$

and variance

$$\frac{1 - \frac{\nu - 1}{6r} - \frac{\nu^2 + \nu - 2}{18r^2} - \frac{4\nu^3 - 9\nu^2 - 228\nu + 233}{216r^3}}{r}$$

where  $r = \nu + \lambda$ .

(iii)  $\left[\chi_\nu'^2\lambda/(\nu + \lambda)\right]^h$  has a normal distribution with mean

$$1 + \frac{h(h - 1)(\nu + 2\lambda)}{(\nu + \lambda)^2} - \frac{h(h - 1)(2 - h)(1 - 3h)(\nu + 2\lambda)^2}{2(\nu + \lambda)^4}$$

and variance

$$\left[\frac{2h^2(\nu + 2\lambda)}{(\nu + \lambda)^2}\right] \left[1 - (1 - h)(1 - 3h)\frac{(\nu + 2\lambda)}{(\nu + \lambda)^2}\right],$$

where

$$h = 1 - \frac{2(\nu + \lambda)(\nu + 3\lambda)}{3(\nu + 2\lambda)^2}.$$

We will call these approximations as Sankaran's 1st, 2nd and 3rd approximations, respectively.

Johnson (1959) developed a simple normal approximation of  $\chi_\nu'^2(\lambda)$  via

$$\Pr\left[\chi_\nu'^2(\lambda) \leq t\right] \approx \Pr\left\{Z \leq (t - \nu - \lambda + 1)/[2(\nu + 2\lambda)]^{1/2}\right\} \quad (4)$$

by applying a normal approximation to the right hand side of the equation

$$\Pr\left[\chi_\nu'^2(\lambda) \leq t\right] = \Pr[X_1 - X_2 \geq \nu/2],$$

where  $X_1$  and  $X_2$  are independent Poisson variables with mean  $t/2$  and  $\lambda/2$ , respectively.

Johnson and Kotz (1970, p. 141) suggested a simple normal approximation of  $\chi_{\nu}^{\prime 2}(\lambda)$  by its direct standardization via

$$\Pr \left[ \chi_{\nu}^{\prime}(\lambda) \leq t \right] \approx \Pr \left\{ Z \leq (t - \nu - \lambda) / [2(\nu + 2\lambda)]^{1/2} \right\}. \quad (5)$$

In both approximations (4) and (5) the error as  $(\lambda \rightarrow \infty)$  is  $\mathcal{O}(\lambda^{-1/2})$ , uniformly in  $t$ . Approximations (4) and (5) although extremely simple are not very accurate and have been included here for the sake of completeness.

Bol'shev and Kuznetsov (1963), using a method in which the distribution of  $\chi_{\nu}^{\prime 2}(\lambda)$  is related to the distribution of central  $\chi_{\nu}^2$  with the same number of degrees of freedom, gave an approximation as (see also Johnson *et al.* 1995, pp. 465–466):

$$\Pr \left[ \chi_{\nu}^{\prime 2}(\lambda) \leq t \right] \approx \Pr \left\{ \chi_{\nu}^2 \leq t \left[ 1 - \frac{\lambda}{\nu} + \frac{1}{2} (\lambda^2/\nu^2) \left( 1 + \frac{t}{\nu+2} \right) \mathcal{O}(\lambda^3) \right] \right\}, \quad (6)$$

where  $\mathcal{O}(\lambda^3)$  is uniform in any finite interval of  $t$ ,  $\lambda \Rightarrow 0$ , leading to the approximation:

$$\chi_{\nu,\alpha}^{\prime 2}(\lambda) \approx \chi_{\nu,\alpha}^2 + (\lambda/\nu) \chi_{\nu,\alpha}^2 + \frac{1}{2} (\lambda/\nu)^2 \left[ 1 - \frac{1}{\nu+2} \chi_{\nu,\alpha}^2 \right] \chi_{\nu,\alpha}^2. \quad (7)$$

Approximation (7) is not very accurate, but has been included here for the sake of completeness. For small values of  $\nu$ , the quality of the approximation deteriorates rapidly as  $\lambda$  increases. For very large values of  $\nu$ , the approximation improves somewhat but is still not to be recommended.

Cox and Reid (1987) considered the approximation

$$\Pr \left[ \chi_{\nu}^{\prime 2}(\lambda) \leq t \right] \approx \Pr \left\{ \chi_{\nu}^2 \leq t / (1 + \lambda/\nu) \right\}. \quad (8)$$

Approximation (8) is valid for  $\lambda \rightarrow 0$  as  $\nu \rightarrow \infty$ . Note that ignoring the third term in the brackets in (6) gives an approximation which is asymptotically equivalent to (8).

In addition, Cox and Reid (1987) proposed approximating  $\chi_{\nu}^{\prime 2}(\lambda)$  by the linear combination  $(1 - \lambda/2) \chi_{\nu}^2 + (\lambda/2) \chi_{\nu+2}^2$ , prompted by a result given in Barndorff-Nielsen and Cox (1985, Equation 1.6). We will call these approximations as Cox-Reid's 1st and 2nd approximations, respectively.

Temme (1993) gave a simple and useful approximation, for large values of  $t$  and  $\lambda$ , as

$$\Pr \left[ \chi_{\nu}^{\prime 2}(\lambda) \leq t \right] \approx \begin{cases} (t/\lambda)^{(\nu-1)/4} \left[ 1 - \Phi \left( \sqrt{2\lambda} - \sqrt{2t} \right) \right], & t \leq \lambda \\ 1 - (t/\lambda)^{(\nu-1)/4} \left[ 1 - \Phi \left( \sqrt{2t} - \sqrt{2\lambda} \right) \right], & t > \lambda. \end{cases} \quad (9)$$

Although approximation (9) is quite simple for evaluating a probability expression, it requires the use of an iterative procedure to compute the corresponding percentile value and is not included in our comparative study.

Finally, we note a simple empirical approximation, reported by Tukey (1957), to the 95th percentiles of  $\chi^2_{\nu}(\lambda)$  given by

$$\chi^2_{\nu,0.95} \approx \left[ 1.6449 + \sqrt{\lambda} + 0.51 \frac{\nu - 1}{\sqrt{\lambda} + 1} - 0.024 \frac{(\nu - 5)(\nu - 1)}{\sqrt{\lambda}(\sqrt{\lambda} + 1)} \right]^2. \quad (10)$$

Although approximation (10) is limited to only 95th percentiles, it has been included here for detailed comparison in view of its usefulness, simplicity and accuracy.

### 3 Results

The percentiles of  $\chi^2_{\nu}(\lambda)$  calculated for various approximations as well as the exact values, for selected values of  $\alpha$ ,  $\nu$  and  $\lambda$ , are shown in Table 2. For higher values of percentiles, the Patnaik's central  $\chi^2$  approximation is the most accurate, when  $\nu$  and  $\lambda$  are small. In this situation, the 1st and 3rd approximations of Sankaran are also quite accurate. For lower percentiles, the 3rd approximation of Sankaran performs best; and for small values of  $\nu$  and  $\lambda$ , the Patnaik's central  $\chi^2$  approximation is also quite accurate. Johnson and Johnson-Kotz type approximations differ by a percentile of one unit and give satisfactory results for higher percentiles and large values  $\lambda$ . Bol'shev-Kuznetsov and the two Cox-Reid approximations have extremely poor performance for simultaneously small values of  $\nu$  and large values of  $\lambda$ . However, their accuracy improves dramatically for large values of  $\nu$ , but are still not to be recommended. Tukey's empirical approximation for the 95th percentile, although quite accurate for small values of  $\nu$  progressively degenerates as  $\nu$  increases.

### Acknowledgements

The exact percentiles of the noncentral  $\chi^2$ -distribution were calculated using AMOSLIB, a special function library prepared at Sandia Laboratories, Albuquerque, New Mexico. We are indebted to Dr. Donald E. Amos for his courtesy in providing the pertinent information about the routines including the necessary software to perform the requisite computations. We also thank Professor Constance van Eeden for reading a preliminary draft of the manuscript and making some helpful comments and suggestions. This work would not have been possible without the generous dedication of time and efforts by Rafael Guajardo Panes and Lorena López whose contributions are gratefully acknowledged.

### Referencias

- [1] Abdel-Aty, S.H. (1964) "Approximate formulae for the percentage points and the probability integral of the noncentral  $\chi^2$  distribution", *Biometrika* **41**: 538–540.
- [2] Anderson, T. W. (1984) *An Introduction to Multivariate Statistical Analysis*, 2nd ed. John Wiley and Sons, New York.
- [3] Andrews, F. C. (1964) "Asymptotic behavior of some rank tests for analysis of variance", *Annals of Mathematical Statistics* **25**: 727–736.

- [4] Bagui, S. C. (1996) *CRC Handbook of Percentiles of Noncentral Distributions*. CRC Press, Boca Raton, Florida.
- [5] Bol'shev, L. N.; Kuznetsov, P. I. (1963) "On the calculation of the integral  $p(x,y)=\dots$ ", *Computational and Mathematical Physics* **3**: 557–571. (English translation from Russian).
- [6] Carnahan, B.; Luther, H. A.; Wilkes, J. O. (1969) *Applied Numerical Methods*. John Wiley and Sons, New York.
- [7] Cox, D. R.; Reid, N. (1987) "Approximations to noncentral distributions", *Canadian Journal of Statistics* **15**: 105–114.
- [8] Fisher, R. A. (1928) "The general sampling distribution of the multiple correlation coefficient", *Proceedings of the Royal Society of London, Series A* **121**: 654–673.
- [9] Graybill, F. A. (1976) *Theory and Application of the Linear Model*. Duxbury Press, North Scituate, Massachusetts.
- [10] Johnson, N. L. (1959) "On the extension of the connection between Poisson and  $\chi^2$  distributions", *Biometrika* **46**: 352–363.
- [11] Johnson, N. L.; Kotz, S. (1970) *Continuous Univariate Distributions*, Vol.2. John Wiley and Sons, New York.
- [12] Johnson, N. L.; Kotz, S.; Balkrishnan, N. (1995) *Continuous Univariate Distributions*, Vol. 2, 2nd ed. John Wiley and Sons, New York.
- [13] Lehmann, E. L. (1975) *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day, Inc., Oakland, California.
- [14] Patnaik, P. B. (1949) "The noncentral  $\chi^2$  and  $F$ -distributions and their applications", *Biometrika* **36**: 202–232.
- [15] Pearson, E. S. (1959) "Note on an approximation to the distribution of noncentral  $\chi^2$ ", *Biometrika* **46**: 364.
- [16] Ralston, A.; Wilf, H. (1967) *Mathematical Methods for Digital Computers*. John Wiley and Sons, New York.
- [17] Sahai, H.; Ojeda, M. M. (1998) *Approximations to percentiles of the noncentral  $t$ ,  $\chi^2$  and  $F$  Distributions*. University of Veracruz Press, Xalapa, México.
- [18] Sankaran, M. (1959) "On the Noncentral Chi-Square Distribution", *Biometrika* **46**: 235–237.
- [19] Sankaran, M. (1963) "Approximations to the noncentral chi-square distribution", *Biometrika* **50** 199–204.
- [20] Temme, N. M. (1993) "Asymptotic and numerical aspects of the noncentral chi-square distribution", *Computers in Mathematics and Applications* **25**: 55–63.
- [21] Tiku, M. L. (1985) "Noncentral chi-square distribution", In: S. Kotz & N. L. Johnson (Eds.) *Encyclopedia of Statistical Science*, Vol. 6. John Wiley and Sons, New York: 276–280.
- [22] Tukey, J. W. (1957) "Approximations to the upper 5% points of Fisher's  $B$  distribution and noncentral  $\chi^2$ ", *Biometrika* **44**: 528–530.
- [23] Wilks, S. S. (1962) *Mathematical Statistics*. John Wiley and Sons, New York.
- [24] Wilson, E. B.; Hilferty, M. M. (1931) "The distribution of chi-square", *Proceedings of the National Academy of Sciences (U. S. A.)* **19**: 684–688.