AN APPROXIMATION TO THE NON-CENTRAL CHI-SQUARE DISTRIBUTION

By J. ROY

Indian Statistical Institute

and

JAN MOHAMAD

Jammu and Kashmir University

SUMMARY. Two additional corrective terms to Patnaik's (1949) approximation to the non-central chi-square distribution are derived from a Laguerre series expansion of the density function.

1. Laguerre series expansion

A stochastic variable Y is said to have the non-central chi-square distribution with n degrees of freedom and non-centrality parameter λ, if its probability density function is of the form :

$$\sum_{j=0}^{\infty} f_{n+2j}(y) \cdot p_j(\tfrac{1}{2}\lambda), \qquad \ldots \quad (1.1)$$

where $f_n(y)$ is the probability density function of the central chi-square distribution with n degrees of freedom :

$$f_n(y) = \begin{cases} c_n e^{-\frac{1}{2}y} y^{\frac{1}{2}n-1} & \text{for } 0 \leqslant y < \infty \\ 0, & \text{otherwise} \end{cases} \qquad \ldots \quad (1.2)$$

where $\quad 1/c_n = 2^{\frac{1}{2}n}\Gamma(\tfrac{1}{2}n)$

and $\quad p_j(\tfrac{1}{2}\lambda) = e^{-\frac{1}{2}\lambda}(\tfrac{1}{2}\lambda)^j/j! \qquad \ldots \quad (1.3)$

denotes the j-th term of the Poisson probability distribution with mean $\frac{1}{2}\lambda$, $j = 0, 1, 2, \ldots$.

We shall consider the transformed variable $X = Y/2\rho$ where ρ is a constant whose value will be specified later, and derive a Laguerre series expansion for $\phi(x)$, the probability density function of X. Following Roy and Tiku (1962), we shall write

$$p_m(x) = \begin{cases} \dfrac{1}{\Gamma(m)} e^{-x} x^{m-1}, & 0 \leqslant x < \infty \\ 0, & \text{otherwise} \end{cases} \qquad \ldots \quad (1.4)$$

for the Gamma density function with mean $m > 0$.

The Laguerre polynomials

$$L_r^{(m)}(x) = \sum_{t=0}^{\infty} C_{r,t}^{(m)}(-x)^t/t! \qquad \ldots \quad (1.5)$$

$$\text{where} \quad C_{r,t}^{(m)} = \begin{cases} (m+t)(m+t+1)\ldots(m+r-1)/(r-t)! & \text{for } t = 0, 1, \ldots, r-1 \\ 1 & \text{for } t = r \end{cases} \qquad \ldots \quad (1.6)$$

have the orthogonality property

$$\int_0^\infty L_r^{(m)}(x)L_s^{(m)}(x)p_m(x)dx = \begin{cases} 0 & \text{if } r \neq s \\ C_{r,0}^{(m)} & \text{if } r = s. \end{cases} \quad \ldots \quad (1.7)$$

We then have the formal expansion

$$\phi(x) = p_m(x) \sum_{r=0}^{\infty} a_r^{(m)} L_r^{(m)}(x) \quad \ldots \quad (1.8)$$

where

$$a_r^{(m)} = \int_0^\infty L_r^{(m)}(x)\phi(x)dx / C_{r,0}^{(m)} \quad \ldots \quad (1.9)$$

where m will be specified later.

The coefficients $a_r^{(m)}$ can be obtained in terms of the moments of X from formula (1.9). The cumulants of Y have been calculated by Patnaik (1949) from which we get the first four cumulants of X as

$$K_1 = \tfrac{1}{2}(n+\lambda)\rho^{-1}, \qquad K_2 = \tfrac{1}{2}(n+2\lambda)\rho^{-2}$$
$$K_3 = (n+3\lambda)\rho^{-3}, \qquad K_4 = 3(n+4\lambda)\rho^{-4}. \quad \ldots \quad (1.10)$$

This now enables us to compute the first five coefficients $a_r^{(m)}$ for $r = 0, 1, 2, 3$ and 4. Of course,

$$a_0^{(m)} = 1. \quad \ldots \quad (1.11)$$

We have two disposable parameters ρ and m which we can now fix by setting

$$a_1^{(m)} = a_2^{(m)} = 0. \quad (1.12)$$

This gives

$$\rho = \frac{n+2\lambda}{n+\lambda}, \quad m = \frac{(n+\lambda)^2}{2(n+2\lambda)}. \quad \ldots \quad (1.13)$$

With values of ρ and m so determined, we get after somewhat lengthy but straightforward calculations :

$$a_3^{(m)} = \frac{2\lambda^2}{(m+1)(m+2)(n+2\lambda)^2} \quad \ldots \quad (1.14)$$

$$a_4^{(m)} = \frac{6\lambda^2}{(m+1)(m+2)(m+3)(n+2\lambda)^3} \quad \ldots \quad (1.15)$$

Using only the first five terms of (1.8) we get the approximation :

$$\phi(x) \sim [1+a_3^{(m)}L_3^{(m)}(x)+a_4^{(m)}L_4(x)]p_m(x) \quad \ldots \quad (1.16)$$

which can be put in the alternative form :

$$\begin{aligned} \phi(x) \sim\ & p_m(x) \\ & +b_3^{(m)}[p_m(x)-3p_{m+1}(x)+3p_{m+2}(x)-p_{m+3}(x)] \\ & +b_4^{(m)}[p_m(x)-4p_{m+1}(x)+6p_{m+2}(x)-4p_{m+3}(x)+p_{m+4}(x)] \end{aligned} \quad \ldots \quad (1.17)$$

where

$$b_3^{(m)} = \frac{\lambda^2 m}{3(n+2\lambda)^2} \quad \ldots \quad (1.18)$$

$$b_4^{(m)} = \frac{\lambda^2 m(n+4\lambda)}{4(n+2\lambda)^3} \quad \ldots (1.19)$$

2. Approximations to the distribution function

Let $F_{n,\lambda}(y)$ denote the distribution function of Y, that is, $F_{n,\lambda}(y) = \text{prob}(Y \leqslant y)$.

Then
$$F_{n,\lambda}(y) = \int_0^x \phi(t)dt \qquad \dots \quad (2.1)$$

where $\phi(t)$ is defined by (1.8) and
$$x = y/2\rho. \qquad \dots \quad (2.2)$$

Using approximation (1.17) for $\phi(x)$,

we get
$$\begin{aligned} F_{n,\lambda}(y) \sim\ & P_m(x) \\ & + b_3^{(m)}[P_m(x) - 3P_{m+1}(x) + 3P_{m+2}(x) - P_{m+3}(x)] \\ & + b_4^{(m)}[P_m(x) - 4P_{m+1}(x) + 6P_{m+2}(x) - 4P_{m+3}(x) + P_{m+4}(x)] \end{aligned} \qquad \dots \quad (2.3)$$

where
$$P_m(x) = \int_0^x p_m(t)dt. \qquad \dots \quad (2.4)$$

This can be expressed in the alternative form

$$\begin{aligned} F_{n,\lambda}(y) \sim\ & P_m(x) \\ & + p_m(x)\Big[\ b_3^{(m)} \left\{ \frac{x}{m} - \frac{2x^2}{m(m+1)} + \frac{x^3}{m(m+1)(m+2)} \right\} \\ & + b_4^{(m)} \left\{ \frac{x}{m} - \frac{3x^2}{m(m+1)} + \frac{3x^3}{m(m+1)(m+2)} \right. \\ & \left. - \frac{x^4}{m(m+1)(m+2)(m+3)} \right\} \Big] \end{aligned} \qquad \dots \quad (2.5)$$

We may rewrite equation (2.3) in terms of the distribution function $F_n(x)$ of the central chi-square statistic with n degrees of freedom.

$$F_n(x) = \int_0^x f_n(t)dt \qquad \dots \quad (2.6)$$

Let
$$\nu = 2m = \frac{(n+\lambda)^2}{n+2\lambda}, \qquad \dots \quad (2.7)$$

$$u = 2x = y/\rho. \qquad \dots \quad (2.8)$$

Then
$$\begin{aligned} F_{n,\lambda}(y) \sim\ & F_\nu(u) \\ & + b_3^{(m)}[F_\nu(u) - 3F_{\nu+2}(u) + 3F_{\nu+4}(u) - F_{\nu+6}(u)] \\ & + b_4^{(m)}[F_\nu(u) - 4F_{\nu+2}(u) + 6F_{\nu+4}(u) - 4F_{\nu+6}(u) + F_{\nu+8}(u)] \end{aligned} \qquad \dots \quad (2.9)$$

Here the first term is the approximation given by Patnaik (1949), the second and the third terms can be regarded as additional corrective terms.

3. Illustrative example

We shall illustrate the computation technique by evaluating $F_{n,\lambda}(y)$ for $y = 30$, $n = 15$, $\lambda = 20$. We get

$$\nu = \frac{(n+\lambda)^2}{n+2\lambda} = 22.27273, \qquad \rho = \frac{n+2\lambda}{n+\lambda} = 1.57143, \quad m = \nu/2$$

$$b_3 = \frac{\lambda^2 m}{3(n+2\lambda)^2} = 0.490850, \quad b_4 = \frac{\lambda^2 m(n+4\lambda)}{4(n+2\lambda)^3} = 0.635880,$$

$u = y/\rho = 19.00089$. By interpolation in Table 7 of *Biometrika Tables*, we get the following values of $F_n = F_n(u)$: $F_\nu = 0.344521$, $F_{\nu+2} = 0.239854$, $F_{\nu+4} = 0.157539$, $F_{\nu+6} = 0.097726$ and $F_{\nu+8} = 0.057612$. Thus

$$\Delta_3 = F_\nu - 3F_{\nu+2} + 3F_{\nu+4} - F_{\nu+6} = -0.000150$$

and

$$\Delta_4 = F_\nu - 4F_{\nu+2} + 6F_{\nu+4} - F_{\nu+6} + F_{\nu+8} = -0.002953.$$

We thus have as approximations to $F_{n,\lambda}(y)$:

first approximation (Patnaik's) : $F = 0.34452$

second aproximation : $F + b_3\Delta_3 = 0.34445$

third approximation : $F + b_3\Delta_3 + b_4\Delta_4 = 0.34257$

exact value : $F_{n,\lambda}(y) = 0.34265$.

We present below a few more values of $F_{n,\lambda}(y)$ and the three approximations to it for the purpose of comparison. The values were obtained by interpolation in Table 7 of *Biometrika Tables*.

TABLE. COMPARISON OF VARIOUS APPROXIMATIONS TO NON-CENTRAL CHI-SQUARE DISTRIBUTION

n	λ	y	approximations			exact value*
			first	second	third	
4	4	1.765	0.0387	0.0410	0.0434	0.0500
4	4	17.309	0.9491	0.9476	0.9489	0.9500
7	16	10.257	0.0429	0.0466	0.0493	0.0500
7	16	38.970	0.9382	0.9396	0.9472	0.9500
16	8	30.000	0.7895	0.7875	0.7875	0.7880
16	8	40.000	0.9620	0.9628	0.9634	0.9632
24	24	36.000	0.1560	0.1575	0.1574	0.1567
24	24	72.000	0.9650	0.9604	0.9609	0.9607

* Borrowed from Patnaik (1949).

ACKNOWLEDGMENT

It is a pleasure to acknowledge with thanks the computational assistance received from Sri T. J. Kurian and Sri S. Saha, apprentices in the Data Processing Unit of the Indian Statistical Institute.

REFERENCES

PEARSON, E. S. and HARTLEY, H. O. (Editor) : *Biometrika Tables for Statisticians*, Cambridge University Press.

PATNAIK, P. B. (1949) : The non-central χ^2 and F-distributions and their applications. *Biometrika*, **36**, 202-232.

ROY, J. and TIKU, M. L. (1962) : A Laguerre series appxoximation to the sampling distribution of variance. *Sankhyā*, Series A, **24**, 181-184.

Paper received : November, 1963.