The Annals of Statistics
1982, Vol. 10, No. 3, 1012–1016

ESTIMATION OF THE NON-CENTRALITY PARAMETER OF A CHI SQUARED DISTRIBUTION[1]

By K. M. Lal Saxena and Khursheed Alam

University of Nebraska; Lincoln and Clemson University

Let X be distributed according to a non-central Chi squared distribution with p degrees of freedom. The non-central Chi squared distribution arises in various statistical analyses and the estimation of the non-centrality parameter is of importance in some problems. This paper deals with the admissibility of certain estimates of the non-centrality parameter. It is shown that $(X-p)^+$, the positive part of $X-p$ dominates the maximum likelihood estimator with squared error as the loss function.

1. Introduction and Summary. The non-central Chi squared distribution arises in various statistical analyses, such as the analysis of variance for tests of homogeneity and Pearson's Chi squared test for goodness of fit. For an example in electrical engineering, Spruill (1979) has shown that the measurement of electrical power in a circuit is related to the estimation of the non-centrality parameter of a Chi squared distribution. A discussion of various applications of the distribution is given in Johnson and Kotz (1970, Section 28.9). The Chi squared random variable is linearly related to the gamma random variable. This paper deals with the problem of estimating the non-centrality parameter of a gamma distribution with the squared error as loss function.

There is a standard argument for the choice of a quadratic loss in a general estimation problem. In relation to the Chi squared distribution, often the parameter of direct interest is an increasing function $g(\lambda)$, say, of the non-centrality parameter λ. Let T_1 and T_2 be two estimators of λ. It is known (see Rao, 1973) that a necessary condition for $P\{g(\lambda)-\varepsilon_1 < g(T_1) < g(\lambda)+\varepsilon_2\} \geq P\{g(\lambda)-\varepsilon_1 < g(T_2) < g(\lambda)+\varepsilon_2\}$ for all ε_1 and ε_2 is that $E(T_1-\lambda)^2 \leq E(T_2-\lambda)^2$. This result gives an added argument for the choice of a quadratic loss.

Let X be distributed according to a noncentral gamma distribution with p degrees of freedom and noncentrality parameter λ. It is known that $X-p$ is a uniformly minimum variance unbiased estimator (UMVUE) of λ, whereas $(X-p)^+ = \max(X-p, 0)$ has smaller mean squared error (MSE) than the UMVUE. But $(X-p)^+$ itself is inadmissible since it is not an analytic function of X. More generally $(X-c)^+$ is inadmissible for $c > 0$. We note here, without proof, that the class of estimators $\{(X-c)^+, c \geq p\}$ is irreducible in the sense that there are no two values of c for which one estimator dominates the other.

Perlman and Rasmussen (1975) and Neff and Strawderman (1976) have considered a class of estimators of the form

$$x - p + b/(x+c)^a, \tag{1.1}$$

which have been shown to have smaller MSE than the UMVUE for certain values of the constants in (1.1) for $\lambda > 0$. The values of the constants should be constrained so that $\delta(x) \geq 0$. Otherwise, the positive part of $\delta(x)$ would improve upon $\delta(x)$. Clearly $(x-p)^+$ has smaller MSE than (1.1) in a neighborhood of $\lambda = 0$, provided $b > 0$.

In Section 2 we consider the maximum likelihood estimator (MLE) of λ. The MLE cannot be produced in a closed form (see Meyer, 1976). The maximization of the likelihood

Received January 1981; revised December 1981.

[1] The authors' work was supported by the Office of Naval Research under Contract N00014-75-C-0451.

AMS 1980 *subject classification.* Primary 62C15; secondary 62F10.

Key words and phrases. Chi squared distribution, non-centrality parameter, maximum likelihood, admissible and minimax estimators.

function leads to the solution of an equation involving modified Bessel functions. However our main result shows that the MLE is dominated by $(x - p)^+$. In the derivation of this result we have used certain properties of the Bessel function which are given in the Appendix.

In Section 3 we consider certain admissible and minimax estimators.

2. Maximum likelihood estimator. The noncentral gamma distribution is given by the density function

$$f_\lambda(x) = x^{p-1}e^{-\lambda-x}\sum_{r=0}^{\infty}\frac{(\lambda x)^r}{r!\,\Gamma(p+r)} \tag{2.1}$$
$$= e^{-\lambda-x}(x/\lambda)^{(p-1)/2}I_{p-1}(2\sqrt{\lambda x}), \qquad x > 0,$$

where $I_p(x)$, given by (A.1) in the Appendix, denotes the modified Bessel function. If X has the distribution (2.1) then $2X$ is distributed according to the Chi squared distribution with $2p$ degrees of freedom and non-centrality parameter 2λ. We consider the sample size $n = 1$ and $n > 1$ separately. The main result is in Theorem 2.1 (for $n = 1$) which shows that the MLE is dominated by $(x - p)^+$ for $p \geq \frac{1}{2}$.

A. *Sample size* $n = 1$. Let λ^* denote the MLE. The value of λ^* is obtained by maximizing (2.1) with respect to λ and is given by the solution of the following equation:

$$1 = \frac{x}{\sqrt{\lambda x}}I_p(2\sqrt{\lambda x})/I_{p-1}(2\sqrt{\lambda x}). \tag{2.2}$$

The right side of (2.2) is decreasing in λ by Lemma A.2 and tends to x/p as $\lambda \to 0$. Therefore for $x \leq p$, $\lambda^* = 0$ and for $x > p$, λ^* is the unique solution of (2.2).

LEMMA 2.1. (a) *For* $x \geq p \geq \frac{1}{2}$, $\lambda^* - (x - p)$ *is increasing in* x, *and* $x - p \leq \lambda^* \leq x - p + \frac{1}{2}$. (b) *For large* x, $\lambda^* = x - p + \frac{1}{2} + O(1/x)$.

PROOF. Proof of (b) follows by using (A.2) in (2.2). Details are not given. Proof of (a): From (2.2) and Lemma A.3 we have

$$\lambda = \sqrt{\lambda x}\,I_p(2\sqrt{\lambda x})\bigg/\left\{\frac{p}{\sqrt{\lambda x}}I_p(2\sqrt{\lambda x}) + I_{p+1}(2\sqrt{\lambda x})\right\}$$
$$\geq \sqrt{\lambda x}\,I_{p-1}(2\sqrt{\lambda x})\bigg/\left\{\frac{p}{\sqrt{\lambda x}}I_{p-1}(2\sqrt{\lambda x}) + I_p(2\sqrt{\lambda x})\right\} \tag{2.3}$$
$$= \sqrt{\lambda x}\bigg/\left\{\frac{p}{\sqrt{\lambda x}} + \frac{\sqrt{\lambda}}{\sqrt{x}}\right\}.$$

The last equality gives

$$\lambda^* \geq x - p. \tag{2.4}$$

Let $Z = 2(\lambda^* x)^{1/2}$ and write λ for λ^*. Then

$$\lambda + x\frac{d\lambda}{dx} = \frac{Z}{2}\frac{dZ}{dx} \tag{2.5}$$

and from (2.2)

$$ZI_{p-1}(Z) = 2xI_p(Z). \tag{2.6}$$

Differentiating both sides of (2.6) with respect to x, we get after simplification that

$$\frac{Z}{2}\frac{dZ}{dx} = \frac{\lambda}{\lambda - x + p}. \tag{2.7}$$

Hence, from (2.5) and (2.7) we get

$$x\frac{d\lambda}{dx} = \lambda\left(\frac{1}{\lambda - x + p} - 1\right). \tag{2.8}$$

Differentiating (2.8) with respect to x we get

$$x\frac{d^2\lambda}{dx^2} = -\left\{2 + \frac{x - p}{(\lambda - x + p)^2}\right\}\frac{d\lambda}{dx} + \frac{\lambda}{(\lambda - x + p)^2}. \tag{2.9}$$

Consider the behavior of λ as a function of x. From the first equality in (2.3) we have

$$\frac{\sqrt{\lambda x}\, I_{p+1}(2\sqrt{\lambda x})}{I_p(2\sqrt{\lambda x})} = x - p.$$

The above relation shows that $\lambda \to 0$ and $\lambda/(x - p) \to (p + 1)/p$ as $x \to p + 0$. Then from (2.8) we have that $d\lambda/dx \to (p + 1)/p$, as $x \to p + 0$. Let x_0 denote the smallest value of $x > p$ for which $d\lambda/dx = 0$. Then $d^2\lambda/dx^2 \leq 0$ for $x = x_0$. On the other hand, from (2.9) it is seen that $d^2\lambda/dx^2 > 0$ at $x = x_0$. Therefore $d\lambda/dx > 0$ for all $x > p$. Let x^0 be the smallest values of $x > p$ for which $d\lambda/dx = 1$. For $p > ½$, from (2.8) we find that

$$1 - 2\lambda + 2x^0 - 2p > 0. \tag{2.10}$$

Putting $d\lambda/dx = 1$ in (2.9), after simplification, we get

$$x^0(\lambda - x^0 + p)^2\left.\frac{d^2\lambda}{dx^2}\right|_{x=x^0} = (\lambda - x^0 + p)(1 - 2\lambda + 2x^0 - 2p) > 0 \tag{2.11}$$

due to (2.4) and (2.10). Since $d\lambda/dx \to (p + 1)/p$ as $x \to p + 0$, it follows that $d^2\lambda/dx^2 \leq 0$ at $x = x^0$, contrary to (2.11). Therefore,

$$\frac{d\lambda}{dx} > 1 \quad \text{for all} \quad x > p > ½. \tag{2.12}$$

Hence, $\lambda^* - (x - p)^*$ is increasing in x. Furthermore, using (2.12) in (2.8) we obtain

$$\lambda^* < (x - p)^+ + ½ \quad \text{for} \quad p > ½. \tag{2.13}$$

THEOREM 2.1. *For $p \geq ½$, λ^* is dominated by $(x - p)^+$.*

PROOF. We have

$$\begin{aligned} \text{MSE}(\lambda^*) - \text{MSE}(x - p)^+ &= E\{\lambda^* - (x - p)^+\}\{\lambda^* + (x - p)^+ - 2\lambda\} \\ &\geq E\{\lambda^* - (x - p)^+\}E\{\lambda^* + (x - p)^+ - 2\lambda\} > 0. \end{aligned} \tag{2.14}$$

The first inequality in (2.14) follows from the fact that each of the quantities $\lambda^* - (x - p)^+$ and $\lambda^* + (x + p)^+ - 2\lambda$ is increasing in x. The second inequality follows from (2.4) and from $E\{\lambda^* + (x - p)^+ - 2\lambda\} > 2E(x - p - \lambda) = 0$. This completes the proof of Theorem 2.1.

The Table 1 gives the mean squared errors of λ^* and $(X - p)^+$ for a few values of λ and p. It is seen from the table that $(X - p)^+$ has smaller MSE than λ^*. The relative difference between the MSE values is less than 5% for moderately large values of p and λ.

B. *Sample size $n > 1$.* The computation of the MLE is fairly difficult for $n > 1$. We give below lower and upper bounds on its value. Let $x_1, \cdots, x_n$ denote the sample values. The likelihood equation is given by

$$n = \sum_{i=1}^{n} \frac{x_i}{\sqrt{\lambda x_i}}\{I_p(2\sqrt{\lambda x_i})/I_{p-1}(2\sqrt{\lambda x_i})\}. \tag{2.15}$$

TABLE 1
Mean squared errors of λ^* *and* $(X-p)^+$

p	$\lambda = 0.5$		$\lambda = 1$		$\lambda = 5$		$\lambda = 10$	
	MSE λ^*	MSE $(X-p)^+$	MSE λ^*	MSE $(X-p)^+$	MSE λ^*	MSE $(X-p)^+$	MSE λ^*	MSE $(X-p)^+$
0.5	1.79	1.27	2.88	2.25	10.85	10.46	20.76	20.50
1	2.10	1.56	3.18	2.52	11.34	10.91	21.32	21.00
5	4.43	3.81	5.45	4.74	15.03	14.11	25.57	24.90
10	7.20	6.54	8.19	7.45	18.67	17.63	30.47	29.55
20	12.58	11.87	13.55	12.77	25.04	23.93	39.26	38.12

Each term of the summation on the right hand side of (2.15) is monotone decreasing in λ by Lemma A.2, and the maximum value of the sum corresponding to $\lambda \to 0$ is equal to $n\bar{x}/p$, where $\bar{x} = \sum x_i/n$ denotes the sample mean. Denoting the MLE by λ_n^*, we have that λ_n^* is uniquely given as a solution of (2.15) for $\bar{x} > p$ and that $\lambda_n^* = 0$ for $\bar{x} \le p$. The lower and upper bounds for λ_n^* are given by

$$\left(\frac{1}{n}\sum_{i=1}^n \sqrt{x_i}\right)^2 - p \le \lambda_n^* \le \min\left\{\frac{p+1}{p}(\bar{x}-p), \left(\frac{1}{n}\sum_{i=1}^n \sqrt{x_i}\right)^2\right\}. \tag{2.16}$$

We omit the proofs. For large sample values, λ_n^* is approximated by $(\sum\sqrt{x_i}/n)^2$. We have not been able to show that the lower bound in (2.16) can be made sharper to $(\bar{x}-p)^+$, but we conjecture it to be true.

3. Admissible and minimax estimators. Consider a prior distribution for λ given by

$$g(\lambda) = c^p\lambda^{p-1}e^{-c\lambda}/\Gamma(p), \qquad c, p, \lambda > 0.$$

Then

$$\delta(X) = \frac{p}{1+c} + \frac{X}{(1+c)^2}$$

is a Bayes estimator. The MSE and the Bayes risk $\rho(c)$ of $\delta(X)$ are given by

$$\begin{aligned} \text{MSE}(\delta) &= (1+c)^{-4}\{p + 2\lambda + (2+c)^2(p-c\lambda)^2\} \\ \rho(c) &= (1+c)^{-4}\left\{p + \frac{2p}{c} + p(2+c)^2\right\}. \end{aligned} \tag{3.1}$$

As the prior distribution assigns positive probability to every open interval, and as the Bayes risk is finite, it follows that $\delta(X)$ is admissible. It is interesting to observe that the limit of $\delta(X)$ as $c \to 0$ is $X + p$ which is a Bayes estimator with respect to an improper prior distribution and is dominated by UMVUE; see deWaal (1974) and Perlman and Rasmussen (1975). From (3.1) we see that $\rho(c) \to \infty$ as $c \to 0$. Since $\delta(X)$ is a proper Bayes estimator, it follows that all estimators have unbounded maximum risk. Therefore all estimators are trivially minimax. We show below that $(X-p)^+$ is minimax (nontrivially) for another loss function.

Let the squared error (SE) loss be changed to (SE)$/(\lambda + p/2)$. Then $\delta(X)$ is again a Bayes estimator when the prior distribution of λ is given by the density $h(\lambda) = \{c(2\lambda + p)/p(2+c)\}g(\lambda)$. Moreover, the UMVUE has a constant risk, equal to 2. The Bayes risk of $\delta(X)$ is now given by $\rho^*(c) = \{2c/(2p + pc)\}\rho(c)$. Because $\rho^*(c) \to 2$ as $c \to 0$, the UMVUE is minimax. Since $(X-p)^+$ dominates the UMVUE, it is also minimax.

APPENDIX

Let $I_p(x)$ denote the modified Bessel function:

$$I_p(x) = \left(\frac{1}{2}x\right)^p \sum_{k=0}^{\infty} \frac{(x^2/4)^k}{k!\Gamma(p+k+1)}. \tag{A.1}$$

For large values of x, we have, with $\mu = 4p^2$,

$$I_p(x) \sim \frac{e^x}{\sqrt{(2\pi x)}}\left\{1 - \frac{\mu-1}{8x} + \frac{(\mu-1)(\mu-9)}{2!(8x)^2} - \cdots\right\}. \tag{A.2}$$

Let $Q_p(x) = I_p(x)/I_{p-1}(x)$. The following lemmas can be derived easily from the known properties of Bessel functions (see Abramowitz and Stegun, 1970).

LEMMA A.1. *$Q_p(x) \le 1$ for $p \ge 1$ and $x > 0$.*

LEMMA A.2. *$xQ_p(x)$ is increasing in x and $Q_p(x)/x$ is decreasing in x.*

LEMMA A.3. *$Q_p(x) > Q_{p+1}(x)$ for all p and $x > 0$.*

Acknowledgment. The authors wish to acknowledge the help of the associate editor, the referees, and the editor in revising the paper. They are also grateful to Professor Thomas S. Shores for computing Table 1.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (1970). *Handbook of Mathematical Functions.* Dover, New York.

DEWAAL, D. U. (1974). Bayes estimate of the noncentrality parameter in multivariate analysis. *Comm. Statist.* **3** 73–79.

JOHNSON, N. L. and KOTZ, S. (1970). *Continuous Univariate Distributions*-2, Houghton Mifflin, Boston.

MEYER, P. L. (1967). The maximum likelihood estimate of the non-centrality parameter of a non-central chi-squared variable. *Amer. Statist. Assoc.* **62** 1258–1264.

NEFF, N. and STRAWDERMAN, W. E. (1976). Further remarks on estimating the parameter of a non-central Chi squared distribution. *Comm. Statist.* **A5** 65–76.

PERLMAN, M. D. and RASMUSSEN, V. A. (1975). Some remarks on estimating a noncentrality parameter. *Comm. Statist.* **4** 455–468.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications.* Wiley, New York.

SPRUILL, C. (1979). Estimation of the non-centrality parameter of a chi-squared distribution. Unpublished Report, School of Mathematics, Georgia School of Technology.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEBRASKA—LINCOLN
LINCOLN, NEBRASKA 68588-0323

DEPARTMENT OF MATHEMATICAL SCIENCES
CLEMSON UNIVERSITY
CLEMSON, SOUTH CAROLINA 29631