ON INFORMATION FROM FOLDED DISTRIBUTIONS

A. V. KHARSHIKAR
University of Poona
(*Received in April 1968*)

1. INTRODUCTION

The distribution of $| x |$ or $| x-a |$ is known as folded distribution. Sometimes in practice the algebraic signs are either lost or not recorded or not available. In such situations folded distributions can be applied.

Leone et al [3] have studied the folded normal distribution. They give estimators for the two parameters making use of first two moments. Elandt [1] uses second and fourth moments for this purpose and further compares these estimators with those given in [3]. Johnson [2] has derived the accuracy of maximum likelihood estimators for folded normal distribution.

Purpose of this note is to show that the average amount of information about the known parameters from the folded distribution is always less than or at the most equal to that from the original distribution. A necessary and sufficient condition is given for the equality to hold good. Here the term amount of information is used as the value of the determinant of the information matrix.

2. THE INFORMATION MATRICES

It may be noted first that if a distribution is folded at some known point 'a', then without loss of generality 'a' may be taken to be equal to zero, as is done here.

Let $g(x, \theta_1, ..., \theta_k)$ be the probability density function (p.d.f.) of a random variate X, where $-\infty < x < \infty$ and $\theta_1, ..., \theta_k$ are k unknown parameters, $\theta_1, ..., \theta_k \varepsilon$-.

For the sake of convenience we write $g(x; \theta)$ for $g(x; \theta_1, ..., \theta_k)$.

The information matrix $A=(a_{ij})$ of this distribution is given by

$$(a_{ij}) = \varepsilon_g\left[-\frac{\delta}{\delta\theta_i \delta\theta_j} \log g(x, \theta)\right] \qquad ...(1)$$

$$i, j = 1, ..., k.$$

where by εg we denote mathematical expectation with respect to the p.d f. $g(x ; \theta)$. It is assumed that A is non-singular for all $\theta \varepsilon$-.

The p.d.f. of the corresponding folded variate is

$$f(x ; \theta)=g(x ; \theta)+g(-x ; \theta), \qquad 0 \leqslant x < \infty \qquad \ldots (2)$$

The informntion matrix $B=(b_{ij})$ of this distribution is

$$(b_{ij}) = \varepsilon_f \left[-\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log f(x ; \theta) \right] \qquad .. (3)$$

$$i, j = 1, \ldots, k.$$

Lemma (1)

The matrix $A-B$ is positive semi-definite (p.s.d.)

Proof :

We shall write g_+ for $g(x, \theta)$ and g_- for $g(-x, \theta)$

where $0 \leqslant x < \infty$.

$$a_{ij} = -\int_{-\infty}^{\infty} \left[\frac{\delta^2}{\delta\theta_i \delta\theta_j} \log g(x ; \theta) \right] g(x ; \theta) dx$$

$$= -\int_{-\infty}^{\infty} \frac{\delta^2}{\delta\theta_i \delta\theta_j} g(x ; \theta) dx$$

$$+ \int_{-\infty}^{\infty} \frac{\left[\frac{\delta}{\delta\theta_i} g(x ; \theta) \right] \left[\frac{\delta}{\delta\theta_j} g(x ; \theta) \right]}{g(x ; \theta)} dx$$

$$= -\int_0^{\infty} \frac{\delta^2}{\delta\theta_i \delta\theta_j} g(g_+ + q_{-1}) dx$$

$$+ \int_0^{\infty} \frac{\left(\frac{\delta}{\delta\theta_i} g_+ \right) \left(\frac{\delta}{\delta\theta_j} g_+ \right)}{g_+} dx$$

$$+ \int_0^{\infty} \frac{\left(\frac{\delta}{\delta\theta_i} g_- \right) \left(\frac{\delta}{\delta\theta_j} g_- \right)}{g_-} dx \qquad \ldots (4)$$

Similarly

$$b_{ij}=-\int_0^\infty \left[\frac{\delta^2}{\delta\theta_i\delta\theta_j}\log(g_+ + g_-)\right](g_+ + g_-)dx$$

$$=-\int_0^\infty \frac{\delta^2}{\delta\theta_i\delta\theta_j}(g_+ + g_-)\,dx$$

$$+\int_0^\infty \frac{\left(\frac{\delta}{\delta\theta_i}g_+\right)\left(\frac{\delta}{\delta\theta_j}g_+\right)}{g_+ + g_-}dx$$

$$+\int_0^\infty \frac{\left(\frac{\delta}{\delta\theta_i}g_-\right)\left(\frac{\delta}{\delta\theta_j}g_-\right)}{g_+ + g_-}dx+\int_0^\infty \frac{\left(\frac{\delta}{\delta\theta_i}g_+\right)\left(\frac{\delta}{\delta\theta_j}g_-\right)}{g_+ + g_-}dx$$

$$+\int_0^\infty \frac{\left(\frac{\delta}{\delta\theta_i}g_-\right)\left(\frac{\delta}{\delta\theta_j}g_+\right)}{g_+ + g_-}dx \qquad \ldots(5)$$

Hence

$$e_{ij}=a_{ij}-b_{ij}=\int_0^\infty \left\{\sqrt{\frac{g_-}{g_+(g_+ + g_-)}}\left(\frac{\delta}{\delta\theta_i}g_+\right) - \sqrt{\frac{g_+}{g_-(g_+ + g_-)}}\left(\frac{\delta}{\delta\theta_i}g_-\right)\right\}\times$$

$$\left\{\sqrt{\frac{g_-}{g_+(g_+ + g_-)}}\left(\frac{\delta}{\delta\theta_j}g_+\right) - \sqrt{\frac{g_+}{g_-(g_+ + g_-)}}\left(\frac{\delta}{\delta\theta_i}g_-\right)\right\}dx$$

$$=\int_0^\infty M_i M_j\,dx \qquad \ldots(6)$$

where

$$M_i=\sqrt{\frac{g_-}{g_+(g_+ + g_-)}}\left(\frac{\delta}{\delta\theta_i}g_+\right) - \sqrt{\frac{g_+}{g_-(g_+ + g_-)}}\left(\frac{\delta}{\delta\theta_i}g_-\right) \qquad \ldots(7)$$

$$i=1,2\ldots\ldots k.$$

Now it is easy to see that

$E=(e_{ij})$ is p.s.d,

Theorem 1 :

The average amount of information from the original distribution is always greater than or at the most equal to that from the corresponding folded distribution.

Proof :

Since A is symmetric and non-singular, a non-singular matrix T exists, such that

$$A = T'T$$

Therefore

$$A - B = T'\{I - (T^{-1'} B T^{-1})\} T$$

where both $I - T'^{-1} B T^{-1}$ and $T^{-1} B T^{-1}$ are p.s.d.

Let λ_i $(i = 1, \ldots, k)$ be the eigenvalues of the matrix $T'^{-1} B T^{-1}$. Then $1 - \lambda_i (i = 1, \ldots, k)$ are the eigenvalues of the matrix $I - T'^{-1} B T^{-1}$.

Since both these matrices are p.s.d. we have

$$0 \leqslant \lambda_i \leqslant 1 \text{ for all } i = 1, \ldots, k. \quad .. (8)$$

Hence

$$0 \leqslant \prod_{i=1}^{k} \lambda_i \leqslant 1 \quad ...(9)$$

But $\prod_{i=1}^{k} \lambda_i = T'^{-1} B T^{-1} = \frac{| B |}{| A |}$

And therefore

$$| B | \leqslant | A |$$

Theorem 2 :

There is no loss of information due to folding if and only if

$$g_+ = c g_-$$

where c is independent of the parameters.

Proof :

It follows immediately that if $g_+ = c g_-$, then $| A | = | B |$.

On the other hand if $| A | = | B |$,

we have

$$\lambda_i = 1 \text{ for all } i = 1, \ldots, k.$$

or $$1 - \lambda_i = 0 \text{ for all } i = 1, \ldots, k.$$

Since $1-\lambda_i$ are the eigenvalues of the matrix $I-T'^{-1}BT^{-1}$, we must have

$$I-T'^{-1}BT^{-1}=0$$

where 0 is the null matrix.

It follows that

$$E=T'(I-T'^{-1}BT^{-1})T=0$$

This further implies that

$$M_i=0 \text{ for all } i=1,\ldots,k.$$

where M_i is given by (7).

Hence

$$\sqrt{\frac{g_-}{g_+(g_++g_-)}}\ \frac{\delta}{\delta\theta_i}g_+=\sqrt{\frac{g_+}{g_-(g_++g_-)}}\ \frac{\delta}{\delta\theta_i}g_- \qquad \ldots(4)$$

for all $i=1,\ldots,k$

i.e.,
$$\frac{\delta}{\delta\theta_i}\log g_+=\frac{\delta}{\delta\theta_i}\log g_-$$

for all $i=1,\ldots,k$.

i.e.,
$$g_+=cg.$$

where c is independent of the parameters.

Remark : Results obtained here can be further generalised so as to include the cases when the range of the r.v. X is not the entire real line or when X is a discrete valued r.v.

3. Summary

It is shown that the average amount of information about the unknown parameters from the folded distribution is always less than or at the most equal to that from the original distribution. A necessary and sufficient condition is given for the equality to hold good.

4. Acknowledgement :

The author wishes to acknowledge with gratitude Dr. V S. Huzurbazar's encouragement and guidance. His sincere thanks are due to Dr. D.N. Shanbhag for many helpful suggestions. He also thanks the Government of India for awarding him a Research Training Scholarship.

5. References :

[1] Elandt, Regina C. : 'The folded normal distribution—two methods of estimating parameters from moments', Technometrics, 3, (1961) 551-562.

[2] Johnson, N.L. : 'The folded normal distribution III: Accuracy of estimation by maximum likelihood. Technometrics, 4, (1962), 249–256.

[3] Leone, F.C., Nelson, L.S., and Nottingham, R.B. : 'The folded normal distribution', Technometrics, 3, (1961), 543-550.