Institute of Mathematical Statistics

LECTURE NOTES — MONOGRAPH SERIES

NUISANCE PARAMETER ELIMINATION AND OPTIMAL ESTIMATING FUNCTIONS

T. M. Durairajan and Martin L. William
Loyola College, Madras, India

Abstract

In the context of obtaining optimal estimating functions for interesting parameters in the presence of nuisance parameters in parametric models, a method of elimination of nuisance parameters is proposed in this paper. The proposed method is direct and does not impose any 'factorization' conditions on the likelihood. In this direction, a sequence of lower bounds for the variance-covariance matrix of estimating functions is derived. A recipe which gives a transparent approach for obtaining optimal estimating functions is suggested. It is shown that minimum variance unbiased estimators could be obtained using the recipe.

Keywords and Phrases: Lower bounds, nuisance parameter elimination, optimal estimating function.

1 Introduction

In the theory of estimating functions applied to parametric models involving nuisance parameters, the 'elimination' of nuisance parameters to obtain optimal estimating functions (EF) for interesting parameters is a very important task. In a pioneering work, Godambe (1976) suggested a method of eliminating nuisance parameters by multiplying and adding suitable functions to the score function and formally established the optimality of conditional score function. Lloyd (1987) and Bhapkar and Srinivasan (1993) claimed the optimality of marginal score function. However, that there are errors in the results of Lloyd (1987) and Bhapkar and Srinivasan (1993) has been pointed out by Bhapkar (1995, 1997) who imposed some more conditions and established the optimality of marginal score function. The conditional and marginal factorization properties were used by the above authors in the elimination of nuisance parameters. Heyde (1997) proposed a method of obtaining optimal EF by eliminating nuisance parameters from a suitably chosen function that possesses the 'likelihood score property'. Heyde gives the optimal EF of 'first order theory' but not of the higher orders. The present work is an attempt in this direction.

In this paper, a straight forward recursive method of elimination of nuisance parameters without going into the factorization aspects of the likelihood is proposed. In this direction, a theorem which gives a sequence of lower bounds for the variance-covariance matrix of the EFs is established in Section 2. This is achieved by considering higher order derivatives with respect to the nuisance parameters drawing inspiration from Godambe (1984). Consequently, a recipe which gives a systematic approach for possible elimination of nuisance parameters leading to optimal EF is suggested. Section 3 presents several examples to illustrate the recipe. In Section 4, as another outcome of the main result of Section 2, a sequence of lower bounds for the variance - covariance matrix of unbiased estimators of the interesting parameters is given. This sequence is different from the sequence of Bhattacharya bounds both in context and in content. Further, it is shown that minimum variance bound unbiased estimators of the interesting parameters could be obtained by the suggested recipe.

2 The Main Result and the Recipe

Let X be a random vector with sample space $\mathcal{X}$ and probability density function $p(x;\omega)$ with respect to some σ-finite measure μ on $(\mathcal{X}, B(\mathcal{X}))$. The family of densities is indexed by $\omega = (\theta, \phi) \in \Omega$ with $\theta \in \Omega_1 \subset \Re^r$, $\phi \in \Omega_2 \subset \Re^m$, $\Omega = \Omega_1 \times \Omega_2$. The interesting parameter is θ and the nuisance parameter is ϕ and estimation of θ in the presence of ϕ is considered.

We assume the usual regularity conditions on the density function 'p' and the EFs $g = (g_1, \ldots, g_r)'$: $\mathcal{X} \times \Omega_1 \to \Re^r$ (refer Godambe (1976, 1984), Bhapkar (1995, 1997)). Let $D_g = ((E(\partial g_i/\partial\theta_j)))$. Let the class of EFs satisfying the regularity conditions be denoted as $\mathcal{G}_0$ let $M_g(\omega) = D_g^{-1}E(g.g')(D_g')^{-1}$, the variance-covariance matrix of standardized EFs.

In the sequel, the following notations are used:

$$l_\theta = (\partial \log\ p/\partial\theta_1, \ldots, \partial \log p/\partial\theta_r)' \tag{2.1}$$

$$I_{11} = E(l_\theta \cdot l_\theta') \tag{2.2}$$

$$l_\phi^{(i)} = \left(\frac{1\partial^i p}{p\partial\phi_1^i}, \ldots, \frac{1\partial^i p}{p\partial\phi_m^i}\right)', \quad i = 1, 2, \ldots \tag{2.3}$$

$$L_\theta^{(0)} = l_\theta, L_\theta^{(i)} = L_\theta^{(i-1)} - I_{12}^{(i)} I_{22}^{(i)^{-1}} l_\phi^{(i)} \tag{2.4}$$

with

$$I_{12}^{(i)} = E\left[L_\theta^{(i-1)} l_\phi^{(i)\prime}\right] = I_{21}^{(i)\prime}, \qquad I_{22}^{(i)} = E\left[l_\phi^{(i)} \cdot l_\phi^{(i)\prime}\right] \tag{2.5}$$

I_{11} and $I_{22}^{(i)}$ are assumed non-singular and for simplicity, we write $l_\phi^{(1)} = l_\phi$, $I_{12}^{(1)} = I_{12}$, $I_{21}^{(1)} = I_{21}$, $I_{22}^{(1)} = I_{22}$. Let

$$B_k = \left[I_{11} - \sum_{i=1}^{k} I_{12}^{(i)} I_{22}^{(i)^{-1}} I_{21}^{(i)}\right]^{-1}, \quad k = 1, 2, \ldots \tag{2.6}$$

Since $I_{22}^{(i)}$ are positive definite we have $B_{k+1} \geq B_k$. Also,

$$E\lfloor L_\theta^{(k)} \cdot L_\theta^{(k)\prime} \rfloor = B_k^{-1} \quad \forall\ k. \tag{2.7}$$

Theorem 2.1: For every $g \in \mathcal{G}_0$, $M_g \geq B_k$, $k = 1, 2, \ldots$, with equality if and only if $g = g^{(k)^*} = A(\theta, \phi) \cdot L_\theta^{(k)}$ where $A(\theta, \phi)$ is a non-singluar matrix and the functions $L_\theta^{(k)}$ are defined recursively in (2.4) and (2.5).

Proof: For $g \in \mathcal{G}_0$, we observe that $E[g.l_\phi^{(k)\prime}] = 0$, $E[g.L_\theta^{(k)\prime}] = -D_g \forall k$. Now, considering the n.n.d. matrix

$$E\begin{bmatrix} L_\theta^{(k)} \\ g \end{bmatrix} \begin{bmatrix} L_\theta^{(k)\prime} & g' \end{bmatrix} = \begin{bmatrix} B_k^{-1} & -D_g' \\ -D_g & E(g.g') \end{bmatrix} \tag{2.8}$$

and applying matrix theory arguments, we get

$$\text{Rank} \begin{bmatrix} B_k^{-1} & -D_g' \\ -D_g & E(g.g') \end{bmatrix} = \text{Rank} \begin{bmatrix} B_k^{-1} - M_g^{-1} & 0 \\ 0 & E(g.g') \end{bmatrix}.$$

Also, $B_k^{-1} - M_g^{-1} \geq 0$ by the non-negative definiteness of the full matrix in (2.8). This gives $M_g \geq B_k$. Further, $B_k^{-1} - M_g^{-1} = 0$ if and only if the rank of the matrix in (2.8) is r.

Now, if $g = A(\theta, \phi) \cdot L_\theta^{(k)}$ for some non-singular matrix $A(\theta, \phi)$, then clearly the matrix in (2.8) has rank r so that $M_g = B_k$. Conversely, if $M_g = B_k$ i.e. the rank of the matrix in (2.8) is r, then as B_k^{-1}, D_g and $E(g.g')$ are all non-singular, it is necessary that for some non-singular matrix $A(\theta, \phi)$, $g = A(\theta, \phi) \cdot L_\theta^{(k)}$. Hence the theorem.

Remark 1: If for some $k \geq 1$, B_k is attained by the EF $A(\theta, \phi) L_\theta^{(k)}$ for a suitable choice of $A(\theta, \phi)$ (i.e. $A(\theta, \phi) L_\theta^{(k)}$ is free of ϕ), then we have $I_{12}^{(k+1)} = 0$ which gives $L_\theta^{(k+1)} = L_\theta^{(k)}$ and $B_{k+1} = B_k$. Similarly, $\forall\ s \geq k+1$, $I_{12}^{(s)} = 0$, $L_\theta^{(s)} = L_\theta^{(k)}$ and $B_s = B_k$.

In view of this, we now propose the following.

Definition 2.1: If there exists a $g^* \in \mathcal{G}_0$ such that M_{g^*} attains a lower bound B_k for some k, then g^* is said to be minimum variance bound estimating function (MVBEF).

Based on the theorem established and the above remark, we suggest the following recipe for eliminating nuisance parameters and obtaining MVBEF: "Starting with the score function l_θ, consider recursively the functions $L_\theta^{(1)} = l_\theta - I_{12}I_{22}^{-1}l_\phi$, $L_\theta^{(2)} = L_\theta^{(1)} - I_{12}^{(2)}I_{22}^{(2)^{-1}}l_\phi^{(2)}$ and so forth. If the nuisance parameters are essentially eliminated or appear as a multiplicative factor of an EF in some recursion, stop the process as the EF thus obtained is optimal and further recursions would result in the same EF".

Remark 2: The first order bound in the sequence of bounds derived above namely $B_1 = (I_{11} - I_{12}I_{22}^{-1}I_{21})^{-1}$ was derived as the lower bound by Chandrasekar and Kale (1984) who gave a different line of proof. In a latest book, Heyde (1997) has proposed a 'first order' theory for obtaining optimal EF. He suggests the possibility for higher order theory but does not give any explicit method for the same. The recipe suggested above gives a formal and transparent method to proceed to second and higher order theories and carries out Heyde's suggestion. However, the forms of the optimal EFs obtained in Theorem 2.1 do not follow as a consequence of the technique of Heyde (1997).

3 Applications

In this section, a number of examples are discussed to illustrate the recipe suggested in the previous section. Throughout this section, we reserve the symbol θ for the interesting (real or vector) parameter.

Example 3.1: Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be independent where x_i are i.i.d. with density $\phi \exp(-\phi x)$, $x > 0$ and y_i are i.i.d. with density $\phi\theta^{-1}\exp(-\phi\theta^{-1}y)$, $y > 0$, θ, $\phi > 0$. Here,

$$L_\theta^{(1)} = \frac{\phi}{2\theta^2}\left(\sum_{i=1}^n y_i - \theta\sum_{i=1}^n x_i\right) \text{ so that } g^{(1)^*} = (2\theta^2/\phi)\cdot L_\theta^{(1)}$$

is the MVBEF attaining the bound B_1.

Example 3.2: Let $z_1, \ldots, z_n$ be i.i.d. with $z_i = (x_i, y_{1i}, \ldots, y_{ri})$ where x_i are i.i.d. exponential with mean ϕ and for each fixed $j = 1, \ldots, r$, y_{ji} are i.i.d. exponential with mean $\phi\theta_j$. Denote $x = \sum\limits_{i=1}^{n} x_i$, $y_j = \sum_{i=1}^n y_{ji}$, $j = 1, \ldots, r$.

Here, $\phi L_\theta^{(1)} = (g_1^*, \ldots, g_r^*)$ is MVBEF, with $g_j^* = \frac{y_j}{\theta_j^2} - \frac{x}{(1+r)\theta_j} \frac{1}{(1+r)\theta_j} \sum_{j=1}^{r} \frac{y_j}{\theta_j}$.

Remark 3: Examples 3.1 and 3.2 have been discussed respectively by Lloyd (1987) and Bhapkar and Srinivasan (1993) in the context of marginal factorization of the likelihood. These authors claimed that a marginal score function is the optimal EF. However, from the above discussion, we find that the optimal EFs obtained above do not coincide with the EFs claimed by these authors as optimal. In Example 3.1, we have $M_{g^{(1)*}} = 2\theta^2/n$ whereas for the EF of Lloyd (1987) namely $g_0 = n/\theta - 2n/(\theta + \Sigma y_i/\Sigma x_i)$, we have $M_{g_0} = 2(n+1)\theta^2/n^2 > M_{g^{(1)*}}$. This shows that Lloyd's claim that g_0 is optimal EF is incorrect. The errors in Lloyd (1987) and Bhapkar and Srinivasan (1993) has been pointed out also by Bhapkar (1995, 1997) who has found the correct optimal EF for the model in Example 3.1 but not for Example 3.2. In contrast, the recipe of Section 2 and the explicit form $L_\theta^{(1)}$ for the optimal EF have enabled us to achieve this for Example 3.2 as well in an elegant manner.

Example 3.3: Let $x_1, y_1, \ldots, x_n, y_n$ be independent normal with $E(x_i) = \theta$, $E(y_i) = \theta + \phi_i$, $V(x_i) = V(y_i) = 1$. Here, $L_\theta^{(1)} = \sum_{i=1}^{n}(x_i - \theta)$ is the MVBEF.

Example 3.4: Let $x_1, y_1, \ldots, x_n, y_n$ be independent normal with $E(x_i) = \theta + \phi_i$, $E(y_i) = \phi_i$, $V(x_i) = V(y_i) = 1$. Here, the MVBEF is $L_\theta^{(1)} = n(\bar{x} - \bar{y} - \theta)/2$.

Remark 4: Examples 3.3 and 3.4 were discussed by Godambe (1976) in the context of nuisance parameter elimination when conditional factorization property for the likelihood holds good. He has shown that the same EFs are optimal. In the above discussion, we have demonstrated the straight forward applicability of our recipe without investigating the factorization aspects which is required in Godambe's approach.

Example 3.5: Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_m)$ be independent where x_i are i.i.d. Poisson $(\theta\phi)$, y_j are i.i.d. Poisson (ϕ), $\theta, \phi > 0$. Here, $L_\theta^{(1)} = mn(\bar{x} - \theta\bar{y})/(\theta(n\theta + m))$ is the MVBEF.

This example with $m = n = 1$ has been discussed by Reid (1995) in illustrating the roles of conditioning in inference in the presence of nuisance parameters, wherein the estimation is based on conditioning upon a statistic called a 'cut' (Barndorff-Nielsen 1978) and involves a suitable reparametrization. In contrast, our approach is straight forward and does not require reparametrization.

Example 3.6: Consider the linear model of a randomized block design $y_{ij} = \mu + t_i + b_j + \epsilon_{ij}$, $i = 1, \ldots, k$, $j = 1, \ldots, r$ where ϵ_{ij} are i.i.d. $N(0, \sigma^2)$,

$\Sigma t_i = \Sigma b_j = 0$. Suppose estimation of the effect of the first treatment 't_1' alone is of interest. That is $\theta = t_1$, $\phi = (\mu, t_1, \ldots, t_{k-1}, b_1, \ldots, b_{r-1}, \sigma^2)$. Here, $(\sigma^2(k-1)/(kr))L_\theta^{(1)} = \bar{y}_{1.} - \bar{y}_{..} - t_1$ is MVBEF, where $\bar{y}_{1.} = \sum_j y_{1j}/r$ and $\bar{y}_{..} = \sum_i \sum_j y_{ij}/(kr)$.

Example 3.7: Consider a 2×2 contingency table $((n_{ij}))$, $i, j = 1, 2$, following a multinomial distribution with fixed sample size $n = \Sigma\Sigma n_{ij}$ and with probabilities $((\pi_{ij}))$, $i, j = 1, 2$, $\Sigma\Sigma\pi_{ij} = 1$. Let $\theta = \pi_{11}$ be the parameter of interest with $\phi = (\pi_{12}, \pi_{21})$. Here $L_\theta^{(1)} = (n_{11} - n\theta)/(\theta(1-\theta))$ is the MVBEF.

This example was discussed by Bhapkar (1989) in investigating conditioning and loss of information in the presence of nuisance parameters.

Example 3.8: Let $\{X(t), t \geq 0\}$ and $\{Y(t), t \geq 0\}$ be two independent Poisson processes with parameters $(\theta + \phi)$ and ϕ respectively, θ, $\phi > 0$. Suppose data on the states of $\{X(t)\}$ at times $t_1, \ldots, t_m$, say, $x(t_1), \ldots, x(t_m)$ and data on the states of $\{Y(t)\}$ at times $s_1, \ldots, s_n$, say, $y(s_1), \ldots, y(s_n)$ are available. The likelihood is $C \exp\{-(\theta + \phi)t_m - \phi s_n\}\ (\theta + \phi)^{x(t_m)} \phi^{y(s_n)}$ and $L_\theta^{(1)}/(\theta + \phi) = s_n x(t_m) - t_m y(s_n) - \theta s_n t_m$ is the MVBEF.

Remark 5: In all the above Examples 3.1 to 3.8 the optimal EFs attain the bound B_1. The following are illustrations for EFs that attain the bound B_2.

Example 3.9: Let $x_1, \ldots, x_n$ be i.i.d. normal with mean vector $(\phi_1, \ldots, \phi_r)'$ and variance-covariance matrix $\Sigma = Diag\ [\sigma_1^2, \ldots, \sigma_r^2]$. Let $\theta = (\sigma_1^2, \ldots, \sigma_r^2)$ be the parameter of interest. Denote $x_i = (x_{i1}, \ldots, x_{ir})'$, $i = 1, \ldots, n$. Here $L_\theta^{(1)}$ is not an EF for θ. So we try the next function given by our recipe, $L_\theta^{(2)} = L_\theta^{(1)} - I_{12}^{(2)} I_{22}^{(2)^{-1}} l_\phi^{(2)}$. It is found that

$$L_\theta^{(2)} = \left(\frac{1}{2\sigma_1^4} \sum_{i=1}^n (x_{i_1} - \bar{x}_1)^2 - \frac{n-1}{2\sigma_1^2}, \ldots, \frac{1}{2\sigma_r^4} \sum_{i=1}^n (x_{i_r} - \bar{x}_r)^2 - \frac{n-1}{2\sigma_r^2} \right)'$$

Let $S_j^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{i_j} - \bar{x}_j)^2$. Choosing $A(\theta, \phi) = Diag \left(\frac{2\sigma_1^4}{n-1}, \ldots, \frac{2\sigma_r^4}{n-1} \right)$. We get $g^{(2)^*} = A(\theta, \phi) \cdot L_\theta^{(2)} = (S_1^2 - \sigma_1^2, \ldots, S_r^2 - \sigma_r^2)'$ as the MVBEF attaining the bound B_2.

This MVBEF was shown as optimal EF not attaining the bound B_1 by Chandrasekar and Kale (1984) and in their framework served as an example of an optimal EF which is not MVBEF. From our discussion above, it is found that $g^{(2)^*}$ is MVBEF attaining, however, the higher bound B_2.

Example 3.10: Let $x_1, y_1, \ldots, x_n, y_n$, be independent normal with $E(x_i) = E(y_i) = \phi_i$, $i = 1, \ldots, n$ and $V(x_i) = V(y_i) = \theta\ \ \forall i = 1, \ldots, n$. Here, $L_\theta^{(1)}$ is not an EF for θ and so we proceed to the second recursion to get

$$L_\theta^{(2)} = -\frac{n}{2\theta} + \frac{1}{4\theta^2}\sum_{i=1}^{n}(x_i - y_i)^2 \text{ as MVBEF attaining the bound } B_2.$$

The above example has been discussed by Godambe (1976) in the context of conditional factorization wherein he contrasts the consistency of the solution of the above optimum estimating equation with the inconsistency of the usual maximum likelihood estimate discussed by Neyman and Scott (1948). Again, investigation into factorization aspects is not required in our approach.

Example 3.11: Let $x_1, y_1, \ldots, x_n, y_n$ be independent normal with $E(x_i) = \theta$, $E(y_i) = \phi$, $V(x_i) = V(y_i) = \theta$, $\phi \in \Re$, $\theta > 0$. Here, $L_\theta^{(1)}$ is not an EF. So, we proceed to the second recursion to get $L_\theta^{(2)} = -\frac{1}{2\theta^2}[n\theta^2 + (2n-1)\theta - n(s_x^2 + s_y^2 + \bar{x}^2)]$ as the MVBEF attaining bound B_2. Here, $s_x^2 = \frac{1}{n}\Sigma(x_i - \bar{x})^2$ and $s_y^2 = \frac{1}{n}\Sigma(y_j - \bar{y})^2$. The solution of the equation $L_\theta^{(2)} = 0$ is a consistent estimate namely

$$\hat{\theta} = \frac{\left[(2n-1)^2 + 4n^2\left(s_x^2 + s_y^2 + \bar{x}^2\right)\right]^{1/2} - (2n-1)}{2n}.$$

4 Minimum Variance Unbiased Estimators

The main result of Section 2 leads us to minimum variance bounds for unbiased estimators of the interesting parameters. From Theorem 2.1, it is immediately evident that, for unbiased estimators $T = (T_1, \ldots, T_r)'$ of $\theta = (\theta_1, \ldots, \theta_r)'$

$$Var - Cov(T) \geq B_k, \quad k = 1, 2, \ldots$$

where B_k's are given in (2.6). This is verified by considering EFs of the form $T - \theta$.

If any of the $L_\theta^{(k)}$'s defined recursively in (2.4) and (2.5) is such that $A(\theta, \phi)L_\theta^{(k)}$ is of the form $T^* - \theta$ for a suitable choice of $A(\theta, \phi)$, then T^* is minimum variance unbiased estimator (MVUE) of θ. Thus, the recipe of Section 2 could possibly be of help in finding T^*. The following examples illuminate this point.

Example 4.1: Consider the model in Example 3.3. Here $L_\theta^{(1)}/n = \bar{x} - \theta$ so that $\bar{x}$ is MVUE of θ attaining bound B_1.

Example 4.2: Consider the model in Example 3.4. Here $2L_\theta^{(1)}/n = \bar{x} - \bar{y} - \theta$ so that $\bar{x} - \bar{y}$ is MVUE of θ attaining bound B_1.

Example 4.3: Consider the linear model in Example 3.6. Here, $(\sigma^2(k-1)/kr)L_\theta^{(1)} = \bar{y}_{1.} - \bar{y}_{..} - t_1$ so that $\bar{y}_{1.} - \bar{y}_{..}$ is MVUE of t_1.

Example 4.4: Consider Example 3.7. Hear, $\theta(1-\theta)L_\theta^{(1)}/n = n_{11}/n - \theta$ so that n_{11}/n is MVUE of θ.

Example 4.5: Consider Example 3.8. Here, $((\theta+\phi)s_n t_m)^{-1}L_\theta^{(1)} = 0$ gives $[s_n x(t_m) - t_m y(s_n)]/(s_n t_m)$ as MVUE of θ.

Example 4.6: Consider Example 3.9. Here, $(S_1^2, \ldots, S_r^2)$ is MVUE of $(\sigma_1^2, \ldots, \sigma_r^2)$ attaining bound B_2.

Example 4.7: Consider the Neyman-Scott model in Example 3.10. Here $2\theta^2 L_\theta^{(2)}/n = \Sigma(x_i - y_i)^2/(2n) - \theta$ so that $\Sigma(x_i - y_i)^2/(2n)$ is MVUE of θ attaining bound B_2.

It is noted that the MVBEF need not produce MVUE when it cannot be reduced to the form $T^* - \theta$. This is evident in Examples 3.1, 3.2, 3.5 and 3.11. This fact is also reported by Thavaneswaran and Abraham (1988) who considered EFs for non-linear time-series models.

Acknowledgement

We thank the referee for bringing to our notice some of the important references in this area and for the valuable comments which improved the quality and content of the paper substantially.

References

Barndorff-Nielsen, O.E. (1978) Information and Exponential Families in Statistical Theory. *Wiley*, New York.

Bhapkar, V.P. (1989). Conditioning on ancillary statistics and loss of information in the presence of nuisance parameters. *J. Statist. Plann. Inference*, 21, 139-160.

Bhapkar, V.P. (1990). Conditioning, marginalization and Fisher information functions. *Proc. R.C. Bose Symposium*, Delhi. (Ed. R.R. Bahadur). Wiley Eastern Limited, New Delhi, 123-136.

Bhapkar, V.P. (1991). Loss of information in the presence of nuisance parameters and partial sufficiency. *J. Statist. Plann. Inference*, 28, 185-203.

Bhapkar, V.P. (1995). Completeness and optimality of marginal likelihood estimating equations. *Comm. Statist. A - Theory Methods*, 24, 945-952.

Bhapkar, V.P. (1997). Estimating functions, partial sufficiency and Q-sufficiency in the presence of nuisance parameters. *Proc. Symposium on Estimating Functions*, Georgia. (Ed. I.V. Basawa, V.P. Godambe and R.L. Taylor). *Institute of Mathematical Statistics: Lecture Notes - Monograph Series.* 32, 83-104.

Bhapkar, V.P. and Srinivasan, C. (1993). Estimating functions : Fisher Information and optimality. *Probability and Statistics. Proc. First International Triennial Calcutta Symposium on Probability and Statistics.* (Ed. S.K. Basu and B.K. Sinha). Narosa Publishing House, New Delhi, 165-172.

Bhapkar, V.P. and Srinivasan, C. (1994). On Fisher information inequalities in the presence of nuisance parameters. *Ann. Inst. Statist. Math.* 46, 593-604.

Chandrasekar, B. and Kale, B.K. (1984). Unbiased statistical estimation functions for parameters in presence of nuisance parameters. *J. Statist. Plann. Inference*, 9, 45-54.

Godambe, V.P. (1976). Conditional likelihood and unconditional optimum estimating equations. *Biometrika*, 63, 277-284.

Godambe, V.P. (1984). On ancillarity and Fisher information in the presence of a nuisance parameter. *Biometrika*, 71, 626-629.

Heyde, C.C. (1997). Quasi-Likelihood and its application - A general approach to optimal parameter estimation. *Springer*, New York.

Lloyd, C.J. (1987). Optimality of marginal likelihood estimating equations. *Comm. Statist. A - Theory Methods*, 16, 1733-1741.

Neyman, J. and Scott, E.L. (1948). Consistent estimates based on partially consistent observations. *Econometrica*, 16, 1-32.

Reid, N. (1995). The roles of conditioning in inference. *Statistical Science*, 10, 138 - 157.

Thavaneswaran, A. and Abraham, B. (1988). Estimation for nonlinear time - series using estimating equations. *J. Time Ser. Anal.* 9, 99