Ann. Inst. Statist. Math.
30 (1978), Part A, 415-419

AN EXTENSION OF THE CRAMÉR-RAO INEQUALITY FOR A SEQUENTIAL PROCEDURE WITHOUT ASSUMING REGULARITY CONDITIONS

ANTONIO DORIVAL CAMPOS

(Received May 1, 1978)

1. Introduction

Chapman and Robbins [2] derived an expression for the lower bound of the variance of estimators, which does not depend upon the regularity conditions. Their derivation is the following:

Let $X=(X_1, X_2, \cdots, X_n)$ be a finite sequence of random variables and let $L_n(x; \theta)$, $\theta \in \Omega \subset R$ be the probability densities associated to X. Now, let us consider the following sets:

$$S(\theta)=\{x \mid L_n(x; \theta)>0\}$$

and

$$\bar{S}(\theta)=\{x \mid L_n(x; \theta)=0\} .$$

If θ and $\theta+h$ $(h \neq 0)$ are any two points in Ω such that

$$S(\theta+h) \subset S(\theta) , \tag{1.1}$$

then Chapman and Robbins showed that the following inequality holds:

$$\operatorname{Var}[\delta_n(X) \mid \theta] \geqq \frac{1}{\inf\limits_h \mathrm{E}(J/\theta)} \tag{1.2}$$

where: $\delta_n(X)$ is an unbiased estimator of θ,

$$J=J(\theta, h)=\frac{1}{h^2}\left\{\left[\frac{L_n(x; \theta+h)}{L_n(x; \theta)}\right]^2-1\right\}$$

and the inf being over all $h \neq 0$ such that

$$S(\theta+h) \subset S(\theta) .$$

It should be added that if the set $\{x \mid L_n(x; \theta)>0\}$ is independent of θ, then condition (1.1) is true for all $\theta+h \in \Omega$.

They also present some examples of the inequality (1.2) and its comparison with the Cramér-Rao [3] and [6] inequality.

In the same article it is mentioned the possibility of using the method to more general estimation problems.

Along the same lines, Fraser and Guttman [4] got bounds analogous to those of Battacharyya [1] in both one and more than one parameter cases without assuming regularity conditions.

Also Kiefer [5] has show through examples the use of yet another lower bound, which is longer than Chapman and Robbins.

2. Theorem

Let us consider the sequential procedure described by Wolfowitz [7] *and let $\delta_N(X)$ be an estimator for $\theta \in \Omega \subset R$, with finite variance and such that*

$$\mathrm{E}\,\{\delta_N(X)\} = \Psi(\theta) < \infty \qquad \textit{for all } \theta \in \Omega ,$$

that is,

$$\sum_{j=1}^{\infty} \int_{R_j} \delta_j(x) L_j(x;\theta) \prod_{i=1}^{j} dx_i = \Psi(\theta) , \tag{2.1}$$

where: $L_j(x;\theta)$ is joint density function of $X=(X_1, X_2, \cdots, X_j)$.

Now let θ_1 and θ_2 be any two distinct points in Ω such that:

$$L_j(x;\theta_2)=0 \qquad \textit{implies } L_j(x;\theta_1)=0 . \tag{2.2}$$

Assume that $\sum_{j=1}^{\infty} t_j(\theta_1, \theta_2)$ be absolutely convergent for all $\theta_i \in \Omega$, $i=1, 2$, where:

$$t_j(\theta_1, \theta_2) = \int_{R_j} \frac{\{L_j(x;\theta_1)\}^2}{L_j(x;\theta_2)} \prod_{i=1}^{j} dx_i , \tag{2.3}$$

with $L_j(x;\theta_2)>0$. Under the above conditions, we have:

$$\mathrm{Var}\,\{\delta_N(X)\,|\,\theta_0\} \geqq \max_{\theta \neq \theta_0} \frac{(\theta-\theta_0)^2}{\mathrm{E}\left\{\left[\dfrac{L_N(X;\theta)}{L_N(X;\theta_0)}\right]^2 - 1/\theta_0\right\}} \tag{2.4}$$

where θ and θ_0 are distinct points in Ω, such that

$$L_j(x;\theta_0)=0 \qquad \textit{implies } L_j(x;\theta)=0 .$$

PROOF. From (2.1), we have:

$$\sum_{j=1}^{\infty} \int_{R_j} L_j(x;\theta) \prod_{i=1}^{j} dx_i = 1 . \tag{2.5}$$

Take the difference between the two expressions in (2.5) and let $\theta=\theta_1$ and $\theta=\theta_2$, respectively. We thus have:

$$\sum_{j=1}^{\infty}\int_{R_j}\{L_j(x;\theta_1)-L_j(x;\theta_2)\}\prod_{i=1}^{j}dx_i=0 \tag{2.6}$$

for any $\theta_1, \theta_2 \in \Omega$ and such that (2.2) holds. Similarly, from (2.1) we have:

$$\sum_{j=1}^{\infty}\int_{R_j}\delta_j(x)\{L_j(x;\theta_1)-L_j(x;\theta_2)\}\prod_{i=1}^{j}dx_i=\Psi(\theta_1)-\Psi(\theta_2)\ . \tag{2.7}$$

Now, multiply both sides of (2.6) by $\Psi(\theta_2)$ and subtract it from the result in (2.7), we get:

$$\Psi(\theta_1)-\Psi(\theta_2)=\sum_{j=1}^{\infty}\int_{R_j}\{\delta_j(x)-\Psi(\theta_2)\}\{L_j(x;\theta_1)-L_j(x;\theta_2)\}\prod_{i=1}^{j}dx_i\ . \tag{2.8}$$

From this equation, it follows that

$$\Psi(\theta_1)-\Psi(\theta_2)=\mathrm{E}\left\{[\delta_N(X)-\Psi(\theta_2)]\left[\frac{L_N(X;\theta_1)-L_N(X;\theta_2)}{L_N(X;\theta_2)}\right]\Big/\theta=\theta_2\right\}\ . \tag{2.9}$$

So from equations (2.1) and (2.6), equation (2.9) can be written as

$$\Psi(\theta_1)-\Psi(\theta_2)=\mathrm{cov}\left\{\left[\delta_N(X),\ \frac{L_N(X;\theta_1)-L_N(X;\theta_2)}{L_N(X;\theta_2)}\right]\Big/\theta=\theta_2\right\}\ . \tag{2.10}$$

Letting ρ be the correlation coefficient of $\delta_N(X)$ and $(L_N(X;\theta_1)-L_N(X;\theta_2))/L_N(X;\theta_2)$ for $\theta=\theta_2$, we obtain:

$$\rho^2=\frac{\{\Psi(\theta_1)-\Psi(\theta_2)\}^2}{\mathrm{E}\ \{[\delta_N(X)-\Psi(\theta_2)]^2/\theta=\theta_2\}\ \mathrm{E}\left\{\left[\frac{L_N(X;\theta_1)-L_N(X;\theta_2)}{L_N(X;\theta_2)}\right]^2\Big/\theta=\theta_2\right\}} \tag{2.11}$$

which exists whenever the conditions on $\mathrm{Var}\,\{\delta_N(X)\}$ and the series obtained from (2.3) hold.

From (2.11) we obtain:

$$\mathrm{E}\ \{[\delta_N(X)-\Psi(\theta_2)]^2/\theta=\theta_2\}\geqq\frac{\{\Psi(\theta_1)-\Psi(\theta_2)\}^2}{\sum_{j=1}^{\infty}\int_{R_j}\left\{\frac{L_j(x;\theta_1)-L_j(x;\theta_2)}{L_j(x;\theta_2)}\right\}^2\prod_{i=1}^{j}dx_i}\ . \tag{2.12}$$

Some algebra on the integrand of the equation above leads to

$$\mathrm{Var}\ \{\delta_N(X)/\theta=\theta_2\}\geqq\frac{\{\Psi(\theta_1)-\Psi(\theta_2)\}^2}{\mathrm{E}\left\{\left[\frac{L_N(X;\theta_1)}{L_N(X;\theta_2)}\right]^2-1/\theta=\theta_2\right\}}\ . \tag{2.13}$$

The last one equation is valid for all $\theta_1, \theta_2 \in \Omega$ such that (2.2) holds.

Consequently, it follows that for $\theta_1=\theta$ and $\theta_2=\theta_0$ where θ_0 is the true value of the unknown parameter and such that (2.2) holds, we get:

$$\text{(2.14)} \qquad \text{Var}\,\{\delta_N(X)\,|\,\theta_0\} \geqq \max_{\theta} \frac{\{\Psi(\theta)-\Psi(\theta_0)\}^2}{\text{E}\left\{\left[\dfrac{L_N(X;\theta)}{L_N(X;\theta_0)}\right]^2 - 1/\theta_0\right\}},$$

where max is taken on values of $\theta \in \Omega$, $\theta \neq \theta_0$. Now, if $\delta_N(X)$ is an unbiased estimator of θ, (2.14) can be written as:

$$\text{Var}\,\{\delta_N(X)/\theta_0\} \geqq \max_{\theta\neq\theta_0} \frac{(\theta-\theta_0)^2}{\text{E}\left\{\left[\dfrac{L_N(X;\theta)}{L_N(X;\theta_0)}\right]^2 - 1/\theta_0\right\}}$$

where θ and θ_0 are distinct points in Ω, such that:

$$L_j(x;\theta_0)=0 \qquad \text{implies } L_j(x;\theta)=0\,.$$

3. Application

Consider inequality (2.4) for $N=n$ fixed and $\theta=\theta_0+h$ $(h\neq 0)$, such that

$$\text{(3.1)} \quad L_n(x;\theta_0)=0 \qquad \text{implies } L_n(x;\theta_0+h)=0, \text{ and } \text{E}\,\{\delta_n(X)\}=\theta_0\,.$$

We thus obtain:

$$\text{Var}\,\{\delta_N(X)\,|\,\theta_0\} \geqq \frac{1}{\inf\limits_{h\neq 0} \text{E}\,(J/\theta_0)},$$

which is inequality (1.2) obtained by Chapman and Robbins, where: (3.1) is precisely condition (1.1) and

$$J=J(\theta_0,h)=\frac{1}{h^2}\left\{\left[\frac{L_n(x;\theta_0+h)}{L_n(x;\theta_0)}\right]^2-1\right\}\,.$$

UNIVERSIDADE DE SÃO PAULO

REFERENCES

[1] Bhattacharyya, A. (1946). On some analogues to the amount of information and their uses in statistical estimation, *Sankhyā*, **8**, 1-14.

[2] Chapman, D. G. and Robbins, H. (1951). Minimum variance estimation without regularity assumptions, *Ann. Math. Statist.*, **22**, 581-586.

[3] Cramér, H. (1946). *Mathematical Methods of Statistics*, Princeton Univ. Press, Princeton.

[4] Fraser, D. A. S. and Guttman, I. (1952). Bhattacharyya bounds without regularity assumptions, *Ann. Math. Statist.*, **23**, 629-632.

[5] Kiefer, J. (1952). On minimum variance estimators, *Ann. Math. Statist.*, **23**, 627-629.

[6] Rao, C. R. (1945). Information and accuracy attainable in the estimation of statistical parameter, *Bull. Calcutta Math. Soc.*, **37**, 81-91.

[7] Wolfowitz, J. (1947). The efficiency of sequential estimates and Wald's equation for sequential processes, *Ann. Math. Statist.*, **18**, 215-230.