COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE
20,3 (1979)

SUFFICIENT CONDITIONS FOR THE ASYMPTOTIC EFFICIENCY OF ESTIMATES

Dana VORLÍČKOVÁ

Abstract: In this paper there are considered conditions, under which estimates of parameters of regular densities are efficient, with a special respect to the maximum likelihood estimates.

Key words and phrases: Regular density, asymptotic efficiency, Fisher information, maximum-likelihood estimate.

AMS: 62F20

---

1. Introduction. The lower bound for the variance of the unbiased estimates of unknown parameters - the Cramér-Rao bound - can be found in special families of densities. Unbiased estimates, the variance of which attains the lower bound, are called efficient. However, unbiased estimates need not exist, or they need not be the best ones. Nevertheless, it can be shown, that there exists such a family of densities, that only "almost" unbiased estimates of parameters are good and that the Cramér-Rao bound is "almost" attained by their variance.

Hájek's results concerning these estimates were published in the lecture-note [3], mostly without proofs. This contribution aims to fulfil this lack partially.

2. Asymptotic efficiency. Let us consider a density $f(x,\theta)$ (Lebesgue or discrete) with an unknown parameter θ and let us assume throughout the paper that $f(x,\theta)$ is regular in the following sense.

Definition. We say that a family of densities $f(x,\theta)$, $\theta \in A$, is regular, if $f(x,\theta)$ has for every $x \in R$ the continuous derivative $\dot{f}(x,\theta) = (\partial/\partial\theta)f(x,\theta)$, the Fisher information

$$I(\theta)=\int_{-\infty}^{\infty}\frac{(\dot{f}(x,\theta))^2}{f(x,\theta)}\,dx$$

is continuous in θ, $I(\theta)>0$ for every $\theta \in A$, A is an open set.

These regularity conditions are weaker than the ones usually introduced in the literature (Cramér, Rao). Asymptotic properties of estimates are investigated there for densities having the third derivatives with respect to θ. However, these assumptions are unnecessarily strong. The following proposition, which is a limit version of the Cramér-Rao inequality, holds for estimates of parameters in regular families of densities.

Theorem 1.1. Let T_n be an estimate of $\theta \in A \subset R$, corresponding to the random sample of a size n. Then,

$$(1) \qquad \lim_{\varepsilon\downarrow 0}\ \liminf_{n\to\infty}\ \sup_{|\bar\theta-\theta|<\varepsilon}\,[n\,E_{\bar\theta}(T_n-\bar\theta)^2] \geq I^{-1}(\theta)$$

holds for every $\theta \in A$. The equality in (1) is attained only if

$$(2) \qquad n^{1/2}\,[T_n-\theta-I^{-1}(\theta)Z_n(\theta)] \xrightarrow{P_\theta} 0,$$

where

$$Z_n(\theta) = n^{-1} \sum_{i=1}^{n} \frac{\dot f(x_i, \theta)}{f(x_i, \theta)} .$$

Proof. The assertion follows from [2], Theorem 4.1, when we choose $\ell(y) = y^2$ and use Theorem A 4. Q.E.D.

Definition. An estimate T_n is called asymptotically efficient, if it satisfies (2).

In [3], there is published a theorem (with a brief reference instead of a proof), which yields a criterion how to recognize whether an estimate is asymptotically efficient. We shall introduce the Theorem also here, with a little more detailed proof, because we can use it later for the proof of the asymptotic efficiency of maximum likelihood estimates.

Theorem 2.2. Let T_n be an estimate of θ. The necessary condition for the asymptotic efficiency of T_n is the asymptotic normality $(\theta, n^{-1} I^{-1}(\theta))$. The sufficient condition for the asymptotic efficiency of T_n is the asymptotic normality $(\theta_n, n^{-1} I^{-1}(\theta))$, for every $\theta_n \to \theta$ for $n \to \infty$, $\theta \in A$, where θ_n is a real value of a parameter θ under a sample of a size n.

Proof. It follows from Theorem A.4 of [2], that the regularity conditions imply conditions of the local asymptotic normality of [2] with $\Gamma_t = I(t)$ and $\Delta_{nt} = n^{1/2} Z_n(t)$. We choose again $\ell(y) = y^2$.

Necessity. The asymptotic efficiency means (2), so that the asymptotic distributions of $n^{1/2}(T_n - \theta)$ and $n^{1/2} I^{-1}(\theta) \times$ $\times Z_n(\theta)$, respectively, are the same.
Furthermore, according to [2],

$\mathcal{L}(n^{1/2}Z_n(\theta)) \longrightarrow N(0, n^{-1} I^{-1}(\theta))$.

Sufficiency. Let us assume that (2) does not hold. Then, there exist such a sequence θ_n that the relation (4.2) of Theorem 4.1 from [2] with $\theta = \theta_n$ and $t = \theta$ does not hold, either, so that $n^{1/2}(T_n - \theta_n)$ is not asymptotically $N(0, I^{-1}(\theta))$. Q.E.D.

3. Asymptotic efficiency of maximum likelihood estimates. The asymptotic efficiency of maximum likelihood estimates under the conditions, mentioned in Part 2, which were introduced by Cramér and Rao, respectively, is well-known. The same proposition holds for estimates of parameters of regular densities under a monotonity assumption.

Theorem 3.1. Let $\theta \in A \subset R$. If the function $\dot{f}(x,\theta)/f(x,\theta)$ is nonincreasing in θ for every x, the maximum likelihood estimate of θ is asymptotically efficient and it may be received as a solution of the equation

$$(3) \qquad \sum_{i=1}^{n} \frac{\dot{f}(x_i,\theta)}{f(x_i,\theta)} = 0.$$

Proof. The assumption that $\dot{f}(x,\theta)/f(x,\theta)$ is nonincreasing in θ for every x, implies that $\log f(x,\theta)$ is a concave function of θ. Then, the solution of (3) maximalizes $\log \prod_{i=1}^{n} f(x_i,\theta)$.

It follows from the fact that the density $f(x,\theta)$ is regular, that the assumptions of Proposition 6 of [1] are satisfied. (The function $(f(x,\theta))^{1/2}$ is studied in [2], Appendix.) According to this Proposition with probability tending to 1 for $n \longrightarrow \infty$ there exist such maximum likelihood estimates $\hat{\theta}_n$, that

$$\mathcal{L}(n^{1/2}(\hat{\theta}_n - \theta)/\theta) \to N(0, \frac{\gamma(\theta)}{(\sigma^2(\theta) + \gamma(\theta))^2}), \; n \to \infty,$$

where

$$\sigma^2(\theta) = \limsup_{t \to 0} t^{-2} \int (f^{1/2}(x, \theta) - f^{1/2}(x, \theta + t))^2 \, dx$$

and

$$\gamma(\theta) = E(X'(\theta))^2,$$

where $X'(t)$ is the mean square derivative of $X(t) =$ $= (f(.,t)/f(.,\theta))^{1/2} - 1$.

Applying the Appendix of [2] again, we can compute easily that $\gamma(\theta) = \frac{1}{4} I(\theta) = \sigma^2(\theta)$, so that

$$\mathcal{L}(n^{1/2}(\hat{\theta}_n - \theta)/\theta) \to N(0, I^{-1}(\theta)).$$

The asymptotic efficiency of $\hat{\theta}_n$ follows from Theorem 2.2 and from the regularity assumptions. Q.E.D.

Remark. Analogical propositions can be formulated for $\theta \in A \subset R^k$ (see [3]).

References

[1] LeCAM L.: On the assumptions used to prove asymptotic normality of maximum likelihood estimates, Ann. Math. Statist. 41(1970), 802-828.

[2] HÁJEK J.: Local asymptotic minimax and admissibility in estimation, Proc. of the Sixth Berkeley Symp. on Math. Stat. and Prob., Univ. of California Press, 1972.

[3] HÁJEK J., VORLÍČKOVÁ D.: Matematická statistika (skriptum), SPN 1977.

Matematicko-fyzikální fakulta
Universita Karlova
Sokolovská 83, 18600 Praha 8
Československo

(Oblatum 7.3.1979)