数理解析研究所講究録
第 327 巻 1978 年 52-58

A property of higher order asymptotically sufficient statistics

By Takeru Suzuki

1. Introduction. Suppose that n-dimensional random variable $z_n=(x_1,x_2,\dots,x_n)$ is distributed according to a probability distribution $P_{\theta,n}$ parameterised by $\theta\in\Theta\subset R^1$, and each x_i is independently and identically distributed. In LeCam[1] it was shown that every estimator t_n with the form $t_n=\hat{\theta}_n+n^{-1}\cdot I^{-1}(\hat{\theta}_n)\Phi_n^{(1)}(z_n,\hat{\theta}_n)$ ($I(\theta)$ means Fisher information number), which is constructed using a reasonable estimator $\hat{\theta}_n$ and the logarithmic derivative $\Phi_n^{(1)}(z_n,\theta)$ relative to θ of density of $P_{\theta,n}$, is asymptotically sufficient in the following sense; t_n is sufficient for a family $\{Q_{\theta,n}\ ;\theta\in\Theta\}$ of probability distributions and that

$$\lim_{n\to\infty}\|P_{\theta,n}-Q_{\theta,n}\|=0$$

uniformly on any compact set in Θ (where $\|\cdot\|$ means the totally variation of a measure). This implies that the statistic $(\hat{\theta}_n,\Phi_n^{(1)}(z_n,\hat{\theta}_n))$ is asymptotically sufficient up to order o(1). As a refinement of this result it will be shown in this paper that for $k\geq 1$ the statistic $t_n^*=(\hat{\theta}_n,\Phi_n^{(1)}(z_n,\hat{\theta}_n),\dots,\Phi_n^{(k)}(z_n,\hat{\theta}_n))$, where $\Phi_n^{(i)}(z_n,\theta)$ means the (i-1)-th derivative relative to θ of $\Phi_n^{(1)}(z_n,\theta)$, is asymptotically sufficient up to order $o(n^{-\frac{k-1}{2}})$ in the following sense; t_n^* is sufficient for a family $\{Q_{\theta,n}\ ;\theta\in\Theta\}$ and

$$\lim_{n\to\infty} n^{\frac{k-1}{2}}\|P_{\theta,n}-Q_{\theta,n}\|=0$$

uniformly on any compact subset of Θ. From our result it follows that if we use the maximum likelihood estimator $\hat{\theta}_n^*$ as the initial estimator $\hat{\theta}_n$ then the statistic $(\hat{\theta}_n^*,\Phi_n^{(2)}(z_n,\hat{\theta}_n^*),\dots,\Phi_n^{(k)}(z_n,\hat{\theta}_n^*))$ is asymptotically sufficient up to order $o(n^{-\frac{k-1}{2}})$. In Ghosh and Subramanyam [4] it was mentioned that for exponential family of distributions, $(\hat{\theta}_n^*, \Phi_n^{(2)}(z_n,\hat{\theta}_n^*), \Phi_n^{(3)}(z_n,\hat{\theta}_n^*), \Phi_n^{(4)}(z_n,\hat{\theta}_n^*))$ is asymptotically sufficient up to order $o(n^{-1})$ in pointwise sense relative to θ. Our result is more general and accurate one.

As an application of our result we try to improve arbitrarily given statistical tests or estimators.

2. Notations and assumptions. Let $\Theta (\neq \phi)$ be an open set in R^1. Suppose that for each $\theta \in \Theta$ there corresponds a probability measure P_θ defined on a measurable space $(\underset{\sim}{X},\underset{\sim}{A})$. For each $n \in N=\{1,2,\dots\}$ let $(\underset{\sim}{X}^{(n)},\underset{\sim}{A}^{(n)})$ be the cartesian product of n copies of $(\underset{\sim}{X},\underset{\sim}{A})$, and $P_{\theta,n}$ the product measure of n copies of P_θ. For a function h and a probability measure P, E[h;P] stands for the expectation of h under P.

We assume that the map: $\theta \longmapsto P_\theta$ is one to one, and that for each $\theta \in \Theta$ P_θ has a density $f(\cdot,\theta)$ relative to a σ-finite measure μ on $(\underset{\sim}{X},\underset{\sim}{A})$. We assume also that $f(x,\theta)>0$ for every $x \in \underset{\sim}{X}$ and every $\theta \in \Theta$. We denote by μ_n the product measure of n copies of the same component μ. We define $\Phi(x,\theta)=\log f(x,\theta)$ for each $x \in \underset{\sim}{X}$ and $\theta \in \Theta$, and $\Phi_n(z_n,\theta)=\sum_{\nu=1}^{n}\Phi(x_\nu,\theta)$ for each $n \in N$, each $z_n=(x_1,x_2,\dots,x_n) \in \underset{\sim}{X}^{(n)}$ and each $\theta \in \Theta$. For a positive integer k we consider the following conditions which will be called Condition(C_k) in this paper.

Condition(C_k). (1). $\Phi(x,\theta)$ is (k+2)-times continuously differentiable with respect to θ in Θ for each $x \in \underset{\sim}{X}$. For each j $(1 \leq j \leq k+2)$ we define $\Phi^{(j)}(x,\theta)=\partial^j\Phi(x,\theta)/\partial\theta^j$ and $\Phi_n^{(j)}(z_n,\theta)=\sum_{\nu=1}^{n}\Phi^{(j)}(x_\nu,\theta)$.

(2). For each $\theta \in \Theta$ there exists a positive number ε such that

a. $\sup_{|\tau-\theta|\leq\varepsilon} E[\sup_{|\sigma-\theta|\leq\varepsilon} |\Phi^{(k+2)}(x,\sigma)|^2; P_\tau] < \infty$

b. $\sup_{|\tau-\theta|\leq\varepsilon} E[|\Phi^{(k+1)}(x,\tau)| \cdot u_\varepsilon(x,\tau); P_\tau] < \infty$ and $E[u_\varepsilon(x,\theta); P_\theta] < \infty$

where $u_\varepsilon(x,\tau) = \sup_{|\sigma-\tau|\leq\varepsilon} |f'(x,\sigma)/f(x,\tau)|$

c. $\mathrm{Var}(\Phi^{(k+1)}(x,\tau); P_\tau)$ are positive and finite uniformly for every τ satisfying $|\tau-\theta|\leq\varepsilon$.

(3). Define $\bar{Z}(x;\varepsilon',\sigma) = \sup\{\Phi^{(k+1)}(x,\tau) - E[\Phi^{(k+1)}(x,\tau); P_\tau]; \tau\in\Theta, |\tau-\sigma|\leq\varepsilon'\}$ and $Z^*(x;\varepsilon',\sigma) = -\inf\{\Phi^{(k+1)}(x,\tau) - E[\Phi^{(k+1)}(x,\tau); P_\tau]; \tau\in\Theta, |\tau-\sigma|\leq\varepsilon'\}$ for each $\varepsilon' > 0$ and $\sigma\in\Theta$. For each $\theta\in\Theta$ there exist positive numbers η and ρ such that for every $(t,\varepsilon',\sigma)\in(-\rho,\rho)\times(0,\eta]\times(\theta-\eta,\theta+\eta)$ the moment generating functions of $\bar{Z}(x;\varepsilon',\sigma)$ and $Z^*(x;\varepsilon',\sigma)$ exist and converge uniformly with respect to σ in $(\theta-\eta,\theta+\eta)$.

Remark 1. An example satisfying Condition(C_k) is the following one. Let μ be a σ-finite measure on $(\mathfrak{X},\mathfrak{A})$ and the density function $f(x,\theta)$ of P_θ relative to μ be given by

$$f(x,\theta) = h(x)c(\theta)\exp[\sum_{i=1}^{m} s_i(\theta)t_i(x)]$$

where $c(\theta), s_i(\theta)$ $(1\leq i\leq m)$ are $(k+2)$-times continuously differentiable real valued functions of θ only, and $h(x), t_i(x)$ $(1\leq i\leq m)$ are real valued $\mathfrak{A}$-measurable functions of x independent of θ. Let $S = \{(s_1, s_2, \ldots, s_m)\in R^m; \int_{\mathfrak{X}} \exp[\sum_{i=1}^{m} s_i t_i(x)]h(x)d\mu(x) < \infty\}$ and $S(\Theta) = \{(s_1(\theta), \ldots, s_m(\theta)); \theta\in\Theta\}$. If $S(\Theta) \subsetneqq \mathrm{int}\, S$ (interior of S) and if $\sum_{i=1}^{m}\sum_{j=1}^{m} s_i^{(k+1)}(\theta)s_j^{(k+1)}(\theta)\mathrm{Cov}(t_i, t_j; P_\theta) > 0$ for every $\theta\in\Theta$, then Condition(C_k) is satisfied by the family $\{P_\theta; \theta\in\Theta\}$. Here for each i $s_i^{(k+1)}$ means $(k+1)$-th derivative of s_i.

3. Asymptotically sufficient statistics up to higher orders. An estimator of θ depending on $z_n = (x_1, x_2, \ldots, x_n)\in\mathfrak{X}^{(n)}$ is an $\mathfrak{A}^{(n)}$-measurable function from $\mathfrak{X}^{(n)}$ to R^1. Such estimator will be called strict

if its range is a subset of Θ. For each δ satisfying $0<\delta<1/2$ we denote by $\underset{\sim}{C}(\delta)$ the class of all sequences $\{\hat{\theta}_n\}$ of strict estimators of θ such that for every compact subset K of Θ

$$\sup_{\theta\in K} P_{\theta,n}\ (\sqrt{n}\,|\hat{\theta}_n-\theta|>n^{\delta})=o(n^{-\frac{k-1}{2}}).$$

The notation $o(a_n)$ means that $\lim_{n\to\infty} o(a_n)/a_n=0$.

Remark 2. For every δ $(0<\delta<1/2)$ $\underset{\sim}{C}(\delta)$ does not empty (cf. Pfanzagl[2], Lemma 2). The maximum likelihood estimator is contained in $\bigcap_{\delta>0}\underset{\sim}{C}(\delta)$ under suitable regularity conditions (cf. Pfanzagle[3], Lemma 3).

Let $\delta_0=1/[2(k+2)]$ and $\underset{\sim}{C}=\bigcup_{0<\delta<\delta_0}\underset{\sim}{C}(\delta)$.

Theorem 1. Suppose that Condition(C_k) is satisfied, and that $\{\hat{\theta}_n\}\in\underset{\sim}{C}$ then there exists a sequence $\{Q_{\theta,n};\theta\in\Theta\}$, $n\in N$, of families of probability measures on $(\underset{\sim}{X}^{(n)},\underset{\sim}{A}^{(n)})$ with the following property:
(1) For each $n\in N$, the statistic $t_n^*=(\hat{\theta}_n,\Phi_n^{(1)}(z_n,\hat{\theta}_n),\ldots,\Phi_n^{(k)}(z_n,\hat{\theta}_n))$ is sufficient for $\{Q_{\theta,n};\theta\in\Theta\}$. (2) For every compact set $K\subset\Theta$,

$$\sup_{\theta\in K}\|P_{\theta,n}-Q_{\theta,n}\|=o(n^{-\frac{k-1}{2}}).$$

The proof is omitted.

4. Tests based on asymptotically sufficient statistics. Let $\omega(\neq\phi)$ be a subset of Θ. Suppose that it is desired to test the null hypothesis that $\theta\in\omega$ against the alternative that $\theta\in\Theta-\omega$. For a statistical test ϕ_n based on $z_n\in\underset{\sim}{X}^{(n)}$ we denote by $\beta_n(\theta;\phi_n)$ the power function of ϕ_n, i.e., $\beta_n(\theta;\phi_n)=E[\phi_n;P_{\theta,n}]$. Let $\tilde{\Phi}(\alpha)$ be the class of all test sequences $\{\phi_n\}$ such that for every compact subset K of ω,

$$\sup_{\theta\in K}|\beta_n(\theta;\phi_n)-\alpha|=o(n^{-\frac{k-1}{2}}).$$

In LeCam[1] such a test sequence, in the case of k=1, is called asymptotically similar of size α uniformly on compacts.

Theorem 2. Suppose that Condition(C_K) is satisfied and that $\{\hat{\theta}_n\}$ is a sequence of estimators belonging to $\underset{\sim}{C}$. Then, for any sequence $\{\phi_n\ ;n=1,2,\dots\}$ of statistical tests contained in $\tilde{\Phi}(\alpha)$ there exists a sequence $\{\psi_n\ ;n=1,2,\dots\}$ of statistical tests contained in $\tilde{\Phi}(\alpha)$ with the following properties: (1) For every compact subset K of $\Theta-\omega$

$$\sup_{\theta\in K}\left|\beta_n(\theta;\phi_n)-\beta_n(\theta;\psi_n)\right| = o(n^{-\frac{K-1}{2}}).$$

The proof is omitted.

5. Estimates based on asymptotically sufficient statistics. For each positive number $s\geq 1$ we denote by $\underset{\sim}{D}_s$ the class of all sequences $\{\tilde{\theta}_n\}$ of estimators of θ satisfying the following properties (1) and (2).

(1) For every $\rho > 0$ and every compact subset K of Θ,

$$\sup_{\theta\in K} P_{\theta,n}\ (|\tilde{\theta}_n(z_n)-\theta|>\rho) = o(n^{-\frac{s-1}{2}})$$

(2) For each $\theta\in\Theta$ there exists a probability measure λ_θ on R^1, which is weakly continuous relative to θ, such that $\lambda_\theta(\{0\})\neq 1$ and that for any compact subset K of Θ the distribution of $n^{\frac{1}{2}}(\tilde{\theta}_n-\theta)$ converges weakly to λ_θ uniformly with respect to θ in K.

For any real number $p(p\geq 1)$ we define $k(p)=p+1$ if p is an integer, $=[p]+2$ if p is not integer where [p] means the maximum integer not exceeding p.

Theorem 3. Let $p(\geq 1)$ be any number. Let $\{\hat{\theta}_n\}\in \underset{\sim}{C}$ and let $t_n^* = (\hat{\theta}_n, \tilde{\Phi}_n^{(1)}(z_n, \hat{\theta}_n), \ldots, \tilde{\Phi}_n^{(k(p))}(z_n, \hat{\theta}_n))$. Suppose that $\Theta = R^1$ and that Condition(C_K) is satisfied with $k=k(p)$. Then, for any $\{\tilde{\theta}_n\}\in \underset{\sim}{D}_1$ there exists a sequence $\{\tilde{\theta}_n^*\}$ of estimates of θ satisfying the following properties (a), (b) and (c); (a). $\{\tilde{\theta}_n^*\}$ is locally uniformly consistent (i.e., for any $\rho > 0$ and for any compact subset K of Θ, $\lim_{n\to\infty} \sup_{\theta\in K} P_{\theta,n}(|\tilde{\theta}_n^*(z_n) - \theta| > \rho) = 0$)

(b). For each $n\in N$ $\tilde{\theta}_n^*$ is a function of t_n^*.

(c). For any compact subset K of Θ

$$\lim_{n\to\infty}\sup \sup_{\theta\in K}\left\{\int_{\underset{\sim}{X}^{(n)}} |\tilde{\theta}_n^*(z_n) - \theta|^p \, dP_{\theta,n} \Big/ \int_{\underset{\sim}{X}^{(n)}} |\tilde{\theta}_n(z_n) - \theta|^p \, dP_{\theta,n}\right\} \leqq 1.$$

The proof is omitted.

Remark 3. In the case where Θ is any open set in R^1, we can conclude the same result as above theorem being exchanged the class $\underset{\sim}{D}_1$ for $\underset{\sim}{D}_{p+1}$.

6. Concluding remarks. A similar result to Theorem 1 is obtained in R.Michel [5]. His result asserts that the order of asymptotic sufficiency of t_n^* is of order $o(n^{-\frac{K-2}{2}})$.

References.

[1] L.LeCam: On the asymptotic theory of estimation and testing hypothesis, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, Univesity of California Press, 1956, Vol.1, pp. 129-156.

[2] J.Pfanzagl: Further results on asymptotic normality I, mimeographed note (1971).

[3] J.Pfanzagl: Asymptotic expansions related to minimum contrast estimators, Ann. Statist. Vol.1, No.6(1973), pp. 993-1026.

[4] J.K.Ghosh and K.Subramanyam: Second order efficiency of maximum likelihood estimators, Sankhya Ser.A, 36(1974), pp. 325-358.

[5] R.Michel: Higher order asymptotic sufficiency, mimeographed note.

Department of Mathematics
School of Science
and Engineering
Waseda University
Shinjuku, Tokyo
Japan