著者別名 | 赤平 昌文
---|---
タイトル | アヤメ学の研究
副題 | 研究
誌名 | 通信学研究
巻 | 27
号 | 1
発行年 | 1976-08
DOI | http://hdl.handle.net/2241/119449
A Remark on Asymptotic Sufficiency of Statistics in Non-Regular Cases*

Masafumi AKAHIRA**

Abstract

Suppose that $X_1, X_2, \ldots, X_n, \ldots$ is a sequence of independent identically distributed random variables with the density $f(x: \theta)$ with a compact support, where $\theta$ is a real valued parameter. We suppose that a strongly $\{c_n\}$-consistent estimator of $\theta$ exists. Then we show that a statistic $(\min_{1 \leq i \leq n} X_i, \max_{1 \leq i \leq n} X_i)$ is asymptotically sufficient in non-regular cases.

1. Introduction

A consistent estimator with order $\{c_n\}$ (or a $\{c_n\}$-consistent estimator) is defined and discussed in Akahira [1], where the necessary conditions for the existence of such an estimator are established and the bounds of the orders of convergence of consistent estimators are obtained for non-regular cases. Further the asymptotic accuracies of $\{c_n\}$-consistent estimators are discussed in Akahira [2].

Asymptotic sufficiency has been discussed under regularity conditions by LeCam [4]. In this paper we extend a similar approach to non-regular cases.

Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent identically distributed random variables with the density $f(x: \theta)$ with a compact support, where $\theta$ is a real valued parameter. We suppose that a strongly $\{c_n\}$-consistent estimator of $\theta$ exists. Then we shall obtain that a statistic $(\min_{1 \leq i \leq n} X_i, \max_{1 \leq i \leq n} X_i)$ is asymptotically sufficient in non-regular cases.

2. Notations and definitions

Let $\mathcal{X}$ be an abstract sample space whose generic point is denoted by $x$, $\mathcal{B}$ a $\sigma$-field of subsets of $\mathcal{X}$ and $\{P_\theta: \theta \in \Theta\}$ a set of probability measures on $\mathcal{B}$, where $\Theta$ is called a parameter space. We suppose that $\Theta$ is an open set in a Euclidean $1$-space $\mathbb{R}^1$. Consider $n$-fold direct products $(\mathcal{X}^{(n)}, \mathcal{B}^{(n)})$ of $(\mathcal{X}, \mathcal{B})$ and the corresponding product measure $P_\theta^{(n)}$ of $P_\theta$. For each $n=1, 2, \ldots$, the points of $\mathcal{X}^{(n)}$ will be denoted by $\bar{x}_n=(x_1, \ldots, x_n)$ and the corresponding random variable by $\bar{X}_n$. An estimator of $\theta$ is defined to be a sequence $\{\hat{\theta}_n\}$ of $\mathcal{B}^{(n)}$-measurable functions $\hat{\theta}_n$ on $\mathcal{X}^{(n)}$ into $\Theta$. For a sequence of positive numbers $\{c_n\}$ ($c_n$ tending to infinity) an estimator $\{\hat{\theta}_n\}$ is called strongly consistent with order $\{c_n\}$ (or strongly $\{c_n\}$-consistent for short) if for every $\varepsilon>0$ and for every compact subset $K$ of $\Theta$, there exists a sufficiently large positive number $L$ satisfying the following:

$$\lim_{n \to \infty} \sup_{\theta \in K} P_\theta^{(n)}(c_n | \hat{\theta}_n - \theta | \geq L) < \varepsilon.$$
A weaker definition of a \([c_n]\)-consistent estimator than that of the above form has been given in Akahira [1].

We suppose that every \(P_\theta(\cdot) (\theta \in \Theta)\) is absolutely continuous with respect to a \(\sigma\)-finite measure \(\mu\). Then we denote the density \(dP_\theta/d\mu\) by \(f(\cdot : \theta)\). If the distribution of \(\bar{x}_n\) is the product measure \(P_\theta^{(n)}\), then the corresponding density with respect to the product measure \(\mu^{(n)}\) will be denoted by \(\prod_{i=1}^n f(x_i : \theta)\). A statistic \(T_n(\bar{X}_n)\) is called asymptotically sufficient if there exist a nonnegative function \(p_n(\bar{x}_n : \theta)\), each the product of a function of \(\bar{x}_n\) only by a function of \(T_n\) and \(\theta\) only such that

\[
\lim_{n \to \infty} \sup_{\theta \in K} \left| \prod_{i=1}^n f(x_i : \theta) - p_n(\bar{x}_n : \theta) \right| d\mu^{(n)} = 0
\]

for any compact subset \(K\) of \(\Theta\) (LeCam [4]).

3. Asymptotically sufficient statistics

Before discussing the asymptotic sufficiency in detail we shall give a definition and a lemma.

Definition. (Generalized from Gnedenko and Kolmogorov [3]) For each \(\theta \in \Theta\) the sums

\[
Y_n(\theta) = X_1(\theta) + X_2(\theta) + \cdots + X_n(\theta)
\]

of positive independent random variables \(X_1(\theta), X_2(\theta), \cdots, X_n(\theta), \cdots\) are said to be uniformly relatively stable for constants \(B_n(\theta)\) if there exist positive constants \(B_1(\theta), B_2(\theta), \cdots, B_n(\theta), \cdots\) such that for any \(\varepsilon > 0\)

\[
P_\theta^{(n)}\left( \left| \frac{Y_n(\theta)}{B_n(\theta)} - 1 \right| > \varepsilon \right) \to 0
\]

as \(n \to \infty\) uniformly in any compact subset of \(\Theta\).

In the subsequent lemma we use the notation that for each \(k\) and each \(\theta \in \Theta\), \(F_{\theta k}(x)\) is the distribution function of \(X_k(\theta)\).

Lemma. (Gnedenko and Kolmogorov [3]).

For each \(\theta \in \Theta\), let \(X_1(\theta), X_2(\theta), \cdots, X_n(\theta), \cdots\) be a sequence of positive independent random variables. The sums

\[
Y_n(\theta) = X_1(\theta) + X_2(\theta) + \cdots + X_n(\theta)
\]

are uniformly relatively stable for constants \(B_n(\theta)\), if there exists a sequence of positive constants \(B_1(\theta), B_2(\theta), \cdots, B_n(\theta), \cdots\) such that for any \(\varepsilon > 0\)

\[
\sum_{k=1}^n \sum_{x \in B_k} dF_{\theta k}(x) \to 0
\]

as \(n \to \infty\) uniformly in any compact subset of \(\Theta\),

\[
\frac{1}{B_n(\theta)} \sum_{k=1}^n \sum_{x \in B_k} x dF_{\theta k}(x) \to 1
\]

as \(n \to \infty\) uniformly in any compact subset of \(\Theta\).

Let \(X = R^1\). Now we suppose that every \(P_\theta(\cdot) (\theta \in \Theta)\) is absolutely continuous with respect to a Lebesgue measure \(m\). Then we denote the density \(dP_\theta/dm\) by \(f(\cdot : \theta)\) and by \(A(\theta) \subseteq X\) the set of points in the space of \(X\) for which \(f(x : \theta) > 0\) and suppose \(f(x : \theta) = f(x - \theta)\). We make the following assumptions (A), (B) and (C).

Assumption (A). \(f(x) > 0\) for \(a \leq x \leq b\); \(f(x) = 0\) for \(x < a, x > b\), and \(f(a)\) and \(f(b)\) are finite.
Assumption (B). \( f(x) \) is twice continuously differentiable in the interval \((a, b)\).

Define

\[
\phi(\theta) = \int_0^\infty w dF(w : \theta),
\]

where \( F(w : \theta) \) is the distribution function of

\[
W(X : \theta) = \chi_{(a, b) \cap A(\theta)}(X) \left| \log \frac{f(x - \theta)}{f'(x)} \right|
\]

(\( \chi_{(a, b) \cap A(\theta)}(\cdot) \) denotes the indicator of \((a, b) \cap A(\theta))\).

Let \( T_n = (Y, Z) \), where \( Y = \min X_i \) and \( Z = \max X_i \). We suppose that \( \{ \hat{\theta}_n(T_n) \} \) is a \( \{ c_n \} \)-consistent estimator. The existence of the estimator is guaranteed (See Theorem 4.1 of [1]). Then for any \( \delta > 0 \) and any compact subset \( K \) of \( \Theta \) there exists a sufficiently large positive number \( L \) satisfying the following:

\[
\limsup_{n \to \infty} \sup_{\theta \in K} P_{\theta, \gamma}(|\hat{\theta}_n(T_n) - \theta| > L c_n^{-1}) < \delta.
\]

Assumption (C). The following (3.2)\( \sim \) (3.4) hold:

\[
\lim_{n \to \infty} n \varphi(L c_n^{-1}) = 0
\]

for any \( \varepsilon > 0 \) and any compact subset \( K \) of \( \Theta \);

\[
\lim_{n \to \infty} \sup_{\theta \in K} \int_{\varphi(L c_n^{-1})}^\infty dF(w : \theta) = 0
\]

for any \( \varepsilon > 0 \) and any compact subset \( K \) of \( \Theta \).

Theorem. Under Assumptions (A), (B) and (C), the statistic \( T_n \), i.e. \( (\min X_i, \max X_i) \), is asymptotically sufficient.

Proof. Let \( \varepsilon \) be an arbitrary positive number. We define \( h(T_n, \theta) \) and \( g(\hat{\theta}_n(T_n)) \) as follows:

\[
h(T_n, \theta) = \chi_{\varphi(y, z)} = \begin{cases} 1, & \text{if } z - b < \theta < y - a; \\ 0, & \text{otherwise} \end{cases}
\]

\[
g(\hat{\theta}_n(T_n)) = \prod_{i=1}^n f(x_i - \hat{\theta}_n(T_n)).
\]

It follows from (3.3), (3.4) and Lemma that \( \sum_{i=1}^n W(X_i : \hat{\theta}_n(T_n) - \theta) \) is uniformly relatively stable for \( n \varphi(L c_n^{-1}) \). Hence we have for any compact subset \( K \) of \( \Theta \)

\[
\liminf_{n \to \infty} P_{\theta, \gamma}(A_n(\hat{\theta}_n(T_n) - \theta : \varepsilon)) = 1,
\]

where \( A_n(\hat{\theta}_n(T_n) - \theta : \varepsilon) = \left\{ x_n : \left| \frac{1}{n \varphi(L c_n^{-1})} \sum_{i=1}^n W(x_i : \hat{\theta}_n(T_n) - \theta) \right| < \varepsilon \right\} \).

It follows from (3.1), (3.2) and (3.5)\( \sim \) (3.7) that for any compact subset \( K \) of \( \Theta \)

\[
\limsup_{n \to \infty} \sup_{\theta \in K} \left( \int_{X_n} \left| \prod_{i=1}^n f(x_i - \theta) - h(T_n, \theta)g(\hat{\theta}_n(T_n)) \right| d\lambda \right)
\]

\[
\leq \limsup_{n \to \infty} \sup_{\theta \in K} \left( \int_{\{ |\hat{\theta}_n(T_n) - \theta| > L c_n^{-1} \}} + \int_{\{ |\hat{\theta}_n(T_n) - \theta| \leq L c_n^{-1} \} \cap A_n(\hat{\theta}_n(T_n) - \theta : \varepsilon)} \right)
\]

\[
\cdot \left| \prod_{i=1}^n f(x_i - \theta) - h(T_n, \theta)g(\hat{\theta}_n(T_n)) \right| d\lambda.
\]
Letting $\delta \to 0$, we complete the proof of the theorem.

**Acknowledgements**

The author wishes to thank Professor K. Takeuchi of Tokyo University for valuable suggestions and Professor T. Homma for his encouragement.

**References**


