Ann. Inst. Statist. Math.
35 (1983), Part A, 229-233

A NOTE ON A CONSISTENT ESTIMATOR OF A MIXING DISTRIBUTION FUNCTION

JŌGI HENNA

(Received Feb. 19, 1981; revised Oct. 28, 1982)

Summary

In this note, we will study a consistent estimator of a mixing distribution function (mixing d.f.). The estimator discussed in this note is that of Choi and Bulgren [4]. Since there is some doubt about the way of proving Lemma in [4] which is used for showing the consistency of the estimator in [2], [3] and [4], we will give different lemmas. We will show that their result (which is still true by using our lemmas) holds under a weaker assumption than theirs. The existence of the estimator is not discussed in [4]. So, we will give conditions under which the existence is guaranteed.

1. Construction of estimator $\hat{G}_n$ and consistency of $\hat{G}_n$

Let $\mathcal{F}=\{F_\theta(x):\ \theta\in R_1\}$ be a family of known d.f.'s on the real line and $G(\theta)$ any d.f. such that $\mu_G(R_1)=1$, where μ_G is the probability measure induced by G and R_1 a compact subset of the real line. Let $F_\theta(x)$ be continuous in x for each θ. We define $P_G(x)$ by

$$P_G(x)=\int_{R_1} F_\theta(x)dG(\theta)\ . \tag{1}$$

It can be easily seen that $P_G(x)$ is a continuous d.f. The problem we are concerned here is to estimate the mixing d.f. G on the basis of the independent random sample $X=(X_1, X_2,\cdots, X_n)$ from the distribution (1). For the mixing d.f. G being estimable, it is obvious that the identifiability condition (which is investigated in [1], [7] and [8]) should be satisfied.

Let $G_n(\theta)$ be any discrete n-point d.f. (with jump g_j at $\theta_j\in R_1$, $j=1, 2,\cdots, n$). The estimator proposed by Choi and Bulgren [4], denoted by $\hat{G}_n(\theta)$, is any $G_n(\theta)$ which minimizes

The research was supported in part by Scientific Research Fund from the Ministry of Education of Japan. No. 564076.

$$(2)\qquad S_n(G_n)=\int \{P_{G_n}(x)-F_n(x)\}^2 dF_n(x)=\frac{1}{n}\sum_{i=1}^{n}\left\{\sum_{j=1}^{m} g_j F_{\theta_j}(X_{(i)})-\frac{i}{n}\right\}^2,$$

where $F_n(x)$ and $X_{(i)}$ are the empirical d.f. and the ith order statistic of X respectively. Assume that $F_\theta(x)$ is continuous in θ for each x, then the existence of $\hat{G}_n$ is guaranteed. In [4], it is assumed that R_1 is an open subset of the real line. But, in this note, we will assume that R_1 is a compact subset of the real line to ensure the existence of $\hat{G}_n$.

We will show the consistency of $\hat{G}_n$ to G under the assumption that $F_\theta(x)$ is continuous in θ for each x and continuous in x for each θ. This is weaker than the assumption (in [4]) that $F_\theta(x)$ is uniformly continuous in (x, θ). We will show first the following two lemmas.

LEMMA 1. *Let $F(x)$ be any continuous d.f. and $H(x)$ any d.f. Let I_1 be the support of μ_F. If there exists x satisfying the inequality*

$$(3)\qquad |H(x)-F(x)|>\delta$$

for some δ (>0), then there exists x in I_1 satisfying (3).

PROOF. Assume that the conclusion does not hold. Then there exists (at least one) x_0 in $I_2=R-I_1$ satisfying (3), where R is the real line. We study first the case $H(x_0)>F(x_0)$. Let $x_1=\sup\{x: F(x)=F(x_0)\}$ and $x_2=\inf\{x: F(x)=H(x_0)-\delta\}$. Then $x_0\leqq x_1<x_2$ and $\{x: x_1<x<x_2\}\cap I_1$ is a non-empty set by the continuity of F. If $x_1<x^*<x_2$, then $F(x^*)<F(x_2)=H(x_0)-\delta$. On the other hand, if $x^*\in I_1$, then $|H(x^*)-F(x^*)|\leqq\delta$. So, we have $H(x^*)\leqq F(x^*)+\delta<H(x_0)$, $(x_0<x^*)$, contradicting the assumption that H is a d.f.

When $H(x_0)<F(x_0)$, we can show a contradiction in the same way as in the first case.

LEMMA 2. *Let $\{H_n(x)\}_{n=1}^{\infty}$ be any sequence of d.f.'s and $F_n(x)$ the empirical d.f. of the sample of size n from any continuous d.f. $F(x)$. If*

$$\int \{H_n(x)-F_n(x)\}^2 dF_n(x)\to 0$$

with probability one as $n\to\infty$, then

$$\|H_n-F_n\|\to 0$$

with probability one, where $\|\ \ \|$ denotes the sup *norm.*

PROOF. Assume that the conclusion does not hold. Then there exists a Borel subset A of the infinite-dimensional Euclidean space R^∞ such that $\mu_F^{(\infty)}(A)>0$ and, if $(X_1, X_2, \cdots)\in A$, then $\sup_x |F_n(x)-F(x)|\to 0$

as $n \to \infty$ (by the Glivenko-Cantelli theorem) and $\sup_x |H_n(x) - F_n(x)| > \delta$ for some δ (>0) and an infinite number of n's, where δ depends on $(X_1, X_2, \cdots)$. Then there exists $x_{0,n}$ holding $|F_n(x_{0,n}) - F(x_{0,n})| < \delta/4$ and $|H_n(x_{0,n}) - F_n(x_{0,n})| > \delta$ for an infinite number of n's. Then

$$|H_n(x_{0,n}) - F(x_{0,n})| > \frac{3}{4}\delta \, . \tag{4}$$

For any fixed $x_{0,n} \in I_1$ satisfying (4) (by Lemma 1), we consider two cases, namely, $H_n(x_{0,n}) > F(x_{0,n})$ and $H_n(x_{0,n}) < F(x_{0,n})$ for an infinite number of n's. We deal with only the first case as the latter case is similar. By the continuity of F, there exists $x_{1,n}$ such that $x_{1,n} = \inf\{x: F(x) - F(x_{0,n}) = \delta/4\}$. Then, for any $x \in (x_{0,n}, x_{1,n}]$, $F(x) \leqq F(x_{1,n}) = F(x_{0,n}) + (1/4)\delta < H_n(x_{0,n}) \leqq H_n(x)$. Then, for $x_{0,n} < x \leqq x_{1,n}$,

$$\begin{aligned}
|H_n(x) - F_n(x)| &\geqq |H_n(x) - F(x)| - |F(x) - F_n(x)| \\
&\geqq H_n(x_{0,n}) - F(x_{1,n}) - \frac{1}{4}\delta \\
&\geqq |H_n(x_{0,n}) - F(x_{0,n})| - |F(x_{0,n}) - F(x_{1,n})| - \frac{1}{4}\delta \\
&> \frac{3}{4}\delta - \frac{1}{4}\delta - \frac{1}{4}\delta = \frac{1}{4}\delta > 0 \, .
\end{aligned}$$

So, we have

$$\int \{H_n(x) - F_n(x)\}^2 dF_n(x) \geqq \left(\frac{1}{4}\delta\right)^2 \int_{(x_{0,n}, x_{1,n}]} dF_n(x) \, .$$

On the other hand, we have

$$\int_{(x_{0,n}, x_{1,n}]} dF_n(x) \to \int_{(x_{0,n}, x_{1,n}]} dF(x) = \frac{1}{4}\delta \, .$$

Accordingly, if $(X_1, X_2, \cdots) \in A$, then

$$\int \{H_n(x) - F_n(x)\}^2 dF_n(x) \geqq \left(\frac{1}{4}\delta\right)^2 \left(\frac{1}{4}\delta - \varepsilon\right) > 0$$

for any fixed ε ($<\delta/4$) and an infinite number of n's. This is contradictory to the assumption that

$$\int \{H_n(x) - F_n(x)\}^2 dF_n(x) \to 0$$

with probability one.

THEOREM. *Assume that $F_\theta(x)$ is continuous in x for each θ and continuous in θ for each x. Then*

$$\mu_{P_G}^{(\infty)}\{\lim_{n\to\infty}\hat{G}_n=G \text{ at every continuity point } \theta \text{ of } G\}=1 .$$

PROOF. For any discrete n-point d.f. G_n^*, we have

$$0\leqq S_n(\hat{G}_n)\leqq S_n(G_n^*)\leqq\int\{P_{G_n^*}(x)-P_G(x)\}^2dF_n(x) + 2\|P_G-F_n\|+\|P_G-F_n\|^2 . \tag{5}$$

Let

$$\theta_0<\theta_1<\cdots<\theta_n , \tag{6}$$

where $\theta_0<\min R_1<\theta_1$, $\theta_{n-1}<\max R_1<\theta_n$ and each θ_i $(i=0, 1, 2,\cdots, n)$ is a continuity point of G. Without loss of generality, assume that $R_1\cap(\theta_{j-1}, \theta_j]\neq\phi$ for each j. Let G_n^* be the d.f. with jump g_j^* at $\theta_j^*\in R_1\cap(\theta_{j-1}, \theta_j]$, where $g_j^*=\mu_G(\theta_{j-1}, \theta_j]$. Then $P_{G_n^*}(x)\to P_G(x)$ uniformly in x if $\delta(\Delta)\to 0$ as $n\to\infty$ by the definition of the Lebesgue-Stieltjes integral and the Polya's theorem (see [5], p. 120), where $\delta(\Delta)=\max_{1\leqq j\leqq n}(\theta_j-\theta_{j-1})$. Hence

$$\int\{P_{G_n^*}(x)-P_G(x)\}^2dF_n(x)\leqq\varepsilon^2 \tag{7}$$

for any given ε (>0). So we have

$$\int\{P_{\hat{G}_n}(x)-F_n(x)\}^2dF_n(x)\to 0$$

with probability one by (5), (7) and the Glivenko-Cantelli theorem. Hence $\|P_{\hat{G}_n}-F_n\|\to 0$ with probability one by Lemma 2 with $H_n(x)=P_{\hat{G}_n}(x)$ and $F(x)=P_G(x)$. Therefore $\|P_{\hat{G}_n}-P_G\|\to 0$ with probability one by $\|P_{\hat{G}_n}-P_G\|\leqq\|P_{\hat{G}_n}-F_n\|+\|F_n-P_G\|$. Accordingly, we have the conclusion by a simple modification of the proof of Theorem 2 of Robbins [6].

Acknowledgement

The author would like to acknowledge the continuing guidance and encouragement of Professor M. Huzii. Also the author is very grateful to Dr. Y. Yajima for his valuable advices and discussions. The author wishes thank the referee for his helpful advices.

RYUKYU UNIVERSITY

REFERENCES

[1] Blum, J. R. and Susarla, V. (1977). Estimation of a mixing distribution function, *Ann. Prob.*, 5, 200-209.

[2] Choi, K. (1969). Estimators for the parameters of a finite mixture of distributions, *Ann. Inst. Statist. Math.*, **21**, 107-116.

[3] Choi, K. (1973). A consistent estimator of the parameters of continuous compound distribution function, *Ann. Inst. Statist. Math.*, **25**, 521-532.

[4] Choi, K. and Bulgren, W. G. (1968). An estimation procedure for mixtures of distributions, *J. R. Statist. Soc.*, B, **30**, 444-460.

[5] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, John Wiley & Sons, New York.

[6] Robbins, H. (1964). The empirical bayes approach to statistical decision problems, *Ann. Math. Statist.*, **35**, 1-20.

[7] Teicher, H. (1960). On the mixture of distributions, *Ann. Math. Statist.*, **31**, 55-73.

[8] Teicher, H. (1963). Identifiability of finite mixtures, *Ann. Math. Statist.*, **34**, 1265-1269.