Youngnam Statistical Letters
June 1992, Vol. 3, No. 1, 25-31

A Note on Complete Convergence in $C_o(R)$ and $L^1(R)$ with Application to Kernel Density Function Estimators

Sungho Lee *

ABSTRACT

Some results relating to $C_o(R)$ and $L^1(R)$ spaces with application to kernel density estimators will be introduced. First, random elements in $C_o(R)$ and $L^1(R)$ are discussed. Then, complete convergence limit theorems are given to show that these results can be used in establishing uniformly consistency and L^1 consistency.

1. Introduction

The estimation of a probability density function $f(t)$ by the kernel method has generated a vast area of research and challenging problems since the pioneering work of Rosenblatt (1956) and Parzen (1962). The measure of deviation(or closedness) of the estimator $\hat{f}_n(t)$ from $f(t)$ has many choices of both local and global nature as well as modes and rates of convergence. For example, the uniform distance, $\sup_t \left|\hat{f}_n(t) - f(t)\right|$, is a random variable which is a global measure of deviation while $\left|\hat{f}_n(t) - f(t)\right|$ is a r.v. which is a measure of deviation at each t. Other commonly used measures of deviation include

$$MSE(\hat{f}_n(t)) = E(\hat{f}_n(t) - f(t))^2,$$
$$IMSE(f_n(t)) = \int E(\hat{f}_n(t) - f(t))^2 \, dt, \quad \text{and}$$
$$\int \left|\hat{f}_n(t) - f(t)\right|^p dt, \qquad 1 \le p < \infty.$$

* Department of Statistics, Taegu University, Kyungsan 713-714, Korea

A general function space expression for the kernel density estimators is

$$S_n = \sum_{k=1}^{n} a_{n_k} X_{n_k},$$

where the array $\{a_{n_k}\}$ consists of weights (possibly r.v.'s) and $\{X_{n_k}\}$ is an array of r.v.'s which take values in a function space determined by the chioce of K and where the norm relates to the measure of deviation of $\hat{f}_n$ to f. Thus, it appears that laws of large numbers for Banach spaces of functions could be used to establish the general asymtotic properties of kernel density estimators. In general, laws of large numbers hold in separable Banach spaces under the assumption of tightness of distribution or geometric conditions on spaces (Radamacher type p spaces). Results relating to tightness are in Taylor and Wei (1979) and Daffer and Taylor (1982). Several authors, most notably Beck (1963), Hoffmann-Jorgensen and Pisier (1976) and Woyczyski (1980), have obtained results relating the laws of large number for sequences of independent r.v.'s to type p spaces, $1 < p \leq 2$.

Notice that $C_o(R) = \{f : f \text{ is continuous and } \lim_{|t| \to \infty} f(t) = 0\}$ is a separable Banach space with the sup norm $\|f\| = \sup_{-\infty < t < \infty} |f(t)|$. Thus the kernel density estimator (cf : Parzen (1962) and Rosenblatt (1971))

$$\hat{f}_n(t) = \frac{1}{nh_n} \sum_{k=1}^{n} K\left(\frac{t - X_k}{h_n}\right) \tag{1.1}$$

is a random variable taking values in $C_o(R)$ if $K \in C_o(R)$ (cf:Taylor and Hu(1986)). Similar results can be obtained in $L^1(R)$ (cf : Lee (1990)).

The form of $\hat{f}_n(t)$ in (1.1) necessitates the consideration of laws of large numbers for arrays of random elements in $C_o(R)$ and $L^1(R)$ which are rowwise independent. However, $C_o(R)$ and $L^1(R)$ is only of type 1 (cf : Lee (1990)). Thus, different techniques other than type p and/or tightness are needed. In this paper, random elements in $C_o(R)$ and $L^1(R)$ will be introduced. Then complete convergence limit theorems are given to show that these results can be used in establishing uniformly consistency and L^1 consistency.

2. Random Elements in $C_o(R)$ and $L^1(R)$

Let $(\Omega,\mathcal{A},P)$ be a probability space and E be a separable Banach space with norm denoted by $\|\ \|$. A random element (r.e.) X in E is a function from Ω into E which is $\mathcal{A}$-measurable with respect to the Borel subsets of E. The expected value of X is defined by the Bochner integral and is denoted by EX. A random variable X is said to be subgaussian (with parameter α) if for some $\alpha > 0$

$$E\left[\exp(tX)\right] \le \exp\left(\frac{t^2\alpha^2}{2}\right) \qquad \forall t \in R.$$

Definitions are directly extended to random elements in $C_o(R)$ and $L^1(R)$ and similar results are obtained (cf : Taylor and Hu (1986), Lee (1990)).

Lemma 2.1. (Taylor (1978)) If E is a separable Banach space, then a function $X : \Omega \to E$ is a random element if and only if $f(X)$ is a random variable for each $f \in E^*$, where E^* is the dual space of E.

Let X be a mapping from $(\Omega,\mathcal{A},P)$ into $C_o(R)$. For each ω in Ω, $X(\omega)$ is an element of $C_o(R)$ whose value at t we denote by $X(t,\omega)$(*or* $X_t(\omega)$). For fixed t, let $X(t)$ denote the real valued function on Ω with value $X(t,\omega)$ at $\omega; X(t)$ is the composition $\Pi_t X$. Similarly, let $(X(t_1),\cdots,X(t_k))$ denote the mapping from Ω into R^k with value $(X(t_1,\omega),\cdots,X(t_k,\omega))$ at ω.

Lemma 2.2. (cf : Billingsley (1968)) Let X be a random element in $C_o(R)$ if and only if for each $t \in R$, $X(t)$ is a random variable.

Similarly, the following characterization of random elements in $L^1(R)$ is indicated in Taylor and Lee (1990).

Lemma 2.3. (cf : Taylor and Lee (1990))

(a) Let X be a function from $R \times \Omega$ into R such that

(i) $\forall t \in R$ $X(t,\cdot) : \omega \to X(t,\omega)$ is a random variable,

(ii) $\forall \omega \in \Omega$ $X(\cdot,\omega) : t \to X(t,\omega)$ is a Riemann integrable function. If $\forall \omega \in \Omega$ $X(\cdot,\omega)$ is identified with $\tilde{X}(\cdot,\omega)$, the equivalence class of $X(\cdot,\omega)$, then $\tilde{X}$ is a random element in $L^1(R)$.

(b) Let $\tilde{X}$ be a random element in $L^1(R)$. Then there exists a function $X : R \times \Omega \to R$ such that

(i) $\forall \omega \in \Omega \;\; X(\cdot, \omega)$ is a Lebesgue integrable function,

(ii) $\forall t \in R \;\; X(t, \cdot)$ is an extended random variable.

Remark. For each random element $\tilde{X}$ in $L^1(R)$, X will denote the function from $R \times \Omega$ into R with properties described by (i) and (ii) of Lemma 2.3(b).

The expected value for a random element in a normed linear space is defined by the Pettis integral (cf : Taylor(1978)). That is, X has expected value $EX \in E$ if $f(EX) = E(f(X)) \;\; \forall f \in E^*$. In a separable Banach space, the Pettis integral is equal to the Bochner integral when the Bochner integral exists. In particular a random element X has a Bochner integral $EX \in E$ if and only if $E\|X\| < \infty$ (cf : Taylor (1978)). Thus, if X is a random element in $C_o(R)$, the expected value of X is defined by the Bochner integral EX and $(EX)(t) = E(X(t))$ (cf : Taylor and Hu). A similar characterization of expected values in $L^1(R)$ is obtained by Taylor and Lee (1990).

3. Complete Convergence

In this section complete convergence results for sums of triangular arrays of random elements in $C_o(R)$ and $L^1(R)$ are compared. These results were designed to establish uniformly consistency and L^1 consistency of the kernel density estimates. Examples are given to show how the results can be applied.

Lemma 3.1. (cf : Taylor and Hu (1986)) Let X be a random element in $C_o(R)$ and let $E|X(u) - X(v)|^\gamma \leq H|u - v|^\delta \;\; \forall u, v \in R$ and $\gamma > 0, \; \delta > 1, \; H > 0$. Then for $\epsilon > 0, \; a \in R$ and $b > 0$

$$P\left[\sup_{a \leq s \leq a+b} |X(s) - X(a)| \geq \epsilon\right] \leq HK_{\delta,\gamma}\frac{b^\delta}{\epsilon^\gamma}$$

where $K_{\delta,\gamma}$ is a constant depending on γ and δ.

Theorem 3.2. (cf : Taylor and Hu (1986)) Let $\{X_{n_k} : 1 \leq k \leq n, \; n \geq 1\}$ be a triangular array of random elements in $C_o(R)$ such that $EX_{n_k} = 0 \;\; \forall n$ and k and let $\frac{1}{2} < d \leq 1$. Suppose that for every n,

(i) $X_{n_k}(t) = 0$ for $|t| > Bn^s$, $s \geq 0$

(ii) $X_{n_1}, \cdots, X_{n_n}$ are independent, subgaussian random elements where the corresponding α_{n_k} satisfy

$$\alpha_n^2(t) = \sum_{k=1}^{n} \alpha_{n_k}^2(t) \leq An^p, \qquad 2d > p$$

(iii) $E\left|X_{n_k}(u) - X_{n_k}(v)\right| \leq Hn^\beta \left|u - v\right|^\delta$, $\beta \geq 0$, $\delta > 1$. Then,
$\left\|\frac{1}{n^d} \sum_{k=1}^{n} X_{n_k}\right\| \to 0$ completely.

Example 3.1. Let $X_1, \cdots, X_n$ be random sample having the same density $f(t)$ belonging to $C_o(R)$. If $E\left|X_1\right|^{2r} < \infty$ for some $r > 0$, then the kernel density estimator $\widehat{f}_n(t) = \frac{1}{nh_n} K\left(\frac{t - X_k}{h_n}\right)$ uniformly converges to $f(t)$ completely, where $K(t)$ is bounded with compact support and satisfies (1.2) and $h_n = O(n^{d-1})$ for some $\frac{1}{2} < d < 1$.

For each n and k, let $X_{n_k}(t) = K\left(\frac{t - X_k}{h_n}\right)$, $Y_{n_k}(t) = K\left(\frac{t - Y_k}{h_n}\right)$ where $Y_k = X_k \mathbf{I}_{[|X_k|^r \leq n]}$. Since $E\left|X_1\right|^{2r} < \infty$, it suffices to show that

$$\sum_{n=1}^{\infty} P\left[\left\|n^{-d} \sum_{k=1}^{n} Y_{n_k} - f\right\| > \epsilon\right] < \infty.$$

By Theorem 3.2, the result follows. $L^1(R)$ counterpart to Theorem 3.2 is given under more strict assumptions as follows.

Theorem 3.3. (cf : Taylor and Lee (1990)) Let $\{X_{n_k} : 1 \leq k \leq n,\ n \geq 1\}$ be a triangular array of random elements in $L^1(R)$ such that $\widetilde{EX}_{n_k} = \tilde{0}$ $\forall n$ and k, and let $\frac{1}{2} < d \leq 1$. Suppose that for every n,

(i) $\tilde{X}_{n_1}, \cdots, \tilde{X}_{n_n}$ are independent, subgaussian random elements where the corresponding functions α_{n_k} satisfy

$$\alpha_n{}^2(t) = \sum_{k=1}^{n} \alpha_{n_k}^2(t) \leq An^p,$$

(ii) $X_{n_k}(t) = 0$ for $|t| > Bn^s$, $0 \le s < d - \frac{p}{2}$,

(iii) $|X_{n_k}(u,\omega) - X_{n_k}(v,\omega)| \le C_n |u-v|^\alpha$ a.s., $0 < \alpha \le 1$, $C_n = O(n^\beta)$, $\beta \ge 0$. Then,

$$\left\| \frac{1}{n^d} \sum_{k=1}^{n} \tilde{X}_{n_k} \right\| \to 0 \text{ completely if } d > s + \tfrac{p}{2}.$$

Example 3.2. Let $X_1, \cdots, X_n$ be a random sample have the same density $f(t)$. If $E|X_1|^r < \infty$ for some $r > 2$, then the L^1-error, $\int |f_n(t) - f(t)|\, dt$, of the kernel density estimator,

$$\widehat{f}_n(t) = \frac{1}{nh_n} \sum_{k=1}^{n} K\left(\frac{t - X_k}{h_n}\right),$$

$$h_n = O\left(n^{d-1}\right), \qquad \frac{1}{2} < d \le 1,$$

converges to zero completely if $2d > 1 + \frac{2}{r}$, where $K(t)$ is bounded with compact support and satisfies (1.2).

For each n and k, let

$$X_{n_k}(t) = K\left(\frac{t - X_k}{h_n}\right) - EK\left(\frac{t - X_k}{h_n}\right),$$

$$Y_{n_k}(t) = K\left(\frac{t - Y_k}{h_n}\right) - EK\left(\frac{t - Y_k}{h_n}\right),$$

where $Y_k = X_k \mathbf{I}_{[|X_k|^r \le n]}$. Then

$$\sum_{n=1}^{\infty} P\left[\int \left|\widehat{f}_n(t) - f(t)\right| dt > \epsilon\right] \le \sum_{n=1}^{\infty} P\left[\left\|n^{-d} \sum X_{n_k}\right\| > \frac{\epsilon}{2}\right] + \sum_{n=1}^{\infty} P\left[\int \left|f(t) - E\widehat{f}_n(t)\right| dt > \frac{\epsilon}{2}\right] < \infty$$

by Theorem 3.3 and Lemma 1 (Devroye, 1983).

References

1. Beck, A. (1963). On the Strong Law of Large Numbers. *Ergodic Theory.* Academic Press, New York, 21-53.
2. Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.
3. Chow, Y. S. (1966). Some Convergence Theorems for Independent Random Variables. *Ann.Math.Statist.*, 35, 1482-1493.
4. Daffer, P. and Taylor, R. (1982). Tightness and Strong Laws of Large Numbers in Banach Spaces. *Bull. Inst. Math. Acad. Sinica*, 10, 251-262.
5. Devroye, L. (1983). The Equivalence of Weak, Strong and Complete Convergence in L^1 for Kernel Density Estimates. *Ann. Statist.*, 11, 896-904.
6. Hoffmann-Jorgensen, J. and Pisier, G. (1976). The Law of Large Numbers and the Central Limit Theorems in Banach Spaces. *Ann. Prob.*, 4, 587-599.
7. Lee, S. (1990). *Laws of Large Numbers for $L^1(R)$.* Ph. D. dissertation, Univ. of Georgia, Athens.
8. Parzen, E. (1962). On Estimation of a Probability Density and Mode. *Ann. Math. Statist.*, 33, 1065-1076.
9. Rosenblatt, M. (1971). Curve Estimates. *Ann. math. Statist.*, 42, 1815-1842.
10. Taylor, R. L. (1978). *Stochastic Convergence of Weighted Sums of Random Elements in Linear Spaces.* Lecture notes in mathematics 672, Springer Verlag, New York.
11. Taylor, R. L., and Hu, T. C. (1987). Consistency of Kernel Density Estimators and Laws of Large Numbers in $C_o(R)$. *Proceeding of sixth Pannonian symposium on mathematical statistics*, D. Reidel publ. comp., dordrecht.
12. Taylor, R. L and Lee, S. (1990). Laws of Large Numbers in $L^1(R)$. Univ. of Georgia Statistics Technical report no. 133.
13. Taylor, R. L. and Wei, D. (1979). Laws of Large Numbers for Tight Random Elements in Normed Linear Spaces. *Ann. Prob.* , 7, 150-155.
14. Woyczynski, W. A. (1980). On Marcinkiewicz-Zygmund Laws of Large Numbers in Banach Space and Related Rates of Convergence. *Prob and math. Statist.* , 1, 117-131.