NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

Uniformly Consistent Estimators of the Characteristic Function

by

Tatiana Lebedeva and Nikolai Ushakov

PREPRINT
STATISTICS NO. 7/2007

NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL http://www.math.ntnu.no/preprint/statistics/2007/S7-2007.ps
Nikolai Ushakov has homepage: http://www.math.ntnu.no/~ushakov
E-mail: ushakov@stat.ntnu.no
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491
Trondheim, Norway.

Uniformly Consistent Estimators of the Characteristic Function

Tatiana Lebedeva
Department of Mathematical Statistics, Moscow State University, Moscow, Russia
and
Nikolai Ushakov
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim, Norway

Abstract

In this work, we suggest some estimators of the characteristic function, which are consistent uniformly over the whole real line under condition that the underlying distribution does not contain a singular component in the Lebesgue decomposition.

Key words: Empirical characteristic function, consistency

1. Introduction

Let $X_1, ..., X_n$ be a random sample (independent and identically distributed random variables) from a distribution function $F(x)$. Denote the characteristic function of X_j and the empirical characteristic function of the sample by $f(t)$ and $f_n(t)$, respectively. The empirical characteristic function is a strongly consistent estimator of the underlying characteristic function. Moreover, the Glivenko-Cantelli theorem implies that $f_n(t)$ is strongly consistent uniformly on each bounded subset. Csörgő and Totik (1983) proved that

$$P\left(\lim_{n\to\infty} \sup_{|t|\le T_n} |f_n(t) - f(t)| = 0\right) = 1$$

if $T_n \to \infty$ so that

$$\lim_{n\to\infty} \frac{\log T_n}{n} = 0.$$

For some classes of underlying distributions this result can be improved, see Ushakov (1999). For example, if $f(t)$ is the characteristic function of a discrete distribution, then $f_n(t)$ is strongly consistent uniformly over the whole real line, see Feuerverger and Mureika (1977). In the general case however, the result of Csörgő and Totik is sharp. In particular, if the underlying distribution is absolutely continuous, then $f_n(t)$ cannot be consistent uniformly on the whole real line. However, in this case, the empirical characteristic function can be modified in such a way that the resulting estimator is uniformly consistent. For instance, the estimator

$$f_n^*(t) = \begin{cases} f_n(t) & \text{for } |t| \le T_n, \\ 0 & \text{for } |t| > T_n, \end{cases}$$

where $T_n \to \infty$ and $\log T_n/n \to 0$ as $n \to \infty$, is strongly consistent uniformly on the whole real line. The defect of this estimator is that its realizations are never characteristic functions. In this work, we show that if the underlying distribution does not contain the singular component in the Lebesgue decomposition, then there exists an estimator of the characteristic function which is strongly consistent uniformly on the whole real line, and whose realizations are always characteristic functions.

If the underlying distribution contains the singular component, situation is unclear. Perhaps uniformly consisten estimator does not exist. This case however is not of practical interest because it is doubtful that such distributions can arise in applications.

2. Absolutely continuous case

In this section, we suppose that $F(x)$, the distribution function of observations X_k, is absolutely continuous. It is well known that in this case $|f(t)| \to 0$ as $|t| \to \infty$. Let $\varphi(t)$ be the characteristic function of an arbitrary absolutely continuous distribution. Consider the following estimator of $f(t)$

$$f_n(t; h) = f_n(t)\varphi(ht)$$

where $h = h_n$ is a positive parameter (depending on n). The following theorem gives necessary and sufficient conditions under which $f_n(t; h)$ is a strongly consistent estimator of $f(t)$ uniformly over the whole real line.

Theorem 1. *$f_n(t;h)$ almost surely converges to $f(t)$ uniformly on the real line:*

$$\sup_t |f_n(t;h) - f(t)| \xrightarrow{a.s.} 0, \quad n \to \infty \tag{1}$$

if and only if

$$h_n \to 0, \quad n \to \infty, \tag{2}$$

and

$$\frac{-\log h_n}{n} \to 0, \quad n \to \infty. \tag{3}$$

Proof. Sufficiency. Consider two arbitrary sequences of positive numbers $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ satisfying conditions:

(i) $a_n h_n \to \infty, \quad n \to \infty,$

(ii) $b_n \to \infty, \quad n \to \infty,$

(iii) $b_n h_n \to 0, \quad n \to \infty.$

These sequences evidently exist. One can put for example $a_n = -h_n^{-1} \log h_n$, and $b_n = -h_n^{-1}(\log h_n)^{-1}$. We have

$$\begin{aligned}
\sup_t |f_n(t;h) - f(t)| &\le \sup_t |f_n(t)\varphi(ht) - f(t)\varphi(ht)| + \sup_t |f(t)\varphi(ht) - f(t)| \\
&= \sup_t (|\varphi(ht)| \cdot |f_n(t) - f(t)|) + \sup_t (|f(t)| \cdot |\varphi(ht) - 1|) \\
&\le \sup_{|t| \le a_n} |f_n(t) - f(t)| + 2 \sup_{|t| > a_n} |\varphi(ht)| + \sup_{|t| \le b_n} |\varphi(ht) - 1| + 2 \sup_{|t| > b_n} |f(t)|.
\end{aligned}$$

Each of the four summands in the right hand side converges to zero as $n \to \infty$: the first summand due to theorem 1 by Csörgő and Totik (1983); the second one because $\varphi(t)$ is the characteristic function of an absolutely continuous distribution and hence $\varphi(t) \to 0$ as $|t| \to \infty$; the third one due to continuity of $\varphi(t)$, condition $\varphi(0) = 1$ and condition (iii); the fourth one because $f(t)$ is the characteristic function of an absolutely continuous distribution (so, $f(t) \to 0$ as $|t| \to \infty$) and due to condition (ii).

Necessity. Prove first necessity of condition (3). Suppose that this condition is not satisfied i.¡e. $\limsup_{n\to\infty} n^{-1} \log(1/h_n) > 0$. We prove that in this case (1) does not hold. Suppose the contrary: (1) holds. Consider a sequence $T_1, T_2, \ldots$ such that $\limsup_{n\to\infty} n^{-1} \log T_n > 0$ and

$$T_n h_n \to 0, \quad n \to \infty. \tag{4}$$

Such a sequence evidently exists, one can put for example $T_n = (h_n \log(1/h_n))^{-1}$. We have

$$\sup_{|t| \le T_n} |f_n(t) - f(t)| \le \sup_{|t| \le T_n} |f_n(t) - f_n(t)\varphi(ht)| + \sup_{|t| \le T_n} |f_n(t)\varphi(ht) - f(t)|. \tag{5}$$

The second summand in the right hand side almost surely converges to 0 as $n \to \infty$ due to our assumption that (1) holds. For the first summand we have

$$\sup_{|t| \le T_n} |f_n(t) - f_n(t)\varphi(ht)| \le \sup_{|t| \le T_n} |\varphi(ht) - 1| = \sup_{|t| \le T_n h_n} |\varphi(t) - 1| \to 0 \text{ as } n \to \infty,$$

due to (4), continuity of $\varphi(t)$ and condition $\varphi(0) = 1$. Thus the right hand side of (5) almost surely converges to 0 as $n \to \infty$, therefore

$$\sup_{|t| \le T_n} |f_n(t) - f(t)| \xrightarrow{a.s.} 0 \text{ as } n \to \infty.$$

But this is in contradiction with theorem 2 by Csörgő and Totik (1983).

Prove now necessity of condition (2). Suppose that (2) does not hold. Then there exist $\varepsilon > 0$ and a subsequence h_{n_k} such that $h_{n_k} > \varepsilon$. Without loss of generality we can assume that

$$h_n > \varepsilon \tag{6}$$

for all n.

Since $\varphi(t)$ is the characteristic function of an absolutely continuous distribution, for any $\Delta > 0$ there exists $\delta > 0$ such that

$$\sup_{|t|\geq\Delta} |\varphi(t)| \leq 1 - \delta. \tag{7}$$

Let us fix an arbitrary $t_0 > 0$ such that $f(t_0) \neq 0$. It follows from (6) and (7) that for some positive δ the inequality $|\varphi(ht_0)| < 1 - \delta$ holds for all n. This inequality implies that

$$|1 - \varphi(ht_0)| > \delta. \tag{8}$$

We have

$$|f_n(t_0; h) - f(t_0)| \geq |f_n(t_0)\varphi(ht_0) - f_n(t_0)| - |f_n(t_0) - f(t_0)| =$$
$$= |f_n(t_0)| \cdot |\varphi(ht_0) - 1| - |f_n(t_0) - f(t_0)|.$$

Taking into account (8) we obtain from this inequality

$$\limsup_{n\to\infty} |f_n(t_0; h) - f(t_0)| \geq \delta \lim_{n\to\infty} |f_n(t_0)| - \lim_{n\to\infty} |f_n(t_0) - f(t_0)| = \delta |f(t_0)| > 0$$

i.e. $f_n(t_0; h)$ does not converge to $f(t_0)$. Thus, necessity of condition (2) is proved. ■

3. Mixtures

In this section, we consider arbitrary mixtures of discrete and absolutely continuous distributions i.e. all distributions without singular component in the Lebesgue decomposition. Distributions, having both the absolutely continuous and discrete components, arise in applications, for example, in financial statistics, see Bowers et al. (1986).

For any k and n ($k \leq n$) define

$$T_{kn} = \sum_{j=1}^{n} I_{(X_k = X_j)} \quad \text{and} \quad N_{kn} = I_{(T_{kn} > 1)}$$

(I_A denotes the indicator of the event A). Let $\varphi(t)$ be the characteristic function of any absolutely continuous distribution, and $h = h_n$ be a sequence of positive numbers satisfying conditions $h_n \to 0$, $-n^{-1} \log h_n \to 0$ as $n \to \infty$. Define

$$\hat{f}_n(t) = \frac{1}{n} \sum_{k=1}^{n} N_{kn} e^{itX_k} + \frac{\varphi(ht)}{n} \sum_{k=1}^{n} (1 - N_{kn}) e^{itX_k}.$$

Theorem 2. *If $F(x)$ does not contain singular component, then*

$$P\left(\lim_{n\to\infty} \sup_{t \in R^1} \left| \hat{f}_n(t) - f(t) \right| = 0 \right) = 1.$$

Proof. Since $F(x)$ does not contain singular component, it is represented in the form

$$F(x) = \alpha F_d(x) + (1-\alpha)F_{ac}(x),$$

where $F_d(x)$ is a purely discrete distribution function, $F_{ac}(x)$ is an absolutely continuous distribution function, and $0 \le \alpha \le 1$. Without loss of generality we can suppose that for all $k = 1, 2, ...,$

$$X_k = \nu_k Y_k + (1-\nu_k)Z_k,$$

where, for each fixed k, random variables Y_k and Z_k have distribution functions $F_d(x)$ and $F_{ac}(x)$ respectively, they (these two random variables) do not depend on ν_k, and $P(\nu_k = 1) = \alpha$, $P(\nu_k = 0) = 1-\alpha$

Denote characteristic functions, corresponding to distribution functions $F_d(x)$ and $F_{ac}(x)$ by $f_d(t)$ and f_{ac}, respectively. Then $f(t) = \alpha f_d(t) + (1-\alpha)f_{ac}(t)$. First let us prove that for each fixed $k = 1, 2, ...,$

$$N_{kn} \xrightarrow{a.s.} \nu_k \tag{9}$$

as $n \to \infty$. Note that the sequence N_{kn}, $n = k, k+1, ...,$ increases. Indeed,

$$\sum_{j=1}^{n} I_{(X_k = X_j)} \le \sum_{j=1}^{n+1} I_{(X_k = X_j)},$$

therefore $\{T_{kn} > 1\} \subseteq \{T_{k(n+1)} > 1\}$ i.e. $N_{kn} \le N_{k(n+1)}$. Hence there exists a random variable N_k such that $N_{kn} \uparrow N_k, \quad n \to \infty$. Prove that $N_k = \nu_k$ (a.s.). N_k takes two values, 0 and 1 (because each N_{kn} takes only these two values). Show that

$$P(N_k = 1|\nu_k = 1) = 1 \tag{10}$$

and

$$P(N_k = 1|\nu_k = 0) = 0. \tag{11}$$

Let $x_1, x_2, ...$ be values which Y_k takes with positive probabilities:

$$P(Y_k = x_m) = p_m > 0, \quad m = 1, 2, ..., \quad \sum_{m=1}^{\infty} p_m = 1.$$

Then for each $j = 1, 2, ...$ and each $m = 1, 2, ...,$

$$P(X_j = x_m) = P(X_j = x_m|\nu_j = 1)P(\nu_j = 1) + P(X_j = x_m|\nu_j = 0)P(\nu_j = 0)$$

$$\ge P(X_j = x_m|\nu_j = 1)P(\nu_j = 1) = P(Y_j = x_m)P(\nu_j = 1) = \alpha p_m > 0.$$

We have

$$P(N_k = 1|\nu_k = 1) = \lim_{n\to\infty} P(N_{kn} = 1|\nu_k = 1) = \lim_{n\to\infty} P\left(\bigcup_{j\ne k}^{n}\{X_k = X_j\}|\nu_k = 1\right)$$

$$= P\left(\bigcup_{j\ne k}^{\infty}\{X_k = X_j\}|\nu_k = 1\right) = P\left(\bigcup_{j\ne k}^{\infty}\{Y_k = X_j\}|\nu_k = 1\right)$$

$$= P\left(\bigcup_{j\neq k}^{\infty}\{Y_k = X_j\}\right) = \sum_{m=1}^{\infty} P\left(\bigcup_{j\neq k}^{\infty}\{Y_k = x_m, X_j = x_m\}\right)$$

$$= \sum_{m=1}^{\infty} P\left(\{Y_k = x_m\} \cap \bigcup_{j\neq k}^{\infty}\{X_j = x_m\}\right) = \sum_{m=1}^{\infty} P(Y_k = x_m) = 1$$

because for every m

$$P\left(\bigcup_{j\neq k}^{\infty}\{X_j = x_m\}\right) = 1 - P\left(\bigcap_{j\neq k}^{\infty}\{X_j \neq x_m\}\right) = 1 - \prod_{j\neq k}^{\infty} P(X_j \neq x_m)$$

$$= 1 - \prod_{j\neq k}^{\infty}(1 - P(X_j = x_m)) \geq 1 - \prod_{j\neq k}^{\infty}(1 - \alpha p_m) = 1.$$

Thus (10) is proved. Prove (11). We have

$$P(N_k = 1|\nu_k = 0) = \lim_{n\to\infty} P(N_{kn} = 1|\nu_k = 0) = \lim_{n\to\infty} P\left(\bigcup_{j\neq k}^{n}\{X_k = X_j\}|\nu_k = 0\right)$$

$$= P\left(\bigcup_{j\neq k}^{\infty}\{X_k = X_j\}|\nu_k = 0\right) = P\left(\bigcup_{j\neq k}^{\infty}\{Z_k = X_j\}\right) \leq \sum_{j\neq k}^{\infty} P(Z_k = X_j) = 0$$

because for any j $P(Z_k = X_j) = P(Z_k - X_j = 0) = 0$ since $Z_k - X_j$ is a continuous random variable (Z_k is continuous and Z_k and X_j are independent). So, (11) is also proved. (10) and (11) imply that $N_k = \nu_k$ (a.s.).

Now we have

$$|\hat{f}_n(t) - f(t)|$$

$$\leq \left|\frac{1}{n}\sum_{k=1}^{n} N_{kn}e^{itX_k} - \alpha f_d(t)\right| + \left|\frac{\varphi(ht)}{n}\sum_{k=1}^{n}(1 - N_{kn})e^{itX_k} - (1-\alpha)f_{ac}(t)\right|$$

$$\leq \left|\frac{1}{n}\sum_{k=1}^{n} N_{kn}e^{itX_k} - \frac{1}{n}\sum_{k=1}^{n}\nu_k e^{itX_k}\right| + \left|\frac{1}{n}\sum_{k=1}^{n}\nu_k e^{itX_k} - \alpha f_d(t)\right|$$

$$+ \left|\frac{\varphi(ht)}{n}\sum_{k=1}^{n}(1 - N_{kn})e^{itX_k} - \frac{\varphi(ht)}{n}\sum_{k=1}^{n}(1-\nu_k)e^{itX_k}\right|$$

$$+ \left|\frac{\varphi(ht)}{n}\sum_{k=1}^{n}(1 - \nu_k)e^{itX_k} - (1-\alpha)f_{ac}(t)\right|$$

$$\leq \frac{2}{n}\sum_{k=1}^{n}|N_{kn} - \nu_k| + \left|\frac{1}{n}\sum_{k=1}^{n}\nu_k e^{itX_k} - \alpha f_d(t)\right| + \left|\frac{\varphi(ht)}{n}\sum_{k=1}^{n}(1-\nu_k)e^{itX_k} - (1-\alpha)f_{ac}(t)\right|$$

$$\leq \frac{2}{n}\sum_{k=1}^{n}|N_{kn} - \nu_k| + \left|\frac{1}{n}\sum_{k=1}^{n}\nu_k e^{itY_k} - \alpha f_d(t)\right| + \left|\frac{\varphi(ht)}{n}\sum_{k=1}^{n}(1-\nu_k)e^{itZ_k} - (1-\alpha)f_{ac}(t)\right|$$

$$\leq \frac{2}{n}\sum_{k=1}^{n}|N_{kn} - \nu_k| + \left|\frac{1}{n}\sum_{k=1}^{n}(\nu_k - \alpha)e^{itY_k}\right| + |\alpha| \cdot \left|\frac{1}{n}\sum_{k=1}^{n} e^{itY_k} - f_d(t)\right|$$

$$+\left|\frac{1}{n}\sum_{k=1}^{n}(\nu_k-\alpha)e^{itZ_k}\right|+|1-\alpha|\cdot\left|\frac{\varphi(ht)}{n}\sum_{k=1}^{n}e^{itZ_k}-f_{ac}(t)\right|.$$

Each of the five summands in the right hand side uniformly converges to zero as $n\to\infty$. The first one due to (9). For the second and fourth summands, this is proved using standard technique, see, for example, proof of the Glivenko-Cantelli theorem. The third one due to the uniform consistency of the empirical characteristic function in the discrete case. The fifth one due to Theorem 1. ■

References

Bowers, N.L., Hickman, J.C., Gerber, H.U., Jones, D.A., and Nesbitt, C.J., 1997. Actuarial Mathematics. 2nd ed. The Society of Actuaries.

Csörgő, S. and Totik, V., 1983. On how long interval is the empirical characteristic function uniformly consistent? Acta Sci. Math., 45, 141-149.

Feuerverger, A. and Mureika, R.A., 1977. The empirical characteristic function and its applications. Ann. Statist., 5, 88-97.

Ushakov, N.G., 1999. Selected Topics in Characteristic Functions. VSP, Utrecht.