

WEAK CONVERGENCE OF THE EMPIRICAL CHARACTERISTIC FUNCTION

J. E. YUKICH

ABSTRACT. Let P be a probability measure on \mathbf{R} such that the density $f(x)$ for P exists and there exists $x_0 > 0$ such that $f(x) + f(-x)$ is decreasing for all $|x| \geq x_0$. Let $c(t)$ be the characteristic function for P , $c_n(t)$ the empirical characteristic function, and let $C_n(t) := n^{1/2}(c_n(t) - c(t))$. New necessary and sufficient metric entropy conditions are obtained for the weak convergence of $C_n(t)$ on the space of continuous complex valued functions on $[-\frac{1}{2}, \frac{1}{2}]$. The result is used to characterize the weak convergence of $C_n(t)$ in terms of the tail behavior of P and it also provides the first step towards a generalization of the Borisov-Dudley-Durst theorem. It also provides a partial response to a challenge raised by Dudley.

Let X be a real valued random variable with distribution function $F(x) = P(X \leq x)$ and characteristic function $c(t) := \int e^{itx} dF(x)$. Define the n th empirical distribution function $F_n(x) := n^{-1} \sum_{k=1}^n I_{[X_k, \infty)}(x)$ where I_A is the characteristic function of the set A and X_1, X_2, \dots are independent and identically distributed copies of X . Let $c_n(t) := \int e^{itx} dF_n(x)$ be the n th empirical characteristic function and let

$$C_n(t) := n^{1/2}(c_n(t) - c(t))$$

be the normalized empirical characteristic function. Let $C([-\frac{1}{2}, \frac{1}{2}])$ be the space of continuous, complex valued functions on $[-\frac{1}{2}, \frac{1}{2}]$, with the usual sup norm ($\|\cdot\|_\infty := \sup_{t \in [-\frac{1}{2}, \frac{1}{2}]} |\cdot|$). As in [2 and 5], we say that $C_n(t)$ converges weakly on $C([-\frac{1}{2}, \frac{1}{2}])$ if the measures induced by $C_n(t)$ on $C([-\frac{1}{2}, \frac{1}{2}])$ converge weakly. Recall that $C_n(t)$ converges weakly to a Gaussian process iff $\mathcal{G} := \{x \rightarrow e^{itx}, |t| \leq 1/2\}$ is a functional Donsker class for P (see [2, 3, and 5] for details).

Before presenting the main results, we recall two definitions of metric entropy.

DEFINITION. Given a measurable space (A, \mathcal{A}) , let $\mathcal{L}^0(A, \mathcal{A})$ denote the set of all real-valued \mathcal{A} -measurable functions on A . Given $f, g \in \mathcal{L}^0(A, \mathcal{A})$ let $[f, g] := \{h \in \mathcal{L}^0(A, \mathcal{A}) : f \leq h \leq g\}$ (empty unless $f \leq g$). Given a probability space (A, \mathcal{A}, P) , a class of functions $\mathcal{F} \subset \mathcal{L}^1(A, \mathcal{A}, P)$, and $\varepsilon > 0$, let $N^B(\varepsilon, \mathcal{F}, P)$ denote the smallest m such that for some f_1, \dots, f_m in $\mathcal{L}^1(A, \mathcal{A}, P)$

$$(1) \quad \mathcal{F} \subset \bigcup_{i,j} \left\{ [f_i, f_j] : \int |f_j - f_i| dP \leq \varepsilon \right\}.$$

Here a set $[f_i, f_j]$ is called a bracket and $\log N^B(\varepsilon, \mathcal{F}, P)$ is called metric entropy with bracketing [3].

Received by the editors December 13, 1984.

1980 *Mathematics Subject Classification.* Primary 60B10, 60F05.

Key words and phrases. Empirical characteristic function, central limit theorem on $C([-\frac{1}{2}, \frac{1}{2}])$.

©1985 American Mathematical Society

0002-9939/85 \$1.00 + \$.25 per page

In the case that \mathcal{F} consists of complex valued functions, let $N^B(\varepsilon, \mathcal{F}, P)$ denote the smallest m such that (1) holds with \mathcal{F} replaced by $\{\operatorname{Re}(f): f \in \mathcal{F}\} \cup \{\operatorname{Im}(f): f \in \mathcal{F}\}$.

DEFINITION. Given a probability space (A, \mathcal{A}, P) , $\mathcal{F} \subset \mathcal{L}^2(A, \mathcal{A}, P)$, and $\varepsilon > 0$, let $N(\varepsilon, \mathcal{F}, P) := \inf\{m: \exists f_1, \dots, f_m \text{ in } \mathcal{L}^2(A, \mathcal{A}, P) \text{ such that } \forall f \in \mathcal{F} \exists 1 \leq i \leq m \text{ with } (\int |f - f_i|^2 dP)^{1/2} \leq \varepsilon\}$.

Say that the probability measure P satisfies condition (A) if

- (A) P has a density $f(x)$ and there exists $x_0 > 0$ such that $f(x) + f(-x)$ is decreasing for all $|x| \geq x_0$.

Without loss of generality we will assume that $x_0 \geq 3$.

Throughout, let $\mathcal{G} := \{x \rightarrow e^{itx}: |t| \leq 1/2\}$. We now present the main results.

As noted in [3], Marcus [5] proved for any probability measure P on \mathbf{R} that the following are equivalent:

- (i) $C_n(t)$ converges weakly to a Gaussian process, and
- (ii) $\int_0^1 (\log N(\varepsilon, \mathcal{G}, P))^{1/2} d\varepsilon < \infty$.

Here, the main new result was to show that (ii) implies (i). Since then, Dudley and others have asked (cf. the “challenge” in Chapter 11 of [3]) whether this implication follows from a more general result in empirical processes. Under condition (A), we provide a solution to this “challenge” and give a surprisingly short and simple proof of (ii) \Rightarrow (i).

We actually show more. In fact, under condition (A) we have

THEOREM 1. $C_n(t)$ converges weakly to a Gaussian process iff

$$(2) \quad \int_0^1 (\log N^B(\varepsilon^2, \mathcal{G}, P))^{1/2} d\varepsilon < \infty.$$

Apart from very small classes of functions which satisfy both (2) and the P -functional Donsker class property for all P , the only known class of functions for which (2) is a necessary and sufficient condition (NASC) for the P -functional Donsker class property is the class $2^{\mathbf{N}}$ of all subsets of \mathbf{N} . This is the Borisov-Dudley-Durst theorem, cf. Theorem 6.3.1 of [3] and also [1 and 4]. Theorem 1 provides a second class of functions for which (2) is a NASC for the P -functional Donsker class property. A complete characterization of those \mathcal{G} for which (2) is a NASC for the P -functional Donsker class property still remains an open problem.

Additionally, we characterize the weak convergence of $C_n(t)$ in terms of the tail behavior of P . Given a probability measure P on \mathbf{R} and $\varepsilon > 0$, let

$$M := M(\varepsilon) := \inf\{M: M \geq 1 \text{ and } P(|X| \geq M) \leq \varepsilon^2\}.$$

THEOREM 2. Let P be a probability measure satisfying (A). Then \mathcal{G} is a functional Donsker class for P iff

$$\int_0^1 (\log M(\varepsilon))^{1/2} d\varepsilon < \infty.$$

The proofs of the above theorems center upon a crucial inequality, expressed in the following

LEMMA. *Let P be a probability measure satisfying (A). Let $\varepsilon > 0$ and $M(\varepsilon)$ as above. Then with x_0 as in condition (A), we have*

$$(3) \quad 4\pi x_0 N(\varepsilon, \mathcal{G}, P) \geq M(\varepsilon) \geq \varepsilon^2 N^B(3\varepsilon^2, \mathcal{G}, P)/16, \quad 0 < \varepsilon \leq 1/2.$$

The “only if” implication of Theorem 1 now follows from (ii) and (3); the “if” step follows from a trivial modification of Theorem 6.2.1 of [3] to classes of complex valued functions. Theorem 2 is an immediate consequence of (ii), (2), and (3). Thus it remains to prove (3).

Actually, as we shall see below, the second inequality in (3) is true without condition (A) and therefore the same is true for the “if” part of Theorem 2.

PROOF OF LEMMA. We first prove the first inequality. Let $\varepsilon > 0$ and $N := N(\varepsilon, \mathcal{G}, P)$. Find $t_j \in [-\frac{1}{2}, \frac{1}{2}]$, $j = 1, 2, \dots, N$, $t_j < t_{j+1}$, as in the definition of N . Then for any e^{itx} , $t \in [-\frac{1}{2}, \frac{1}{2}]$, there is a $t_j := t_j(t)$, $1 \leq j \leq N$, with

$$\int |e^{itx} - e^{it_j x}|^2 dP(x) = 4 \int \sin^2(|t - t_j|x/2) dP(x) \leq \varepsilon^2.$$

Let $\alpha := \alpha(t) := \pi/|t - t_j| \geq 1$. With x_0 as in condition (A), we may clearly find an odd integer $x_0 \leq k \leq 2x_0$, k independent of t , such that $f(x) + f(-x)$ is decreasing for all $|x| \geq \alpha k$. Then

$$\varepsilon^2 \geq 4 \int_{|x| \geq \alpha k} \sin^2(|t - t_j|x/2) dP(x) \geq \int_{|x| \geq \alpha k} dP(x),$$

since $f(x) + f(-x)$ is decreasing for $|x| \geq \alpha k$ and since $\sin(|t - t_j|x/2) = 1$ when $x = \alpha k$.

By minimality of $M(\varepsilon)$, we have $2\alpha x_0 \geq \alpha k \geq M(\varepsilon)$ and thus $|t - t_j| \leq 2\pi x_0/M(\varepsilon)$. Since this last inequality holds for all t we clearly have

$$\sup_{1 \leq j \leq N-1} (t_{j+1} - t_j) \leq 4x_0\pi/M(\varepsilon),$$

whence the desired result.

The proof of the second inequality centers around a suitable construction of brackets, similar to that used in [6]. Partition $[-\frac{1}{2}, \frac{1}{2}]$ into $m := [2M(\varepsilon)/\varepsilon^2] + 1$ subintervals of equal length less than or equal to $\varepsilon^2/2M(\varepsilon)$ and with endpoints t_0, t_1, \dots, t_m ; $t_0 = -\frac{1}{2}$ and $t_m = \frac{1}{2}$. (Here $[\cdot]$ denotes the integer part.) For all $j = 1, \dots, m$, define the bracket functions

$$g_j^\pm(x) = (\pm \varepsilon^2/2 + \operatorname{Re} e^{it_j x})1_{[-M, M]}(x) \pm 1_{[-M, M]^c}(x),$$

and

$$h_j^\pm(x) = (\pm \varepsilon^2/2 + \operatorname{Im} e^{it_j x})1_{[-M, M]}(x) \pm 1_{[-M, M]^c}(x).$$

Given e^{itx} find the smallest t_j such that $t \leq t_j$. We claim that $\forall x \in \mathbf{R}$

$$(4) \quad g_j^-(x) \leq \operatorname{Re} e^{itx} \leq g_j^+(x),$$

$$(5) \quad h_j^-(x) \leq \operatorname{Im} e^{itx} \leq h_j^+(x), \text{ and}$$

$$(6) \quad \int |g_j^+ - g_j^-| dP \leq 3\varepsilon^2 \quad \text{and} \quad \int |h_j^+ - h_j^-| dP \leq 3\varepsilon^2,$$

showing that $N^B(3\varepsilon^2, \mathcal{G}, P) \leq 4m$, and thus implying the desired result.

To prove (4) it suffices to show that

$$|\cos(t_j x) - \cos(tx)| \leq \varepsilon^2/2 \quad \forall x \in [-M, M].$$

Now $|\cos(t_j x) - \cos(tx)| \leq 2|\sin(t_j - t)x/2| \leq 2(t_j - t)|x|/2 \leq \varepsilon^2/2$ for $|x| \leq M$. Clearly (5) is proved in a similar fashion. Finally, by considering integrals over $[-M, M]$ and $[-M, M]^c$, we obtain (6) as well. Q.E.D.

ACKNOWLEDGMENT. It is my pleasure to thank the referee for several helpful suggestions which resulted in an improved version of the original draft.

NOTE ADDED IN PROOF. After the completion of this paper Professor E. Giné kindly showed the author a second response to Dudley's challenge, one based on Theorem 5.4 of *Some limit theorems for empirical processes* (E. Giné and J. Zinn, Ann. Probab. **12** (1984)); if one moves expectations inside random entropy integrals this approach will handle the general case.

REFERENCES

1. I. S. Borisov, *Problem of accuracy of approximation in the central limit theorem for empirical measures*, Sibirsk. Mat. Zh. **24**, no. 6(1983), 14–25 = Siberian Math. J. July issue (1984), 833–843.
2. S. Csörgö, *Limit behavior of the empirical characteristic function*, Ann. Probab. **9** (1981), 130–144.
3. R. M. Dudley, *A course on empirical processes*, Lecture Notes in Math., vol. 1097, Springer, Berlin and New York.
4. M. Durst and R. M. Dudley, *Empirical processes, Vapnik-Červonenkis classes, and Poisson processes*, Prob. Math. Statist. (Wrocław) **1** (1981), 109–115.
5. M. Marcus, *Weak convergence of the empirical characteristic function*, Ann. Probab. **9** (1981), 194–201.
6. J. E. Yukich, *Laws of large numbers for classes of functions*, Multivariate Anal. (to appear).

UNIVERSITÉ LOUIS PASTEUR, 7 RUE RENÉ DESCARTES, 67084 STRASBOURG, FRANCE

Current address: Department of Mathematics, Lehigh University, Bethlehem, Pennsylvania 18015