Lévy-like Continuity Theorems for Convergence in Distribution

by
Theodore P. Hill[1]
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332
and
Ulrich Krengel
Institut für Mathematische Stochastik
Universität Göttingen
Lotzestrasse 13
Göttingen, Germany 37083

Abstract

Lévy's classical continuity theorem states that if the pointwise limit of a sequence of characteristic functions exists, then the limit function itself is a characteristic function if and only if the limit function satisfies a single universal limit condition (in his case, the limit at zero is one), in which case the underlying measures converge weakly to the probability measure represented by the limit function. It is the purpose of this article to give a number of direct analogs of Lévy's theorem for other probability-representing functions including moment sequences, maximal moment sequences, mean-residual-life functions, Hardy-Littlewood maximal functions, and failure-rate functions. In each of these cases the single crucial condition on the limit function often relates to conservation of mass or moment, but a general theory encompassing all of these examples is still missing.

§1 Introduction

In general, limits of characteristic functions of probability measures are not characteristic functions, limits of moment sequences are not moment sequences, limits of densities are not densities, and limits of many maximal functions such as failure rate functions or Hardy-Littlewood functions are themselves not failure rate functions or Hardy-Littlewood functions. On the other hand, for each of these examples and a number of analogous results, the only additional requirement for the limit function to itself be a member of the

[1] Research partially supported by NSF grant DMS-9971146 and the Göttingen Academy of Sciences.

representing class of functions is that the limit function satisfy one universal limit condition. With Lévy's continuity theorem as prototype, the purpose of this note is to establish a number of similar continuity theorems for weak convergence of probability measures.

The underlying intuition is that many of the most basic defining properties of probability-representing functions are preserved by limits. For example, the limit of convex functions is convex, the limit of measurable functions is measurable, the limit of non-negative definite functions is non-negative definite, the limit of non-decreasing functions is non-decreasing, the limit of d-periodic functions is d-periodic, and the limit of functions with values in $[a, b]$ has values in $[a, b]$. *Continuity*, however, is in general *not* preserved under limits, and therefore if continuity is necessary (in addition to non-negative definiteness, convexity, etc.) for a function to be an element of a representing class, then that continuity condition is often the only additional property the limit function must satisfy in order for it to be in the class.

The crucial continuity conditions in the theorems below often relate to conservation of mass or moment (tightness or uniform integrability), but a general theory containing all these results is still missing.

§2 General Probability Distributions

Throughout this note: $\mathbb{N}$ denotes the natural numbers; $\mathbb{Z}$ the integers; $\mathbb{R}$ and $\mathbb{R}^k$ real and k-dimensional Euclidean space, respectively; $\mathbb{C}$ the complex plane; $\mathbb{B}$ and $\mathbb{B}^k$ the real and k-dimensional Borel subsets of $\mathbb{R}$ and $\mathbb{R}^k$; $\mathbb{R}^+$ the non-negative reals $[0, \infty)$; $\delta_{(a)}$ the Dirac (point) mass at a; $a \wedge b$ the minimum of a and b; $E(X)$ the expected value of the random variable X; and $\mathcal{L}(X)$ the distribution (law) of the random variable X.

As a first example of Lévy-like continuity theorems, consider the well-known generalization of the classical one-dimensional case of Lévy's theorem to the general finite dimensional setting. (Recall that the characteristic function of an $\mathbb{R}^k$-valued random vector X is given by $\psi_x(\vec{v}) = E(e^{i\langle w, X\rangle})$, where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product.)

Theorem 2.1. *Suppose $(P_n)_{n=1}^{\infty}$ are Borel probability measures on $\mathbb{R}^k$ with characteristic functions (ψ_k) respectively. If $\lim\limits_{n\to\infty} \psi_k = g$, then there is a Borel probability measure P on $\mathbb{R}^k$ with characteristic function g if and only if $\lim\limits_{\vec{x}\to\vec{0}} g(\vec{x}) = \vec{1}$, in which case $P_n \xrightarrow{w} P$.*

Proof. (See Theorem 18.21 of Fristedt and Gray (1997). Note also that the proof that distinct Borel probability measures on $\mathbb{R}^k$ have distinct characteristic functions uses the Parseval relation when $k > 1$.) $\square$

Not all classical probability-representing functions share this property that all pointwise limits of such functions are themselves probability-representing functions of that same type if and only if they satisfy a single universal limit condition. For instance, consider the case of classical empirical distribution functions.

Example 2.2. Let $(F_n)_{n=1}^{\infty}$ and $(G_n)_{n=1}^{\infty}$ be the cdf's for the Dirac point masses $(\delta_{(\frac{1}{n})})$ and $(\delta_{(\frac{-1}{n})})$ respectively. Then $\lim_{n\to\infty} F_n = g_1$ and $\lim_{n\to\infty} G_n = g_2$ both exist, but only g_2 is a cdf, since g_1 is not right continuous at zero, even though $\delta_{(\frac{1}{n})}$ and $\delta_{(\frac{-1}{n})}$ both converge weakly to $\delta_{(0)}$.

The problem in Example 2.2, however, is simply the artificial convention of right continuity in the classical definition of cdf, and that is easy to repair in the following way.

Definition 2.3. For a non-decreasing function $g : \mathbb{R} \to [0, 1]$, let $\hat{g}$ denote the equivalence class of all non-decreasing functions from $\mathbb{R}$ to $[0, 1]$ which agree with g at all continuity points of g. (In Example 2.2 above, for instance, $g_1 \in \hat{g}_2$.) Conversely, given such an equivalence class $\hat{g}$, let g denote the unique right continuous element in $\hat{g}$. Also, let $\lim \hat{g}_n$ denote the equivalence class of $\lim g_n$, if it exists.

Clearly there is a one-to-one correspondence between Borel probability measures and equivalence classes of cdf's, and this correspondence leads to the following very simple Lévy-like theorem.

Theorem 2.4. *Suppose $(F_n)_{n=1}^{\infty}$ are the cdf's for Borel probability measures $(P_n)_{n=1}^{\infty}$. If $\lim_{n\to\infty} \hat{F}_n = \hat{g}$, then there is a Borel probability measure P with generalized cdf $\hat{g}$ if and only if $\lim_{x\to\infty} g(x) - g(-x) = 1$, in which case $P_n \xrightarrow{w} P$.*

Proof. Clearly g is non-decreasing with values in $[0,1]$, each F_n is. By definition, g is right continuous, so g is the cdf for a Borel probability measure iff $\lim_{x\to-\infty} g(x) = 0$ and $\lim_{x\to\infty} g(x) = 1$ iff $\lim_{x\to\infty} g(x) - g(-x) = 1$, in which case $P_n \xrightarrow{w} P$ since $F_n \to g$ at all continuity points of g. $\square$

An analogous result for half-space representing functions can be formulated using essentially the same arguments, since a Borel probability measure on $\mathbb{R}^k$ is determined by its measures of half-spaces (e.g., Feller (1971)). That is, P is uniquely determined by the functions $(\varphi_P(\vec{a}, b))_{\vec{a}\in\mathbb{R}^k, b\in\mathbb{R}}$, where $\varphi_P(\vec{a}, b) = P(H^+(\vec{a}, b))$, and $H^+(\vec{a}, b) = \{\vec{x} \in \mathbb{R}^k : \sum a_i x_i > b\}$. The corresponding result for this setting, roughly speaking, says that if $\varphi_{P_n}(\vec{a}, b)$ converges for all a, b as $n \to \infty$, then the limiting functions $\{g(\vec{a}, b)\}$ correspond to a Borel probability (and $P_n \xrightarrow{w} P$) if and only if $\lim_{b\to\infty} g(\vec{a}, b) - g(\vec{a}, -b) = 1$ for all $\vec{a} \in \mathbb{R}^k$.

The final result in this section is a Lévy-like continuity theorem for completely general Borel probabilities (e.g., no moment or continuity conditions) supported on the positive reals.

Definition 2.5. For a Borel probability measure P with support in $\mathbb{R}^+$, let $\psi_P : \mathbb{R}^+ \to \mathbb{R}^+$ be given by

$$\psi_P(t) = E(X \wedge t),$$

(where $\mathcal{L}(X) = P$).

Theorem 2.6. *Suppose $(P_n)_{n=1}^{\infty}$ are Borel probability measures with support in $\mathbb{R}^+$. If $\lim_{n\to\infty} \psi_{P_n} = g$, then there exists a Borel probability P with $\psi_P = g$ if and only if $\lim_{t\to\infty} \psi(t+1) - \psi(t) = 0$, in which case $P_n \xrightarrow{w} P$.*

The proof will be based on the following lemma, which is an analog of a result of Gilat (1987) for the maximal function $E(X \vee t)$ and distributions with finite mean.

Lemma 2.7. *There is a one-to-one correspondence between Borel probability measures on $\mathbb{R}^+$ and the set of all functions $\psi : \mathbb{R} \to \mathbb{R}$ satisfying*

$$\psi \text{ is concave} \tag{1a}$$

(1b) $$\psi \text{ is non-decreasing}$$

(1c) $$\psi(t) = t \qquad \text{for } t \leq 0$$

(1d) $$\lim_{t\to\infty} \psi(t+1) - \psi(t) = 0$$

Moreover, this correspondence is given by $P \mapsto \{E(X \wedge t), t \in \mathbb{R}\}$*, where* $\mathcal{L}(X) = P$*; and* $\psi'_-(t) \mapsto P([t,\infty))$*, where* ψ'_- *is the left-hand derivative* $\left(\lim_{h\nearrow 0} \frac{(\psi(t+h)-\psi(t))}{h}\right)$ *of* ψ*.*

Proof of Lemma 2.7. Suppose P is a Borel probability measure with support in $\mathbb{R}^+$, and let X be a random variable with $\mathcal{L}(X) = P$. Clearly $\psi_P(t) := E(X \wedge t)$ satisfies (1a)–(1d).

Conversely, suppose ψ satisfies (1a)–(1d). Then (1a) implies that ψ'_- exists and is left-continuous everywhere; (1b) implies that $\psi'_- \geq 0$ everywhere; (1c) implies that $\psi'_-(0) = 1$; and (1a,d) imply that $\psi'_-(t) \to 0$ as $t \to \infty$. Hence ψ'_- is the left-continuous version, $P([t,\infty))$, of the distribution function for a unique Borel probability P. $\square$

Remark. $\lim_{t\to\infty} \psi_P(t) = m < \infty$ if and only if P has finite mean m, in which case (1d) is trivial.

Proof of Theorem 2.6. By Lemma 2.7, ψ_{P_n} satisfies (1a–d), with $\psi = \psi_{P_n}$, for all $n \in \mathbb{N}$. Thus $g = \lim \psi_{P_n}$ is non-decreasing and concave, and satisfies $g(t) = t$ for $t \leq 0$. Thus by Lemma 2.7 again, there exists a Borel probability P with support in $\mathbb{R}^+$ and satisfying $\psi_P = g$ if and only if $g(t+1) - g(t) \to 0$ as $t \to \infty$.

It remains only to show that $\psi_{P_n} \to \psi_P$ implies that $P_n \xrightarrow{w} P$. This follows from the correspondence given in Lemma 2.7, which implies that the cdf's of P_n converge to the cdf of P at all continuity points of P. $\square$

As the following example shows, if the supports of (P_n) are not bounded below, then $P_n \xrightarrow{w} P$ does not imply that $\psi_{P_n} \to \psi_P$.

Example 2.8. Let $P_n = \frac{1}{n}\delta_{(-n)} + \left(\frac{n-1}{n}\right)\delta_{(1)}$ for $n \in \mathbb{N}$, and let $P = \delta_{(1)}$. Clearly $P_n \xrightarrow{w} P$, but $\lim \psi_{P_n}(0) = -1 \neq 0 = \psi_P(0)$.

Replacing $E(X \wedge t)$ by $E(X-t)^+$ in Definition 2.5 will yield analogous results for *mean residual life* functions. A related theorem in Krengel and Lin (1987) shows that if $E(X_n - t)^+$ and $E(X_n + t)^-$ converge to finite limits for all t as $n \to \infty$, then P_n $(= \mathcal{L}(X_n))$ always converges in distribution. Without the continuity condition, however, the limiting functions may in general not be $E(X-t)^+$ and $E(X+t)^-$.

§3 Distributions with Finite Means

In this section, several Lévy-like continuity theorems will be established for distributions with finite means.

For probability-representing functions such as the *potential function*

$$U_P(t) := E|X - t|$$

or the *mean residual life* function

$$L_P(t) := E(X-t)^+,$$

(cf. Van der Vecht (1986)), direct analogs of Theorem 2.6 for distributions with finite means can be established by replacing $E(X \wedge t)$ by $E|X-t|$ or $E|X-t|^+$, and replacing the hypothesis of positive support by the hypothesis of integrability. In a similar vein, the next theorem gives an analog for another classical representation function, the Hardy-Littlewood maximal function.

Definition 3.1. For a real Borel probability measure P with finite mean, let $H_P : (0,1) \to \mathbb{R}$ denote the *Hardy-Littlewood maximal function*

$$H_P(t) = \frac{1}{t} \int_0^t F^{-1}(s) ds,$$

where F^{-1} is the generalized upper inverse, or quantile function, corresponding to the cdf F for P.

Theorem 3.2. *Suppose (P_n) are real Borel probability measures with finite means and Hardy-Littlewood maximal functions (H_n), respectively. If $\lim_{n\to\infty} H_n(t) = g(t)$ for all $t \in (0,1)$, then there exists a Borel probability P with finite mean and satisfying $H_P = g$ if and only if $\lim_{t \nearrow 1} g(t) < \infty$, in which case $P_n \xrightarrow{w} P$.*

Sketch of Proof. First check that a Borel probability with finite mean is uniquely determined by its Hardy-Littlewood maximal function (via $F^{-1}(t) = H(t) + tH'(t)$), and that a function $h : (0,1) \to \mathbb{R}$ is a Hardy-Littlewood maximal function for some probability distribution with finite mean iff h is non-decreasing, $\lim\limits_{t \nearrow 1} h(t) < \infty$, and $th(t)$ is convex in t. Since it is the limit of such functions, g itself is non-decreasing and $tg(t)$ is convex in t, so the only remaining condition is finiteness of the left limit of g at 1. The conclusion $P_n \xrightarrow{w} P$ follows since $H_{P_n} \to H_P$ implies $F_{P_n}^{-1} \to F_P^{-1}$. $\square$

The next result, the main theorem in (Hill and Spruill (1994)), is based on a representation theorem of Hoeffding for maximal moments.

Definition 3.3. For a real Borel probability measure P with finite mean, $(M_k)_{k=1}^{\infty} = (M_k(P))_{k=1}^{1}$ are the *maximal moments*

$$M_k(P) = E(X_1 \vee \cdots \vee X_k)$$

where $X_1, X_2, \ldots, X_k$ are i.i.d. with $\mathcal{L}(X) = P$. (Note that $E|X| < \infty \Rightarrow M_k(P) < \infty$ for all k, where $\mathcal{L}(X) = P$.)

Theorem 3.4. *Suppose (P_n) are Borel probability measures with support in $\mathbb{R}^+$ and finite means. If $\lim\limits_{n\to\infty} M_k(P_n) = g_k$ for all $k \in \mathbb{N}$, then there exists a Borel probability P with $M_k(P) = g_k$ for all $k \in \mathbb{N}$ if and only if $\lim\limits_{k\to\infty} \frac{g_k}{k} = 0$, in which case $P_n \xrightarrow{w} P$.*

Proof. (See Hill and Spruill (1994); for an alternative proof based on characterizations of expectation sequences of maximal order statistics, see Kolodynski (2000).) $\square$

In some Lévy-like continuity theorems, the single necessary and sufficient condition on the limit function that it represent a probability distribution is most easily expressed in non-limit terms. In the next theorem, the crucial condition is easily expressed in terms of a sequence of equalities, a limit version of which is possible but cumbersome.

Definition 3.5. For a real Borel probability measure P with finite mean, let the *sequential barycenter array* $(m_{i,k}(P))_{i=1}^{\infty}\,{}_{k=1}^{z^i-1}$ be given inductively as follows:

$$
\begin{aligned}
m_{1,1}(P) &= \int xP(dx) \\
m_{i,2j}(P) &= m_{i-1,j} \\
m_{i,2j-1}(P) &= E(X \mid X \in (m_{i-1,j-1}, m_{i-1,j}]) && \text{if } F(m_{i-1,j}) > F(m_{i-1,j-1}) \\
&= m_{i-1,j}(P) && \text{otherwise}
\end{aligned}
$$

with the convention that $m_{i,0} = -\infty$ and $m_{i,2^i} = \infty$, where $\mathcal{L}(X) = P$ and F is the cdf for P.

Theorem 3.6. *Suppose $(P_n)_{n=1}^{\infty}$ are real Borel probability measures with finite means. If $\lim_{n\to\infty} m_{i,k}(P_n) = g_{i,k}$ for all i and k, then there exists a Borel probability P with $m_{i,k}(P) = g_{i,k}$ for all i and k if and only if*

$$
g_{i,4k-3} = g_{i,4k-2} \Leftrightarrow g_{i,4k-1} = g_{i,4k-2} \qquad \textit{for all } i \textit{ and } k, \tag{2}
$$

(or equivalently,

$$
\lim_{i\to\infty} \sum_{k=1}^{2^{i-2}} \big|\mathrm{sgn}|g_{i,4k-2} - g_{i,4k-3}| - \mathrm{sgn}|g_{i,4k-1} - g_{i,4k-2}|\big| = 0
$$

where sgn $a = -1$ *if* $a < 0$, $= 0$ *if* $a = 0$, $= 1$ *if* $a > 0$*) in which case* $P_n \xrightarrow{w} P$.

Proof. By (Hill and Monticino (1998), Theorem 2.9), a triangular array $(m_{i,k})_{i=1}^{\infty}\,{}_{k=1}^{2^i-1}$ is a sequential barycenter array for some P, that is, $m_{i,k}(P) = m_{i,k}$ for all i and k, if and only if

$$
m_{i,2j} = m_{i-1,j} \quad \text{for all } i \in \mathbb{N} \text{ and } j = 1, \ldots, 2^{i-1} - 1, \tag{3a}
$$

$$
m_{i,k-1} \le m_{i,k} \qquad \text{for all } i \in \mathbb{N} \text{ and } k = 1, \ldots, 2^i, \tag{3b}
$$

and

$$
m_{i,4k-3} = m_{i,4k-2} \Leftrightarrow m_{i,4k-1} = m_{i,4k-2} \quad \text{for all } i \text{ and } k. \tag{3c}
$$

Since the limit of a sequence of arrays satisfying (3a,b) also satisfies (3a,b), there is a Borel probability measure P with $m_{i,k}(P) = g_{i,k}$ for all i and k if and only if g satisfies (2). To see that this implies that $P_n \xrightarrow{w} P$, note that by the inversion theorem in (Hill and Monticino (1998), Theorem 2.7), $m_{i,k}(P_n) \to m_{i,k}(P)$ for all i and k implies that $F_n \to F$ at all continuity points of F. $\square$

On the other hand, $P_n \xrightarrow{w} P$ does not in general imply that $m_{i,k}(P_n) \to m_{i,k}(P)$ for all i, k, even for distributions with bounded support, as the next example shows.

Example 3.7. Let $P_n = \left(\frac{n-1}{n}\right)\delta_{(0)} + \frac{1}{n}\,\delta_{(1)}$, $P = \delta_{(0)}$. Then $P_n \xrightarrow{w} P$, but $m_{2,3}(P_n) \equiv 1 \neq 0 = m_{2,3}(P)$.

The final result in this section is a Lévy-like theorem for Borel probability measures with finite moments of all orders.

Definition 3.8. For a Borel probability measure P and positive integer k with $\int |x|^k dP(x) < \infty$, let $\mu_k(P)$ denote the k^{th} moment $\mu_k(P) = \int_{\mathbb{R}} x^k dP(x)$ of P, and let

$$\mathcal{C} = \left\{ P : \limsup_{k\to\infty} \frac{\mu_{2k}(P)^{\frac{1}{2k}}}{k} < \infty \right\}.$$

Theorem 3.9. *Suppose $(P_n)_{n=1}^{\infty}$ are in $\mathcal{C}$. If $\lim_{n\to\infty} \mu_k(P_n) = g_k$ is finite for all $k \in \mathbb{N}$, then there exists a P in $\mathcal{C}$ with $\mu_k(P) = g_k$ for all $k \in \mathbb{N}$ if and only if $\limsup_{k\to\infty} \frac{g_{2k}^{\frac{1}{2k}}}{k} < \infty$, in which case $P_n \xrightarrow{w} P$.*

Proof. Fix $(P_n)_{n=1}^{\infty}$, and let $\mu_{n,k} = \int_{\mathbb{R}} x^k dP_n(x)$ for all $n, k \in \mathbb{N}$. By hypothesis

$$\lim_{n\to\infty} \mu_{n,k} = g_k \qquad \text{for all } k \in \mathbb{N}. \tag{4}$$

By the classical result for the Hamburger moment problem (e.g., Kawata (1972), Theorem 11.1.5), given a sequence of real numbers (g_k), there exists a (not necessarily unique) Borel probability measure P satisfying

$$\mu_k(P) = g_k \qquad \text{for all } k \in \mathbb{N} \tag{5}$$

if and only if the sequence $\{g_k\}$ is non-negative definite, that is, if and only if

$$|g_{i+j}|_{i,j=0}^{n} \geq 0 \qquad \text{for all } n = 0, 1, 2 \ldots, \tag{6}$$

where $|g_{i+j}|$ is the determinant of the $(n+1) \times (n+1)$ matrix with (i,j)-th entry g_{i+j} (and $g_0 := 1$).

Thus, since P_n has finite kth moments for all n and k, it follows that the sequence $(\mu_{n,k})_{k=1}^{\infty}$ is non-negative definite for each n. By (4) and the continuity of the determinant function, the limit sequence (g_k) satisfies (6), which (by the classical result again) implies there is at least one Borel measure P satisfying (5).

Next, suppose that (g_k) satisfies

$$\limsup_{k\to\infty} \frac{g_{2k}^{\frac{1}{2k}}}{k} < \infty. \tag{7}$$

Then there exist $M > 0$ with $\frac{g_{2k}^{\frac{1}{2k}}}{k} \leq M$ for all $k \in \mathbb{N}$, so

$$\sum_{k=1}^{\infty} \frac{1}{g_{2k}^{\frac{1}{2k}}} \geq \sum_{k=1}^{\infty} \frac{1}{kM} = \infty.$$

By Carleman's theorem (e.g., Kawata (1972) Theorem 11.1.7), this implies that the P satisfying (5) is unique, and by Polyá's theorem (Polyá (1920)), (7) implies that $P_n \xrightarrow{w} P$.

The converse is trivial by definition of $\mathcal{C}$. $\square$

Remarks. It easily follows that the limit of moment sequences of distributions with the same compact support is itself always a moment sequence of a unique Borel probability distribution which is the weak limit of the original probability distributions. Even with unbounded support, as the proof shows, the (finite) limit of moment sequences is still the moment sequence for a distribution, but uniqueness (and hence weak convergence) may fail. This may easily be seen by letting P, Q be distinct Borel measures with $|\mu_k(P)| < \infty$ and $\mu_k(P) = \mu_k(Q)$ for all $k \in \mathbb{N}$, (cf., Feller, Vol. II (1971)), and then taking $P_n \equiv P$ for all $n \in \mathbb{N}$. Then $\lim_{n\to\infty} \mu_k(P_n) = \mu_k(Q)$ for all k, but $P_n \overset{w}{\not\to} Q$.

§4 Discrete and Absolutely Continuous Distributions

For a Borel probability measure P on $\mathbb{R}$, let $\varphi_P : \mathbb{R} \to \mathbb{C}$ denote the characteristic function $\varphi_P(t) = E(e^{itX})$, where $\mathcal{L}(X) = P$. (This simply is the classical 1-dimensional case of Definition 3.1.) The next theorem is a simple example of a continuity theorem for discrete distributions.

Theorem 4.1. *Suppose $(P_n)_{n=1}^{\infty}$ are Borel probability measures with support in $\mathbb{Z}$. If $\varphi_{P_n} \to g$, then there exists a Borel probability P with support in $\mathbb{Z}$, and satisfying $\varphi_P = g$, if and only if g is continuous, in which case $P_n \xrightarrow{w} P$.*

Proof. As is well known (cf., Fristedt and Gray (1997), Theorem 13.13), a function $\varphi : \mathbb{R} \to \mathbb{C}$ is the characteristic function of a $\mathbb{Z}$-valued random variable if and only if

$$\varphi(0) = 1, \tag{8a}$$

$$\sum_{k=1}^{n} \sum_{j=1}^{n} \varphi(v_k - v_j) z_j \bar{z}_k \geq 0, \tag{8b}$$

for all $n = 1, 2, \ldots$, all complex n-tuples $(z_1, \ldots, z_n)$, and all real n-tuples $(v_1, \ldots, v_n)$,

$$\varphi \text{ is periodic with period } 2\pi, \tag{8c}$$

$$\varphi \text{ is continuous.} \tag{8d}$$

Since limits of functions satisfying (8a)–(8c) clearly satisfy those same properties, there is a Borel probability P with $\varphi_P = g$ if and only if φ is continuous, in which case $P_n \xrightarrow{w} P$ as in the classical Lévy continuity theorem. □

The next two theorems record easy examples of continuity theorems for absolutely continuous distributions.

Theorem 4.2. *Suppose $(P_n)_{n=1}^{\infty}$ are absolutely continuous probability measures with density functions (f_n) respectively. If $\lim_{n\to\infty} f_n = g$, then there exists an absolutely continuous probability P with density g if and only if $\lim_{N\to\infty} \int_{-N}^{N} g\, dx = 1$, in which case $P_n \xrightarrow{w} P$.*

Proof. For each $n \in \mathbb{N}$, f_n is Borel measurable and (almost surely) non-negative, so the limit function g is also measurable and non-negative. Thus if $\int_{\mathbb{R}} g = 1$, defining P by $P(A) = \int_A g\,dx$ for all $A \in \mathbb{B}$ it follows that P is an absolutely continuous probability measure with density g, in which case $f_n \to f$ implies that $P_n \xrightarrow{w} P$ (cf., Feller (1971) p. 25). $\square$

Remark. Even if $P_n \xrightarrow{w} P$, where each P is absolutely continuous and P is also a.c. with *continuous* density f, it does not follow that $f_n \to f$, as the example in Feller (*Ibid*) shows.

Definition 4.3. For an absolutely continuous probability measure P with density f and cdf F, and with support in $[0,\infty)$, the *failure rate function* for P, $r_P : \mathbb{R}^+ \to \mathbb{R}^+$, is given by

$$r(t) = r_P(t) = \frac{f(t)}{1 - F(t)}.$$

Theorem 4.4. *Suppose* $(P_n)_{n=1}^{\infty}$ *are absolutely continuous Borel probability measures with support in* $[0,\infty)$*, and failure rate functions* (r_n) *respectively. If* $\lim_{n\to\infty} r_n = g$*, then there exists an absolutely continuous probability* P *with failure rate* $r_P = g$ *if and only if* $\lim_{N\to\infty} \int_0^N g(x)dx = \infty$*, in which case* $P_n \xrightarrow{w} P$.

Proof. If r is the failure rate function for an a.c. probability P, then

$$F(t) = 1 - e^{-\int_{-\infty}^{t} r(x)dx}$$

(since $\int r = \int \frac{f}{1-F} = -\ln(1-F)$ implies that $e^{-\int r} = 1 - F$).

Thus $g : \mathbb{R}^+ \to \mathbb{R}$ is a failure rate function (for some P) if and only if g is Borel measurable, almost surely non-negative, and $\int_0^\infty g(x)dx = \infty$. The limit function g is non-negative and measurable since the (r_n) are, so it is a failure rate function if and only if it has unbounded integral. In this case, since $r_n \to g$ implies that $F_n \to F$ everywhere, clearly $P_n \xrightarrow{w} P$. $\square$

§5 Applications

The Lévy-like continuity theorems above are useful in establishing weak limit theorems by choosing a representation which facilitates the analysis of the distributions in question. Just as the characteristic function is useful in establishing the classical central limit theorem precisely because the characteristic function of a sum of iid variables is easy to calculate, so the maximal moment method (Theorem 3.4) facilitates proof of classical weak laws in extreme value theory simply because the maximal moments of sample maxima are easy to calculate (cf., Hill and Spruill (1994), Proposition 4.1).

Acknowledgment

The first author is grateful to the Göttingen Academy of Sciences for support as a Gauss Professor, and to the Institut für Mathematische Stochastik of the University of Göttingen for its hospitality during Fall Semester 2000.

References

Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

Breiman, L. (1968). *Probability*, Addison-Wesley, Reading.

Dette, H. and Studden, W. (1997). *The Theory of Canonical Moments with Applications*, Wiley, New York.

Feller, W. (1971). *An Introduction to Probability and its Applications*, Vol. II, 2nd Ed., Wiley, New York.

Fristedt, B. and Gray, L. (1997). *A Modern Approach to Probability Theory*, Birkhäuser, Boston.

Gilat, D. (1987). "On the best order of observation in optimal stopping problems," *J. Applied Probab.* **24**, 773–778.

Gilat, D. and Hill, T. (1993). "Quantile-locating functions and the distance between the mean and quantiles," *Statistica Neerlandia* **47**, 279–283.

Hill, T. and Monticino, M. (1998). "Constructions of random distributions via sequential barycenters," *Ann. Statist.* **26**, 1242–1253.

Hill, T. and Spruill, M. (1994). "On the relationship between convergence in distribution and convergence of expected extremes," *Proc. Amer. Math. Soc.* **121**, 1235–1243.

Kawata, T. (1972). *Fourier Analysis in Probability Theory*, Academic Press, New York.

Kolodynski, S. (2000). "A note on the sequence of expected extremes," preprint, University of Tennessee.

Krengel, U. and Lin, M. (1987). "Nonexpansive operators," *Israel J. Math.* **58**, 170–183.

Pollack, M. (1973). "On equal distributions," *Ann. Statist.* **1**, 180–182.

Polyá, G. (1920). *Math. Zeitschrift*.

Van der Vecht (1986). *CWI Tract* 21, Math. Centrum, Amsterdam.