ALTERNATIVE PROOFS OF SOME THEOREMS ON CHARACTERISTIC FUNCTIONS

By P. K. PATHAK
Indian Statistical Institute

SUMMARY. The product of two characteristic functions is absolutely integrable whenever one of them is absolutely integrable. This fact and the inversion formula for the normal distribution are exploited here to provide alternative proofs of the uniqueness theorem, Levy's continuity theorem, Bochner's theorem and the Herglotz lemma.

1. Introduction

The standard proofs of the uniqueness theorem, Levy's continuity theorem and Bochner's theorem on characteristic functions (henceforth to be referred to as the basic theorems) depend heavily on evaluation of limits of integrals of the kind $\int_{-T}^{T} h(x)dx$ as T tends to infinity. Evaluation of such limits is generally uninteresting because the corresponding improper integrals do not always exist. Moreover such integrals tend to obscure the crucial aspects of the proofs of the basic theorems. An interesting question that can be raised now is whether it is possible to give alternative proofs of the basic theorems which would not require the evaluation of limits of the kind $\lim_{T\to\infty} \int_{-T}^{T} h(x)dx$. Dr. Sethuraman, a colleague of the author, believed that this could probably be done by composing characteristic functions under investigation with an absolutely integrable characteristic function and investigating the new characteristic functions so obtained. The main object of this paper is to provide an affirmative answer to the above question and to show that the technique of composing characteristic functions with an absolutely integrable characteristic function leads to considerable simplifications and at the same time brings out more vividly the essential aspects of the proofs of the basic theorems. Our method of attack is to compose the given characteristic function with the characteristic function of the normal distribution and then repeatedly apply the inversion formula for the normal distribution. The choice of the characteristic function of the normal distribution is entirely arbitrary here (at least for uniqueness and continuity theorems) and in fact any absolutely integrable characteristic function could have been used for this purpose. We begin now by giving first a few preliminary definitions and results necessary for our development.

For brevity in exposition we consider throughout distribution functions on the real line, R_1, rather than on a general n-dimensional Euclidean space, $(n = 1, 2, \ldots)$. What is done in the sequel for the real line carries word for word for finite dimensional Euclidean spaces.

Definition 1.1 : A sequence of distribution functions $F_1, F_2, \ldots$ is said to converge weakly to a distribution function F if the sequence $F_1(x), F_2(x), \ldots$ converges to $F(x)$ at each continuity point x of F.

Definition 1.2 : A sequence of distribution functions $F_1, F_2, \ldots$ is called *weakly precompact* if every subsequence from $F_1, F_2, \ldots$ contains a further subsequence which converges weakly to a distribution function.

Definition 1.3 : A distribution function F is composed of distribution function F_1, F_2, written as $F = F_1 * F_2$, if

$$F(x) = \int F_1(x-y)d\ F_2(y) \qquad \ldots \quad (1)$$

for $x \in R_1$.

If $F = F_1 * F_2$ then the characteristic functions φ, φ_1 and φ_2 of F, F_1 and F_2 respectively satisfy

$$\varphi(t) = \varphi_1(t)\,\varphi_2(t) \qquad \ldots \quad (2)$$

for $t \in R_1$.

If $F = F_1 * F_2$, F_1 absolutely continuous with the density function f_1, then F is absolutely continuous. The density function f of F is given by

$$f(x) = \int f_1(x-y)\,dF_2(y) \qquad \ldots \quad (3)$$

for $x \in R_1$.

Let G_σ denote the normal distribution function with mean zero and variance σ^2, i.e.

$$G_\sigma\ (x) = \int_{-\infty}^{x} \exp\left(-u^2/2\sigma^2\right)\,du/\sigma\sqrt{2\pi} \qquad \ldots \quad (4)$$

for $x \in R_1$.

The following, which is an inversion theorem for the Normal, is then well-known

$$\int \exp\left(itx - x^2/2\sigma^2\right)\,dx/\sigma\ \sqrt{2\pi} = \exp\left(-t^2\sigma^2/2\right). \qquad \ldots \quad (5)$$

In the next section we establish the uniqueness theorem.

2. Uniqueness of characteristic functions

Every distribution function clearly determines a unique characteristic function. The converse follows from the following theorem.

Theorem 1 : *Every characteristic function uniquely determines a distribution function.*

Proof : Let $\phi(t)$ be a characteristic function. If possible let F_j $(j = 1, 2)$ be two distribution functions having the characteristic function $\phi(t)$. Let $H_{j_\sigma} = F_j * G_\sigma$. Then H_{j_σ} is absolutely continuous since G_σ is absolutely continuous. The density function h_{j_σ} of H_{j_σ} is given by

$$h_{j_\sigma}(x) = \int \exp\left(-(x-y)^2/2\sigma^2\right)dF_j(y)/\sigma\sqrt{2\pi} \qquad \ldots \quad (6)$$

for $x \in R_1$ $(j = 1, 2)$.

The characteristic function of $H_{j\sigma}$ $(j = 1, 2)$ is given by $\varphi(t)\exp(-t^2\sigma^2/2)$. It is demonstrated below that $h_{1\sigma} \equiv h_{j\sigma}$ for $\sigma > 0$. Actually this is immediate from the followgin equation

$$\begin{aligned}\int \varphi(t)\exp(-itx-t^2\sigma^2/2)\,dt/2\pi \\ &= \int \exp(-itx-t^2\sigma^2/2)\,dt \int \exp(ity)\,dF_j(y)/2\pi \\ &= \int dF_j(y) \int \exp(-it(x-y)-t^2\sigma^2/2)\,dt/2\pi \\ &= \int \exp(-(x-y)^2/2\sigma^2)\,dF_j(y)/\sigma\sqrt{2\pi} \\ &= h_{j\sigma}(x) \qquad \ldots (7)\end{aligned}$$

since the left side of (7) is the same for $j = 1, 2$.

Hence $H_{1\sigma} \equiv H_{2\sigma}$ for $\sigma > 0$. This implies that

$$\int F_1(x-\sigma y)dG_1(y) = \int F_2(x-\sigma y)\,dG_1(y) \qquad \ldots (8)$$

where G_1 is the normal distribution function with zero mean and unit variance.

On letting $\sigma \to 0$ in (8), we obtain $F_1(x) = F_2(x)$ for all common continuity points of F_1 and F_2. Hence $F_1 \equiv F_2$.

This completes the proof of Theorem 1.

3. Proof of Levy's continuity theorem

The following lemmas are considered necessary in order to provide an alternative proof of Levy's continuity theorem.

Lemma 1 : *The sequence of distribution functions $F_1, F_2, \ldots,$ is weakly precompact if and only if for each $\varepsilon > 0$ there exists a number K_ε such that*

$$F_n(K_\varepsilon)-F_n(-K_\varepsilon) > 1-\varepsilon \qquad \ldots (9)$$

for all n.

Proof : Invoke Helly's weak compactness theorem (Loéve, 1962).

Lemma 2 : *Let $F_1, F_2, \ldots$ be a sequence of distribution functions. Let G be a given distribution function. Then the sequence $G, F_1 * G, F_2 * G, \ldots$ is weakly precompact if and only if $F_1, F_2, \ldots$ is weakly precompact.*

Proof : It suffices to prove the necessity part of the lemma. By Lemma 1 if $G, F_1 * G, F_2 * G, \ldots$ is weakly precompact then for each $\varepsilon > 0$ there exists a K_ε such that $G(K_\varepsilon)-G(-K_\varepsilon) > 1-\varepsilon$ and $F_n * G(K_\varepsilon)-F_n * G(-K_\varepsilon) > 1-\varepsilon$ for all n. Hence $F_n(2K_{\varepsilon/2}) - F_n(-2K_{\varepsilon/2}) > 1-\varepsilon$. This proves the stated lemma.

We have thus the following theorem.

Theorem 2 : *A necessary and sufficient condition for a sequence of distribution functions $F_1, F_2, \ldots$ to converge weakly to a distribution F_0 is that the corresponding sequence of characteristic functions $\varphi_1(t), \varphi_2(t), \ldots$ converge pointwise to a function $\varphi_0(t)$ which is continuous at $t = 0$; in which case $\varphi_0(t)$ is the characteristic function of F_0.*

Proof : That the condition is necessary is obvious. To prove the sufficiency it can be seen that it suffices to prove that $F_1, F_2, \ldots$ is weakly precompact if the corresponding sequence of characteristic functions $\varphi_1(t), \varphi_2(t), \ldots$ converges pointwise to a function $\varphi_0(t)$ which is continuous at $t = 0$. To prove this, consider the new sequence of distribution functions $F_1 * G, F_2 * G, \ldots$ where G is the distribution function of the standard normal distribution. The characteristic function of $F_n * G$ is $\psi_n(t) = \varphi_n(t) \exp(-t^2/2)$ $(n = 1, 2, \ldots)$. $F_n * G$ is absolutely continuous because G is absolutely continuous. From (7) the density function of $F_n * G$ is given by

$$f_n(x) = \int \varphi_n(t) \exp(-itx - t^2/2)\, dt/2\pi \quad (n = 1, 2, \ldots). \qquad \ldots (10)$$

By the dominated convergence theorem, we have

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \int \varphi_n(t) \exp(-it\,x - t^2/2)\, dt/2\pi$$

$$= \int \varphi_0(t) \exp(-itx - t^2/2)\, dt/2\pi$$

$$= f_0(x) \quad \text{(say)}. \qquad \ldots (11)$$

Thus the sequence of density functions $f_1(x), f_2(x), \ldots$ converges pointwise to $f_0(x)$. We now prove that $f_0(x)$ is a density function. By the monotone convergence theorem

$$\int f_0(x) dx = \lim_{\delta\to 0} \int f_0(x) \exp(-x^2\delta^2/2)\, dx$$

$$= \lim_{\delta\to 0} \int \exp(-x^2\delta^2/2)\, dx \int \varphi_0(t) \exp(-itx - t^2/2)\, dt/2\pi$$

$$= \lim_{\delta\to 0} \int \varphi_0(t) \exp(-t^2/2)\, dt \int \exp(-itx - x^2\delta^2/2)\, dx/2\pi$$

$$= \lim_{\delta\to 0} \int \varphi_0(t) \exp(-t^2/2 - t^2/2\delta^2)\, dt/\delta\sqrt{2\pi}$$

$$= \lim_{\delta\to 0} \int \varphi_0(u\delta) \exp(-u^2\delta^2/2 - u^2/2)\, du/\sqrt{2\pi}$$

$$= \int \exp(-u^2/2)\, du/\sqrt{2\pi} = 1. \qquad \ldots (12)$$

by virtue of the continuity of $\varphi_0(t)$ at $t = 0$.

This shows that the sequence of density functions of G, $F_1 * G$, $F_2 * G, \ldots$ converges pointwise to a probability density function. Next by (10) $|f_n| \leqslant 1/\sqrt{2\pi}$ $(n = 1, 2, \ldots)$. Hence $\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b f_0(x)\, dx$. This implies that $G, F_1 * G, F_2 * G, \ldots$ is weakly precompact. Consequently by Lemma 2, $F_1, F_2, \ldots$ is weakly precompact.

This completes the proof.

4. Proof of Bochner's theorem

We now state and prove Bochner's theorem.

Theorem 3 : *A function $g(\cdot)$ defined on the real line is non-negative definite and continuous with $g(0) = 1$ if and only if it is a characteristic function.*

Proof : It is recalled that a function is non-negative definite if for each positve integer n, real numbers $t_1, t_2, \ldots, t_n$ and a function $h(\cdot)$ defined on the real line, the following holds

$$\sum_{i=1}^{n} \sum_{j=1}^{n} g(t_i - t_j)\, h(t_i)\, \overline{h(t_j)} \geqslant 0. \qquad \ldots \quad (13)$$

It is easy to verify that a characteristic function is non-negative definite. To prove the converse, it is seen that (13) implies that

$$\int\int g(u-v) \exp\{-i(u-v)x - (u^2+v^2)\sigma^2\}\, du\, dv \geqslant 0 \qquad \ldots \quad (14)$$

for each x and $\sigma^2 > 0$ because the above integral can be approximated by means of sums of the form of (13).

The above integrand can be written as

$$\int\int g(u-v) \exp\left\{-i(u-v)x - \frac{(u-v)^2\sigma^2}{2} - \frac{(u+v)^2\sigma^2}{2}\right\} du\, dv \geqslant 0. \qquad \ldots \quad (15)$$

Letting $t = u-v$ and $s = u+v$ and integrating over s, we obtain

$$f(x, \sigma) = \int g(t) \exp(-itx - t^2\sigma^2/2)\, dt/2\pi \geqslant 0. \qquad \ldots \quad (16)$$

Clearly $f(x, \sigma)$, as a function of x, is a density function because an argument similar to that of (12) gives

$$\int f(x, \sigma)\, dx = \lim_{\delta \to 0} \int f(x, \sigma) \exp(-x^2\delta^2/2)\, dx = 1. \qquad \ldots \quad (17)$$

A similar argument shows that the characteristic function of $f(x, \sigma)$ is given by

$$\int \exp(it_1 x) f(x, \sigma)\, dx = \lim_{\delta \to 0} \int f(x, \sigma) \exp(it_1 x - x^2\delta^2/2)\, dx$$

$$= \lim_{\delta \to 0} \int g(t_1 + u\delta) \exp\left(-\frac{(t_1+u\delta)^2\sigma^2}{2} - \frac{u^2}{2}\right) du/\sqrt{2\pi}$$

$$= g(t_1) \exp(-t_1^2\sigma^2/2) \qquad \ldots \quad (18)$$

by virtue of the continuity of $g(\cdot)$. Hence for each σ, $g(t) \exp(-t^2\sigma^2/2)$ is a characteristic function.

Since $\lim_{\sigma \to 0} g(t) \exp(-t^2\sigma^2/2) = g(t)$ and $g(t)$ is continuous, by Levy's continuity theorem it follows that $g(t)$ is in fact a characteristic function.

This concludes the proof of Bochner's theorem.

Concluding remarks : The Herglotz lemma can be proved in a similar fashion. To prove this one need replace the integrals in the proof of Bochner's theorem by infinite sums and make use of the orthonormality of $\{\exp\ (itx) : t = 0, \pm 1, \pm 2, ...\}$ in $L_2(0, 2\pi)$. Finally it can be by checked that Cramer's theorem too can be proved in a similar manner.

5. Acknowledgement

I am grateful to my colleague, Dr. J. Sethuraman, who suggested me the possible alternative approach for proving the theorems given in this paper.

References

Doob, J. L. (1953) : *Stochastic Processes*, John Wiley and Sons, New York.

Loève, M. (1962) : *Probability Theory*, 3rd ed. Van Nostrand, West Germany.

Paper received : April, 1966.