A LIMIT THEOREM FOR DENSITIES

By R. RANGA RAO

and

V. S. VARADARAJAN

Indian Statistical Institute, Calcutta

SUMMARY. In this paper necessary and sufficient conditions are obtained for the almost everywhere convergence of the density of the normed sum of a sequence of independent and identically distributed random variables, to the density of the standard normal distribution. Extensions to (i) multidimensions and (ii) the case of convergence to stable law are also given. An application to the convergence of conditional distributions is also pointed out.

1. Introduction

Let $\xi_1, \xi_2, \ldots \xi_n, \ldots,$ be a sequence of independent random variables with common distribution function $F(x)$. Let $F_n(x)$ be the distribution function of the normed sums

$$\eta_n = \frac{\xi_1+\xi_2+\ldots+\xi_n}{\sqrt{n}} \qquad \ldots \quad (1.1)$$

and let $p_n(x)$ denote the density of the absolutely continuous component of F_n. It is well-known that $F_n \Longrightarrow \Phi$ ($\Longrightarrow$ denotes weak convergence) if and only if (i) $E\xi_1 = 0$ and (ii) $E\xi_1^2 = \sigma^2 < \infty$. (See Kolmogorov and Gnedenko (1954), p. 181). Here $\Phi(x)$ denotes the standardized normal distribution. Kolmogorov and Gnedenko in their book (1954, p. 223) have given an example which shows that the density $p_n(x)$ need not 'converge' to $\phi(x)$, the density of $\Phi(x)$. However, an examination of the example reveals that convergence fails to take place only in a neighbourhood of the origin. We are therefore naturally led to ask whether it is true that $p_n(x) \to \phi(x)$ almost everywhere. The object of this note is to answer this question. Theorem 1 shows that almost everywhere convergence can indeed be always asserted.

In this connection, it may be pointed out that complete solutions of the problem are already available for certain other definitions of convergence. Thus, for instance, Gnedenko (1954) has shown that necessary and sufficient conditions that

$$\operatorname*{ess\ sup}_{-\infty < x < +\infty} |p_n(x)-\phi(x)| \to 0 \qquad \ldots \quad (1.2)$$

are (i) $F_n(x) \Longrightarrow \Phi(x)$, and (ii) for some n_0, $F_{n_0}(x)$ satisfies the Lipschitz condition of order 1. Again, Prohorov (1952), (see also Smith, (1953)) has shown that necessary and sufficient conditions that

$$\int_{-\infty}^{\infty} |p_n(x)-\phi(x)|\,dx \to 0 \qquad \ldots \quad (1.3)$$

are (i) $F_n(x) \Longrightarrow \Phi(x)$, and (ii) F_n is non-singular for some n. Now, from (1.3) it at once follows that $p_n(x)$ converges to $\phi(x)$ in measure and this makes it plausible that a.e. convergence takes place.

2. THEOREMS

Theorem 1 : *With the above notation the necessary and sufficient conditions that*

$$p_n(x) \to \phi(x) \qquad \ldots \quad (2.1)$$

almost everywhere are (i) $F_n(x) \Longrightarrow \Phi(x)$, *and* (ii) F_n *is non-singular for some* n.

Proof: Necessity. The necessity of (i) follows from Scheffés theorem (1947), if we notice that in view of (2.1), the contribution of the singular component of F_n tends to zero as $n \to \infty$. The necessity of (ii) is obvious.

Sufficiency: Let n_0 be the least integer n for which F_n is non-singular. Such an n_0 exists in view of condition (ii) of the theorem. We can then write $F^{*n_0} = F_A + F_S$ where $F_A(+\infty) = \gamma (> 0)$ is the absolutely continuous component of F^{*n_0}. We can therefore choose k so large that if $E = \{x : p(x) < k\}$, $p(x)$ being the density of F_A, then $\int_E p(x)dx = \alpha$ where $0 < \alpha < \gamma$. We can then define

$$G_1(x) = \frac{1}{\alpha} \int_{-\infty}^{x} p(x)\chi_E dx$$

where χ_E is the characteristic function of the set E, and write

$$F^{*n_0} = \alpha G_1 + \beta G_2 \qquad \ldots \quad (2.2)$$

where $\beta = 1-\alpha$ and G_1, G_2 are distribution functions. By our choice G_1 is absolutely continuous with bounded density and G_2 is not completely singular. Let $g_1(t)$ and $g_2(t)$ be the c.f.'s (characteristic functions) of the two distribution functions $G_1(x)$ and $G_2(x)$ respectively. Then since G_1 has bounded density it follows that $g_1 \in L_2(-\infty, \infty)$. (See for instance Bochner and Chandrasekharan, (1949), p. 20). Further, since both G_1 and G_2 are non-singular, for any $\delta > 0$ we can find a $\rho = \rho(\delta)$ such that $0 < \rho < 1$ and

$$|g_1(t)| < \rho, \; |g_2(t)| < \rho \qquad \ldots \quad (2.3)$$

for all $|t| > \delta$.

Now suppose that $n = n_0 k + r$, where k and r are positive integers with $0 \leqslant r < n_0$. Then (* denoting convolution)

$$\begin{aligned} F^{*n} = F^{*n_0k} * F^{*r} &= (\alpha G_1 + \beta G_2)^{*k} * F^{*r} \\ &= \left[\sum_{s=0}^{k} \binom{k}{s} \alpha^s \beta^{k-s} G_1^{*s} * G_2^{*k-s} \right] * F^{*r} \\ &= \left\{ \sum_{s>2} \right\} * F^{*r} + \left\{ \sum_{s<2} \right\} * F^{*r} \\ &= H_n(x) + K_n(x) \quad \text{say.} \qquad \ldots \quad (2.4) \end{aligned}$$

It is then clear that $H_n(x)$ is absolutely continuous and let $q_n(x)$ denote the density of $H_n(x\sqrt{n})$. The total mass of K_n is at most $k\,\alpha\,\beta^{k-1}+\beta^k \leqslant C\,\beta^{k/2}$ where C is a constant independent of k. This implies that

$$\int |p_n(x)-q_n(x)|\,dx < C\,\beta^{k/2}. \qquad \ldots \quad (2.5)$$

Now since $0 < \beta < 1$, it follows that $\sum\limits_k \beta^{k/2} < \infty$. By the Borel Cantelli lemma we can then assert that

$$|p_n(x)-q_n(x)| \to 0, \text{ as } n \to \infty \qquad \ldots \quad (2.6)$$

for almost all (Lebesgue) x.

We now proceed to show that $q_n(x)$ tends to $\phi(x)$ uniformly in x. Let $h_n(t)$ denote the c.f. of $H_n(x\sqrt{n})$. Then since $g_1 \in L_2$ and as is easily verified $|h_n(t)| \leqslant |\, g_1(t/\sqrt{n})|^2$ it follows that $h(t) \in L_1$. We then have by the inversion theorem

$$q_n(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} h_n(t) e^{-itx}\,dt.$$

Hence
$$|q_n(x)-\phi(x)| \leqslant \frac{1}{2\pi}\int_{-\infty}^{\infty} |\, h_n(t)-e^{-t^2/2}|\,dt$$

$$\leqslant \frac{1}{2\pi}\int_{|t|<A} |h_n(t)-e^{-t^2/2}|\,dt+\frac{1}{2\pi}\int_{A}^{\infty} e^{-t^2/2}dt$$

$$+\frac{1}{2\pi}\int_{A<|t|<\delta\sqrt{n}} |h_n(t)|\,dt+\frac{1}{2\pi}\int_{|t|>\delta\sqrt{n}} |h_n(t)|\,dt$$

$$= I_1+I_2+I_3+I_4 \text{ say.}$$

Since ($f_n(t)$ denoting the c.f. of F_n)

$$|f_n(t)-h_n(t)| \leqslant \beta^k+k\,\alpha\,\beta^{k-1} \qquad \ldots \quad (2.7)$$

and $f_n(t) \to e^{-t^2/2}$ it follows that $h_n(t) \to e^{-t^2/2}$,

that is
$$I_1 = \frac{1}{2\pi}\int_{|t|<A} |h_n(t)-e^{-t^2/2}|\,dt = o(1) \qquad \ldots \quad (2.8)$$

for any fixed A. By choosing A sufficiently large I_2 can be made small.

From $F_n(x) \Longrightarrow \Phi(x)$ we can deduce that $F(x)$ has finite second moment. (See Kolmogorov and Gnedenko (1954), p. 181). Hence we can find $\delta > 0$ such that for $|t| \leqslant \delta$ we have

$$|f(t)| \leqslant \exp\left(-\frac{\sigma^2 t^2}{4}\right)$$

where σ^2 is the variance. From this and (2.7) it follows that

$$I_3 \leqslant \frac{1}{2\pi} \int\limits_{A<|t|<\delta\sqrt{n}} |f_n(t)|\,dt$$

$$+\frac{1}{2\pi}\, 2\delta\sqrt{n}\,(\beta^k+k\alpha\beta^{k-1})$$

$$\leqslant \frac{1}{\pi}\int\limits_A^{\delta\sqrt{n}} e^{-t^2\sigma^2/4}\,dt+o(1)$$

$$\leqslant \frac{1}{\pi}\int\limits_A^{\infty} e^{-t^2\sigma^2/4}\,dt+o(1) \qquad \dots \quad (2.9)$$

i.e., I_3 can also be made small by choosing A sufficiently large.

Finally from (2.3) it follows that

$$|h_n(t)| \leqslant \rho^{k-2}\,|g_1(t/\sqrt{n})|^2$$

so that $$I_4 \leqslant \rho^{k-2} \int\limits_{|t|>\delta\sqrt{n}} |g_1(t/\sqrt{n})|^2\,dt = \rho^{k-2}\sqrt{n} \int\limits_{|t|>\delta} |g_1(t)|^2\,dt. \qquad \dots \quad (2.10)$$

i.e., $I_4 = o(1)$ as $n \to \infty$.

Thus combining (2.8), (2.9) and (2.10), we can conclude that

$$|q_n(x)-\phi(x)| \to 0 \qquad \dots \quad (2.11)$$

uniformly in x. This completes the proof.

By a standard modification of the above argument, Theorem 1 can be generalized to stable distribution functions.

Theorem 2 : *Let* $\xi_1, \xi_2, \dots \xi_n, \dots$, *be a sequence of independent and identically distributed random variables, and let*

$$\eta_n = \frac{\xi_1+\xi_2+\dots+\xi_n-A_n}{B_n}.$$

Then the necessary and sufficient conditions that there exist constants A_n *and* B_n (>0), *such that* $p_n(x) \to \gamma(x)$ *almost everywhere, where* $p_n(x)$ *is the density of the absolutely continuous component of* η_n *and* $\gamma(x)$ *is the density of the stable law* $\Psi(x)$, *are*

(i) F *belongs to the domain of attraction of the stable law* Ψ, *and*

(ii) F_n *is non-singular for some* n.

The proof of the above theorem is entirely analogous to that of Theorem 1 and so is omitted.

3. Concluding remarks

With regard to the second condition of the theorem viz., F_n or what is the same thing F^{*n} is non-singular for some n, we may observe that there exist singular distributions F for which F^{*2} is absolutely continuous. (See Salem (1942)). However a characterization of this class of distributions intrinsically in terms of F is not known.

It is worth while to note that Theorem 1 is valid for distributions in finite-dimensional Euclidean spaces. That is, let $X_1, X_2, \ldots$ be a sequence of independent random vectors in k-dimensional Euclidean space E_k with mean zero and variance covariance matrix Σ. Let $p_n(x)$ denote the density function of the distribution F_n of the normed sum $\xi_n = (X_1+X_2+\ldots+X_n)/\sqrt{n}$ and let $\phi(x)$ denote the density of a Normal distribution $\Phi(x)$ with mean vector zero, and variance covariance matrix Σ. Then the necessary and sufficient conditions that $p_n(x) \to \phi(x)$ almost everywhere are (i) $F_n(x) \Longrightarrow \Phi(x)$, and (ii) $F_n(x)$ is non-singular for some n. The proof is the same as in the one-dimensional case.

Let $\Lambda_1(x)$, $\Lambda_2(x)$ be two independent linear functions on E_k. We adopt the same notation as in the above paragraph and assume further, for the sake of simplicity that the common distribution $F(x)$ is absolutely continuous with respect to Lebesgue measure. Then it is clear that the distribution of the two-dimensional random vector $\zeta_n = [\Lambda_1(\xi_n), \Lambda_2(\xi_n)]$ is also absolutely continuous. By central limit theorem the distribution of ζ_n converges weakly to a two-dimensional normal distribution, say $\Phi(\Lambda_1, \Lambda_2)$. Then we can deduce the following theorem from Theorem 1 and its multi-dimensional analogue.

Theorem 3: *Let $p_n(x/y)$ be the density of the conditional distribution of $\Lambda_1(\xi_n)$ given $\Lambda_2(\xi_n)$. Let $\phi(x/y)$ be the corresponding conditional density of the normal distribution $\Phi(\Lambda_1, \Lambda_2)$. Then*

$$p_n(x/y) \to \phi(x/y) \text{ almost everywhere in } x$$

for almost all (Lebesgue) y. In particular

$$\sup_A |P_n(A/y) - P(A/y)| \to 0, \text{ for almost all } y$$

where the supremum is taken over all Borel sets, and $P_n(A/y)$, $P(A/y)$ represent the corresponding conditional distributions.

The above theorem is an immediate consequence of Theorem 1, its multi-dimensional analogue and the fact that if $p_n(x, y)$, $q_n(y)$ denote the densities of ζ_n and $\Lambda_2(\xi_n)$ respectively, then $p_n(x/y) = p_n(x, y)/q_n(y)$. In this connection we might also refer to a paper of Steck (1957).

The almost sure convergence of Theorem 1 can be replaced by uniform convergence, if we impose further assumptions. This, as stated in the introduction has been done by Gnedenko (1954), the assumption required being that F_n be absolutely continuous and $p_n(x)$ be bounded for some n. He also obtains the estimate

$$\operatorname*{ess\,sup}_{-\infty < x < \infty} |p_n(x)-\phi(x)| = 0\left(\frac{1}{\sqrt{n}}\right)$$

if the third moment is finite. In the general case we can still derive estimates for the Lobesgue measure of the set $\{x : |p_n(x)-\phi(x)| > C\}$.

Theorem 4 : *Let $E|X|^3 < \infty$. Then there exist constants C_1, C_2 and λ (possibly depending on the distribution) such that*

$$L\left\{x : |p_n(x)-\phi(x)| > \frac{C_1}{\sqrt{n}}\right\} < C_2\lambda^n \qquad \ldots \quad (3.1)$$

where $0 < \lambda < 1$, L being Lebesgue measure.

Proof : With the same notation as in the proof of Theorem 1 we have

$$p_n(x) = q_n(x)+r_n(x)$$

where

$$\int_{-\infty}^{\infty} r_n(x)\,dx \leqslant k\,\alpha\,\beta^{k-1}+\beta^k. \qquad \ldots \quad (3.2)$$

Further by standard techniques it is easily shown that

$$\sup_x \; |q_n(x)-\phi(x)| \leqslant \frac{C}{\sqrt{n}} \qquad \ldots \quad (3.3)$$

where C is a constant independent of n. In view of this it is clear that

$$E = \left\{x : |p_n(x)-\phi(x)| > \frac{2C}{\sqrt{n}}\right\} \subset E_1 = \left\{x : r_n(x) > \frac{C}{\sqrt{n}}\right\}. \qquad \ldots \quad (3.4)$$

Thus from (3.2) and (3.4), we deduce

$$L(E) \leqslant L(E_1) \leqslant \frac{\sqrt{n}}{C}\,(k\,\alpha\,\beta^{k-1}+\beta^k). \qquad \ldots \quad (3.5)$$

Since k is determined by the relation $n = n_0\,k+r$ $(0 \leqslant r < n_0)$ it is easy to see that (3.1) is an immediate consequence of (3.5). This proves the theorem.

REFERENCES

BOCHNER, S. and CHANDRASEKHARAN, K. (1949): *Fourier Transforms*, Princeton University Press.

GNEDENKO, B. V. (1954): Local limit theorem for densities. *Doklady Akad. Nauk.* SSSR. **95**, 5–7.

KOLMOGOROV, A.N. and GNEDENKO, B. V. (1954): *Limit Distributions for Sums of Independent Random Variables*, Translated by K. L. Chung, Addison-Wesley.

PROHOROV, YU. V. (1952): On a local limit theorem for densities. *Doklady Akad. Nauk.* SSSR, **83**, 797–800.

SALEM, R. (1942): On sets of multiplicity for trigonometric series. *Amer. J. Math.*, **64**, 531–538.

SCHEFFÉ, H. (1947): A useful convergence theorem for probability distributions. *Ann. Math. Stat.*, **18**, 434–438.

SMITH, W. L. (1953): A frequency function form of the central limit theorem. *Proc. Camb. Phil. Soc.*, **49**, 462–472.

STECK, G. P. (1957): *Limit Theorems for Conditional Distributions*, University of California Publications in Statistics, **2**, Number 12.

Paper received : December, 1959.