

Convergence in probability (‘intermediate’) implies convergence in distribution (‘weak’). We quote this.

There is no converse, but there is a partial converse [which we shall use below]. If  $X_n$  converges in distribution to a *constant*  $c$ , then since the distribution function of the constant  $c$  is 0 to the left of  $c$  and 1 at  $c$  and to the right, it is easy to see that in fact  $X_n \rightarrow c$  in probability.

*Example.* We show by example that convergence in pr does not imply a.s. convergence (a fact known to F. Riesz in 1912). On the *Lebesgue measure space*  $[0, 1]$  (i.e.,  $([0, 1], \Lambda, \lambda)$ ), let

$$s_n := 1/2 + 1/3 + \dots + 1/n \pmod{1}, \quad A_n := [s_{n-1}, s_n], \quad X_n := I_{A_n}.$$

Since the harmonic series diverges, the  $X_n$  endlessly move rightwards through the interval  $[0, 1]$ , exiting right and reappearing left. So the  $X_n$  do not converge anywhere, in particular are not a.s. convergent. But since  $X_n = 0$  except on a set of probability  $1/n$ ,  $X_n \rightarrow 0$  in probability.

*Three classical convergence theorems.* We quote (see e.g. SP L6, 8):

M (Lebesgue’s monotone convergence theorem). If  $X_n \geq 0$ ,  $X_n \uparrow X$ , then  $E[X_n] \uparrow E[X]$ .

F (Fatou’s lemma). If  $X_n \geq 0$ , then  $E[\liminf X_n] \leq \liminf E[X_n]$ .

D (Lebesgue’s dominated convergence theorem). If  $X_n \rightarrow X$  a.s.,  $|X_n| \leq Y$  with  $E[Y] < \infty$ , then  $E[X_n] \rightarrow E[X]$ .

## 2. The Weak Law of Large Numbers (WLLN) and the Central Limit Theorem (CLT).

Recall that by Real Analysis,

$$\left(1 + \frac{x}{n}\right)^n \rightarrow e^x \quad (n \rightarrow \infty)$$

(this expresses compound interest, or exponential growth, as the limit of simple interest as the interest is compounded more and more often). This extends also to complex number  $z$ , and to  $z_n \rightarrow z$ :

$$\left(1 + \frac{z_n}{n}\right)^n \rightarrow e^z \quad (n \rightarrow \infty).$$

The next result is due to Lévy in 1925, and in more general form to the Russian probabilist A. Ya. KHINCHIN (1894-1956) in 1929 and to Kolmogorov in 1928/29.

**Theorem (Weak Law of Large Numbers, WLLN).** If  $X_i$  are iid with mean  $\mu$ ,

$$\frac{1}{n} \sum_{k=1}^n X_k \rightarrow \mu \quad (n \rightarrow \infty) \quad \text{in probability.}$$

*Proof.* If the  $X_k$  have CF  $\phi(t)$ , then as the mean  $\mu$  exists  $\phi(t) = 1 + i\mu t + o(t)$  as  $t \rightarrow 0$ . So  $(X_1 + \dots + X_n)/n$  has CF

$$E \exp\{it(X_1 + \dots + X_n)/n\} = [\phi(t/n)]^n = [1 + \frac{i\mu t}{n} + o(1/n)]^n,$$

for fixed  $t$  and  $n \rightarrow \infty$ . By above, the RHS has limit  $e^{i\mu t}$  as  $n \rightarrow \infty$ . But  $e^{i\mu t}$  is the CF of the constant  $\mu$ . So by Lévy's continuity theorem,

$$(X_1 + \dots + X_n)/n \rightarrow \mu \quad (n \rightarrow \infty) \quad \text{in distribution.}$$

Since the limit  $\mu$  is constant, by II.4 (L11), this gives

$$(X_1 + \dots + X_n)/n \rightarrow \mu \quad (n \rightarrow \infty) \quad \text{in probability.} //$$

As the name implies, the Weak LLN can be strengthened, to the Strong LLN (with a.s. convergence in place of convergence in probability). We turn to this later, but proceed with a refinement of the method above, in which we retain one more term in the Taylor expansion of the CF. Recall first that the CF of the standard normal distribution  $\Phi = N(0, 1)$ , with density  $\phi(x)$  and distribution function  $\Phi(x)$

$$\phi(x) := \frac{e^{-x^2/2}}{\sqrt{2\pi}}, \quad \Phi(x) := \int_{-\infty}^x \phi(u) du$$

is  $e^{-t^2/2}$ .

**Theorem (Central Limit Theorem, CLT).** If  $X_1, \dots, X_n, \dots$  are iid with mean  $\mu$  and variance  $\sigma^2$ , and  $S_n := X_1 + \dots + X_n$ , then

$$(S_n - n\mu)/(\sigma\sqrt{n}) \rightarrow \Phi = N(0, 1) \quad (n \rightarrow \infty) \quad \text{in distribution.}$$