Beginning of Oct. 21, 2022

Theorem: Convergence in distribution vs a.s.

If $F_n \to F$ in distribution, then there exist Y_n, Y with distribution functions F_n, F such that $Y_n \to Y$ almost surely.

Proof. The existence of Y_n, Y have been shown above. We need to only consider $\omega \in [0,1]$ for which $F^{-1}(\omega)$ contains 0 or 1 point. Fix ω and let $t = Y(\omega)$. Then

$$F^{-1}(\omega) = \varnothing \text{ or } \{t\}.$$

Therefore, for such points, for all $\delta > 0$,

$$F(t-\delta) < F(t) < F(t+\delta).$$

Choose δ such that $t \pm \delta$ are continuity points of F. Then, for large n, $F_n(t-\delta) < F(t) < F_n(t+\delta)$, so $t - \delta \leqslant Y_n(\omega) \leqslant t + \delta$, and similarly $t - \delta \leqslant Y(\omega) \leqslant t + \delta$. Since δ is arbitrary, $Y_n \to Y$ a.s., as there can only be countably many exceptions (countable jumps). □

Theorem: D3.2.9, Characterization of Weak Convergence

$X_n \to X$ in distribution iff $\mathbb{E}g(X_n) \to \mathbb{E}g(X)$ for all bounded continuous g.

Proof. Suppose $X_n \to X$ weakly. Take Y_n with the same distribution of X_n and Y similarly, with $Y_n \to Y$ almost surely. Let g be bounded and continuous. Then $g(Y_n) \to g(Y)$ a.s., so $\mathbb{E}g(Y_n) \to \mathbb{E}g(Y)$ by bounded convergence theorem.

Conversely, suppose $\mathbb{E}g(X_n) \to \mathbb{E}g(X)$ for all bounded continuous functions. We want

$$\mathbb{E}1_{(-\infty,x]}(X_n) \to \mathbb{E}1_{(-\infty,x]}(X)$$

for all continuity points x.

$1_{(-\infty,x]}$ isn't continuous, but it can be approximated by 1 on $(-\infty, x-\epsilon)$, 0 on (x,∞), and linear in between. We call this function $g_{x-\epsilon,x}$ and define $g_{x,x+\epsilon}$ similarly. By assumption,

$$\mathbb{E}g_{x-\epsilon,x}(X_n) \to \mathbb{E}g_{x-\epsilon,x}(X) \geqslant F(x-\epsilon)$$

and

$$\mathbb{E}g_{x,x+\epsilon}(X_n) \to \mathbb{E}g_{x,x+\epsilon}(X) \leqslant F(x+\epsilon).$$

Then since $F(x-\epsilon) \leqslant \liminf_{n\to\infty} F_n(x) \leqslant \limsup_{n\to\infty} F_n(x) \leqslant F(x+\epsilon)$, if x is a continuity point, we obtain the claim. □

Remark. Note that we can weaken the assumption and only require g to be continuous a.e.: denote the discontinuity set as D_g; if $\mathbb{P}(X \in D_g) = 0$ and $X_n \to X$ in distribution, then $\mathbb{E}g(X_n) \to \mathbb{E}g(X)$.

Corollary

If $X_n \to X$ in distribution and f is continuous, then $f(X_n) \to f(X)$ in distribution too.

Proof. If g is bounded, then $g \circ f$ is bounded, so $\mathbb{E}g(f(X_n)) \to \mathbb{E}g(f(X))$. Using the previous theorem once more, $f(X_n) \to f(X)$ in distribution. □

Corollary

If $X_n \to X$ almost surely, then $X_n \to X$ in distribution.

We have shown that there exists a metric w.r.t. convergence in probability: $|X - Y|/(1 + |X - Y|)$. There also exists metrics (one example is Lévy metric) for convergence in distribution.

Proposition: Convergence in probability ⇒ in distribution

Slick proof. It suffices to show that for all subsequence, there exists a further subsequence converging almost surely (then such sub-subsequence converges in distribution). And this is true as shown previously. Finally, since there is a metric for convergence in distribution, the full sequence indeed $\to X$ in distribution.

More revealing proof. Let g be bounded continuous, with $|g| \leqslant K$. By uniform continuity on compact sets, given M and ϵ, there exists δ satisfying the uniform continuity criterion on $[-M, M]$. Then

$$\begin{aligned}
|\mathbb{E}g(X_n) - \mathbb{E}g(X)| &= \int_\Omega |g(X_n) - g(X)| \, d\mathbb{P} \\
&\leqslant \int_{|X| \leqslant M, |X_n - X| < \delta} |g(X_n) - g(X)| \, d\mathbb{P} + \int_{|X| > M} |\ldots| \, d\mathbb{P} + \int_{|X_n - X| \geqslant \delta} \ldots \, d\mathbb{P} \\
&\leqslant \int_{|X| \leqslant M, |X_n - X| < \delta} \epsilon \, d\mathbb{P} + \int_{|X| > M} 2K \, d\mathbb{P} + \int_{|X_n - X| \geqslant \delta} 2K \, d\mathbb{P} \\
&\leqslant \epsilon + 2K\mathbb{P}(|X| > M) + 2K\mathbb{P}(|X_n - X| \geqslant \delta).
\end{aligned}$$

Hence,

$$\limsup_{n \to \infty} |\mathbb{E}g(X_n) - \mathbb{E}g(X)| \leqslant \epsilon + 2K\mathbb{P}(|X| > M) \qquad \text{for all } M, \epsilon.$$

Since M, ϵ are arbitrary, we see $\limsup |\mathbb{E}g(X_n) - \mathbb{E}g(X)| = 0$, as claimed. □

Remark: Converse is false. Let X_n, X be i.i.d. $\mathcal{N}(0, 1)$. Then clearly $X_n \to X$ in distribution, but not in probability.
However, (shown in HW), if $X_n \to c$ for some constant c, then indeed $X_n \to c$ in probability.