WEAK CONVERGENCE

Definition Let μ_n, μ be Borel probability measures on $\mathcal{S} = C[0,1]$ (or $C[0,\infty)$ or $\mathbb{R}^d$ or any completely and separably metrizable topological space). We say that $\mu_n \xrightarrow{d} \mu$ if any of the following equivalent conditions hold.

(1) For any *bounded continuous* $L : \mathcal{S} \to \mathbb{R}$, we have $\int L(s)d\mu_n(s) \to \int L(s)d\mu(s)$.
(2) For any open $G \subset \mathcal{S}$, we have $\limsup \mu_n(G) \le \mu(G)$.
(3) For any closed $F \subset \mathcal{S}$, we have $\liminf \mu_n(F) \ge \mu(F)$.
(4) For any Borel subset $A \subset \mathcal{S}$ with $\mu(\partial A) = 0$, we have $\mu_n(A) \to \mu(A)$.
(5) For any $L : \mathcal{S} \to \mathbb{R}$ for which $\mu\{s \in \mathcal{S} : L \text{ is not continuous at } s\} = 0$, we have $\int L(s)d\mu_n(s) \to \int L(s)d\mu(s)$.

The main ingredient in proving these equivalences is the *regularity* of Borel measures on such spaces. Regularity means $\mu(A) = \inf\{\mu(G) : \text{ open } G \supseteq A\} = \sup\{\mu(F) : \text{ closed } F \subseteq A\}$ for any Borel set A. A nice 'probabilistic' way of checking weak convergence (and one that we shall use) is the following.

Proposition Suppose we can construct a probability space $(\Omega, \mathcal{F}, P)$ and $\mathcal{S}$-valued random variables X_n having distribution μ_n and Y_n having distribution μ for each n (that is $PX_n^{-1} = \mu_n$ and $PY_n^{-1} = \mu$ for all n) and such that $X_n - Y_n \xrightarrow{P} 0$. Then, $\mu_n \xrightarrow{d} \mu$.

Proof Let $L : \mathcal{S} \to \mathbb{R}$ be a bounded continuous function. Then, $\int L(s)d\mu_n(s) - \int L(s)d\mu(s) = \mathbf{E}[L(X_n) - L(Y_n)]$ which goes to zero by the Dominated convergence theorem (since $|L(X_n) - L(X)|$ goes to zero in probability, and is bounded by the constant $2\sup|L(s)|$).

Remark Skorokhod proved the (stronger looking) converse. If $\mu_n \xrightarrow{d} \mu$, then one can always construct a probability space and random variables X_n, X such that $PX_n^{-1} = \mu_n$ and $PX^{-1} = \mu$ and such that $X_n \xrightarrow{a.s.} X$ w.r.t. the measure P. Sometimes we just write $X_n \xrightarrow{d} X$, but the statement is always about the distributions of X_n and of X and has nothing to do with the random variables themselves.

Example Let X_i be i.i.d. real-valued random variables with zero mean and unit variance and let $S_n = X_1 + \ldots + X_n$. The Central Limit Theorem says that $\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1)$. Here is how to prove it using Skorokhod embedding.

By Skorokhod embedding theorem, we can find B, a 1-dim BM and stopping times (perhaps w.r.t. an enhanced filtration) $0 = \tau_0 \le \tau_1 \le \tau_2 \le \ldots$ such that

$$(i)\ (\tau_k - \tau_{k-1}, B(\tau_k) - B(\tau_{k-1})) \text{ are i.i.d.} \qquad (ii)\ (B(\tau_k))_k \stackrel{d}{=} (S_k)_k \qquad (iii)\ \mathbf{E}[\tau_k - \tau_{k-1}] = 1.$$

Let $W_n(t) = \frac{B(nt)}{\sqrt{n}}$ so that $\frac{S_n}{\sqrt{n}} \stackrel{d}{=} \frac{B(\tau_n)}{\sqrt{n}} = W_n\left(\frac{\tau_n}{n}\right)$. By the proposition above, it suffices to show that $W_n\left(\frac{\tau_n}{n}\right) - W_n(1) \xrightarrow{P} 0$ (take $X_n = W_n\left(\frac{\tau_n}{n}\right) \stackrel{d}{=} \frac{S_n}{\sqrt{n}}$ and $Y_n = W_n(1) \stackrel{d}{=} N(0,1)$. Fix any $\varepsilon > 0$. Then for any $0 < \delta < 1$,

$$\mathbf{P}\left[\left|W_n\left(\frac{\tau_n}{n}\right) - W_n(1)\right| \ge \varepsilon\right] \le \mathbf{P}\left[\sup_{\substack{s,t\in[0,2]\\|s-t|\le\delta}} |W_n(t) - W_n(s)| > \varepsilon\right] \quad + \quad \mathbf{P}\left[\left|\frac{\tau_n}{n} - 1\right| > \delta\right].$$

The first term does not depend on n and goes to zero as $\delta \downarrow 0$. For fixed $\delta > 0$, by WLLN the second term goes to zero as $n \to \infty$. Thus, by first letting $n \to \infty$ and then $\delta \to 0$, we deduce that $|W_n\left(\frac{\tau_n}{n}\right) - W_n(1)| \xrightarrow{P} 0$.