Convergence Revisited

We've considered several modes of convergence of random variables:

$X_n \to X$ a.s. $\qquad P(\lim_{n\to\infty} X_n = X) = 1 \quad \leftarrow$ need $\mathcal{T} = \bigcap_{n\geq 1} \sigma(X_n, X_{n+1}, \dots)$

$X_n \to X$ in L^p $\qquad 0 = \lim_{n\to\infty} \|X_n - X\|_{L^p}^p = E[|X_n - X|^p] = \int_{\mathbb{R}^2} |y-x|^p \mu_{X_n,X}(dy\,dx)$

$X_n \to_p X$ $\qquad 0 = \lim_{n\to\infty} P(|X_n - X| > \varepsilon) = \int_{\mathbb{R}^2} \mathbb{1}_{|y-x|>\varepsilon}\, \mu_{X_n,X}(dy\,dx)$

All of these require information about the joint distribution of $\{X, X_n\}$.

We're now going to turn to some convergence notions that only care about the individual distributions.

Total Variation

Let $\{\mu_n\}_{n=1}^{\infty}$ be a sequence of probability measures on $(S, \mathcal{B})$.
They're just $\mathbb{R}$-valued functions on $\mathcal{B}$, so we can use any function convergence notion we like.

Def: Let μ, ν be probability measures on $(S, \mathcal{B})$. The **total variation distance** between them is

$$d_{TV}(\mu, \nu) = \sup_{B \in \mathcal{B}} |\mu(B) - \nu(B)|$$

If X, Y are $(S, \mathcal{B})$-valued random variables, we set

$$d_{TV}(X, Y) = d_{TV}(\mu_X, \mu_Y)$$

$$= \sup_{B \in \mathcal{B}} |\mathbb{P}(X \in B) - \mathbb{P}(Y \in B)|$$

Lemma: (Scheffé) If α is a finite measure on $(S, \mathcal{B})$ such that $\mu, \nu \ll \alpha$ with $d\mu = u\, d\alpha$, $d\nu = v\, d\alpha$, then

$$d_{TV}(\mu, \nu) = \frac{1}{2} \|u - v\|_{L^1(\alpha)}.$$

Pf. For $B \in \mathcal{B}$,

$$\left\{ \begin{array}{l} |\mu(B) - \nu(B)| = \left| \int_B u\, d\alpha - \int_B v\, d\alpha \right| \le \int_B |u-v|\, d\alpha \\ |\mu(B^c) - \nu(B^c)| \le \int_{B^c} |u-v|\, d\alpha \\ \quad \parallel \\ |(1-\mu(B)) - (1-\nu(B))| = |\mu(B) - \nu(B)| \end{array} \right.$$

$$2|\mu(B) - \nu(B)| \le \int_B |u-v|\, d\alpha + \int_{B^c} |u-v|\, d\alpha = \int_S |u-v|\, d\alpha = \|u-v\|_{L^1(\alpha)}$$

$A = \{u > v\}$

$$0 = \int u\, d\alpha - \int v\, d\alpha = \int_S (u-v)\, d\alpha = \int_A (u-v)\, d\alpha + \int_{A^c} (u-v)\, d\alpha$$

$$= \int_A |u-v|\, d\alpha - \int_{A^c} |u-v|\, d\alpha$$

$$|\mu(A) - \nu(A)| = \int_A |u-v|\, d\alpha = \int_{A^c} |u-v|\, d\alpha \quad \therefore \quad 2|\mu(A) - \nu(A)| = \int_S |u-v|\, d\alpha \quad ///$$

Note: it is always possible to find such an α.

In fact: if $\{\mu_n\}_{n=1}^{\infty}$ is any countable collection of finite measures,

take $\alpha = \sum_{n=1}^{\infty} 2^{-n} \mu_n$. $\mu_k \ll \alpha$

$\therefore$ by R-N thm, $d\mu_k = u_k \, d\alpha$ $u_k \geq 0$, $\int u_k \, d\alpha = 1$.

Cor: d_{TV} is a complete metric on $Prob(S, \mathcal{B})$.

Pf. (This can be shown directly, but the present approach is slicker.)

Let $\{\mu_n\}_{n=1}^{\infty} \subset Prob(S, \mathcal{B})$. Fix α as above; then $d\mu_n = u_n \, d\alpha$.

- $0 = d_{TV}(\mu_1, \mu_2) = \frac{1}{2} \int |u_1 - u_2| \, d\alpha \Rightarrow u_1 = u_2$ a.s. $[\alpha] \Rightarrow$ a.s. $[\mu_n] \Rightarrow \mu_1 = \mu_2$
- $d_{TV}(\mu_1, \mu_3) = \frac{1}{2} \|u_1 - u_3\|_{L^1} \leq \frac{1}{2}(\|u_1 - u_2\|_{L^1} + \|u_2 - u_3\|_{L^1}) = d_{TV}(\mu_1, \mu_2) + d_{TV}(\mu_2, \mu_3)$
- If $\{\mu_n\}_{n=1}^{\infty}$ is d_{TV}-Cauchy,

$\frac{1}{2}\|u_n - u_m\|_{L^1(\alpha)} = d_{TV}(\mu_n, \mu_m) \to 0$ $\therefore$ $\{u_n\}_{n=1}^{\infty}$ is a Cauchy in $L^1(\alpha)$

$\therefore u_n \to u \in L^1(\alpha)$ Define $d\mu := u \, d\alpha$.

$d_{TV}(\mu_n, \mu) = \frac{1}{2}\|u_n - u\|_{L^1(\alpha)} \to 0$.

$\underbrace{\left| \int u_n \, d\alpha - \int u \, d\alpha \right|}_{=1} \leq \int |u_n - u| \, d\alpha \to 0$

///

Cor: If h is a bounded r.v. on $(S, \mathcal{B})$, then $\forall \mu, \nu \in \text{Prob}(S, \mathcal{B})$

$$\left| \int_S h\,d\mu - \int_S h\,d\nu \right| \leq 2\, d_{TV}(\mu,\nu) \cdot \sup_S |h| .$$

Moreover, $d_{TV}(\mu,\nu) = \frac{1}{2} \sup\left\{ \left| \int_S h\,d\mu - \int_S h\,d\nu \right| : \sup_S |h| \leq 1 \right\}$

Pf. Fix finite meas. α s.t. $d\mu = u\,d\alpha$, $d\nu = v\,d\alpha$.

$$\left| \int_S h\,d\mu - \int_S h\,d\nu \right| = \left| \int_S (hu - hv)\,d\alpha \right| \leq \int_S |h|\,|u-v|\,d\alpha$$

$$\leq \sup_S |h| \cdot \underbrace{\int_S |u-v|\,d\alpha}_{2 d_{TV}(\mu,\nu)} .$$

$\therefore$ If $\sup_S |h| \leq 1$, $\left| \int_S h\,d\mu - \int_S h\,d\nu \right| \leq 2 d_{TV}(\mu,\nu)$

Define $h = \mathbb{1}_{u>v} - \mathbb{1}_{u\leq v}$ $\therefore |h| \leq 1$.

$$\int_S h\,d\mu - \int_S h\,d\nu = \int_S (\mathbb{1}_{u>v} - \mathbb{1}_{u\leq v})(u-v)\,d\alpha = \int_{u>v} |u-v|\,d\alpha + \int_{u\leq v} |u-v|\,d\alpha$$

$$= \int_S |u-v|\,d\alpha = 2 d_{TV}(\mu,\nu).$$

Total variation works well when S is countable.

Lemma: If S is countable, and $\mu, \nu \in Prob(S, \mathcal{B})$, then

$$d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{k \in S} |\mu(\{k\}) - \nu(\{k\})| \quad \checkmark$$

$$\therefore \mu_n \to_{TV} \mu \text{ iff } \mu_n(\{k\}) \to \mu(\{k\}) \quad \forall k \in S.$$

Pf. Fix α s.t. $d\mu = u\,d\alpha$, $d\nu = v\,d\alpha$

$\alpha_k = \alpha(\{k\})$

$\alpha = \sum_{k \in S} \alpha_k \delta_k$.

$$d_{TV}(\mu, \nu) = \frac{1}{2} \| u - v \|_{L^1(\alpha)}$$

$$= \frac{1}{2} \int |u - v| \, d\alpha$$

$$= \frac{1}{2} \sum_{k \in S} \alpha_k |u(k) - v(k)|$$

$$= \frac{1}{2} \sum_{k \in S} |\underbrace{u(k)\alpha_k}_{\mu(\{k\})} - \underbrace{v(k)\alpha_k}_{\nu(\{k\})}|.$$

$(\Rightarrow)$ $\sum_k |\mu_n(\{k\}) - \mu(\{k\})| \to 0$

$\Rightarrow \mu_n(\{k\}) \to \mu(\{k\})$. $\checkmark$

$(\Leftarrow)$ HW.

Eg. $\mu_p \overset{d}{=} \text{Bernoulli}(p)$.

$$d_{TV}(\mu_p, \mu_q) = \frac{1}{2}\sum_{k=0}^{1} |\mu_p\{k\} - \mu_q\{k\}| = \frac{1}{2}\left(|(1-p)-(1-q)| + |p-q|\right) = |p-q|$$

Eg. $\nu_\lambda \overset{d}{=} \text{Poisson}(\lambda)$. $\nu_\lambda(\{k\}) = e^{-\lambda}\lambda^k/k!$

$$d_{TV}(\nu_\lambda, \nu_\eta) = \frac{1}{2}\sum_{k=0}^{\infty} \left|e^{-\lambda}\frac{\lambda^k}{k!} - e^{-\eta}\frac{\eta^k}{k!}\right| \leq |\lambda - \eta| \quad (HW).$$

Eg.
$$d_{TV}(\mu_p, \nu_p) = \frac{1}{2}\sum_{k=0}^{\infty} |\mu_p(\{k\}) - \nu_p(\{k\})|$$
$$= \frac{1}{2}|\mu_p(0) - \nu_p(0)| + \frac{1}{2}|\mu_p(1) - \nu_p(1)| + \underbrace{\frac{1}{2}\sum_{k\geq 2}\nu_p(k)}_{\frac{1}{2}(1-\nu_p(0)-\nu_p(1))}$$
$$= \frac{1}{2}|1-p-e^{-p}| + \frac{1}{2}|p - e^{-p}p| + \frac{1}{2}(1-\nu_p(0)-\nu_p(1))$$
$$= \frac{1}{2}(e^{-p}+p-1) + \frac{1}{2}p - \frac{1}{2}e^{-p}p + \frac{1}{2} - \frac{1}{2}e^{-p} - \frac{1}{2}e^{-p}p$$
$$= p(1-e^{-p}).$$