In non-discrete settings, total variation convergence is usually too much to ask.

Eg. Let $a_n \in \mathbb{R}$, $a_n \to a$. Surely $\delta_{a_n} \to \delta_a$?

In general, if $\mu \perp \nu$, $d_{TV}(\mu, \nu)$

Eg. A discrete approximation of $\mathrm{Unif}([0,1])$:

$$\mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{k/n}$$

Why do we think $\frac{1}{n}\sum_{k=1}^{n}\delta_{k/n} \to \mathrm{Unif}([0,1])$ should be true?

Recall from last lecture: $\left|\int h\,d\mu_n - \int h\,d\nu\right| \leq 2\,d_{TV}(\mu_n,\mu)\cdot \sup|h|$

$\therefore\ d_{TV}(\mu_n,\mu) \to 0 \Rightarrow \int h\,d\mu_n \to \int h\,d\mu$

If $\mu_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{k/n}$, $\int h\,d\mu_n$

$\therefore$ If h is

Def: Let S be a metric space, $\mu_n, \mu \in \mathrm{Prob}(S,\mathcal{B}(S))$.
Say μ_n converges weakly to μ, $\mu_n \to_w \mu$ or $\mu_n \Rightarrow \mu$,
if $\int f\,d\mu_n \to \int f\,d\mu \quad \forall f \in C_b(S)$.

Eg. If $a_n \to a$, then $f(a_n) \to f(a)$ $\forall f \in C_b$

Prop: If $d_{TV}(\mu_n, \mu) \to 0$, then $\mu_n \to_w \mu$.

Pf.

Restricting to continuous test functions $f \in C_b$ allows μ_n to "smear out".

Notation: If X_n, X are S-valued random variables
say $X_n \to_w X$ iff $\mu_{X_n} \to_w \mu_{X_n}$.

Prop: If $X_n, X : (\Omega, \mathcal{F}, \mathbb{P}) \to (S, \mathcal{B}(S))$ and $X_n \to_{\mathbb{P}} X$, then $X_n \to_w X$.

First, a **Lemma**: If $X_n \to_{\mathbb{P}} X$ and $g \in C(S)$, then $g(X_n) \to_{\mathbb{P}} g(X)$.

Pf. For $\varepsilon, \delta > 0$, let $B_{\varepsilon\delta}(g) = \{ x \in S : \exists y \in S \ d(x,y) < \delta, \ |g(x) - g(y)| \geq \varepsilon \}$
Continuity of g means that, for fixed $\varepsilon > 0$, $B_{\varepsilon\delta}(g) \downarrow \emptyset$ as $\delta \downarrow 0$.

Prop: If $X_n, X : (\Omega, \mathcal{F}, \mathbb{P}) \to (S, \mathcal{B}(S))$ and $X_n \to_{\mathbb{P}} X$, then $X_n \to_w X$.

Pf: Let $f \in C_b(S)$. Then

Cor: If $X_n \to X$ a.s., or if $X_n \to X$ in L^p, then $X_n \to_w X$.

What was fundamentally wrong in the example $d_{TV}(\delta_{a_n}, \delta_a) \not\to 0$?
d_{TV} is too sensitive to jumps - to "discontinuity sets".

Def: Let μ be a Borel measure on a metric space S.
An event $B \in \mathcal{B}(S)$ is a continuity set for μ if

$$\mu(\partial A) = 0.$$

Eg. $(-\infty, a]$ is not a continuity set for δ_a

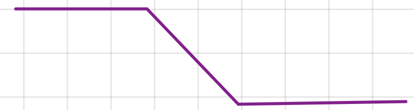
Eg. If $\mu \in \mathrm{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and F_μ is continuous (i.e. $\mu(\{a\}) = 0 \ \forall a \in \mathbb{R}$)
then all intervals are continuity sets for μ.

The Portmanteau Theorem

Let S be a complete, separable metric space.
Let $\mu_n, \mu \in \text{Prob}(S, \mathcal{B}(S))$. TFAE:

1. $\int f \, d\mu_n \to \int f \, d\mu \quad \forall f \in C_b(S)$. I.e. $\mu_n \to_w \mu$.
2. $\int f \, d\mu_n \to \int f \, d\mu \quad \forall f \in \text{Lip}_b(S)$
3. $\limsup_{n \to \infty} \mu_n(F) \leq \mu(F) \quad \forall$ closed $F \subseteq S$.
4. $\liminf_{n \to \infty} \mu_n(G) \geq \mu(G) \quad \forall$ open $G \subseteq S$.
5. $\mu_n(A) \to \mu(A) \quad \forall$ μ-continuity sets $A \in \mathcal{B}(S)$.

Pf. $(1 \Rightarrow 2)$: $\mathrm{Lip}_b(S) \subseteq C_b(S)$

$(2 \Rightarrow 3)$: Let $\psi =$ 

Fix a closed set F. Set $f_k(x) = \psi(k \cdot d(x, F))$

$|f_k(x) - f_k(y)|$

As $k \to \infty$: if $d(x,F) > 0$, $f_k(x) \downarrow \psi(\infty) = 0$

if $d(x,F) = 0$, $f_k(x) = \psi(0) = 1$.

$\therefore$ $f_k \downarrow \mathbb{1}_F$

Thus $\limsup_{n\to\infty} \mu_n(F) = \limsup_{n\to\infty} \int \mathbb{1}_F \, d\mu_n \le \limsup_{n\to\infty} \int f_k \, d\mu_n$

$(3 \Rightarrow 4)$: $1 - \liminf_{n\to\infty} \mu_n(G) = \limsup_{n\to\infty} (1 - \mu_n(G))$

$(3,4 \Rightarrow 5)$ If $\mu(\partial A) = 0$, $\therefore$ $\mu(A) = \mu(\bar{A}) = \mu(\mathrm{Int}(A))$,

$\therefore$ $\limsup_{n\to\infty} \mu_n(A)$

& $\liminf_{n\to\infty} \mu_n(A)$

It remains to prove (5 ⇒ 1): If $\mu_n(A) \to \mu(A)$ $\forall$ μ-continuous A,

then $\int f d\mu_n \to \int f d\mu$ $\forall f \in C_b$.

Let $f \in C_b(S)$; so $a \leq f(x) \leq b$ $\forall x \in S$ for some $a \leq b$ in $\mathbb{R}$.

For each $x \in S$, $f(x) - a = \lambda([a, f(x)])$

$$\therefore \int_S f \, d\mu_n = a + \int_S (f - a) \, d\mu_n$$

Similarly, $\int_S f \, d\mu =$

$\therefore$ Suffices to show that $\mu_n\{f \geq t\} \to \mu\{f \geq t\}$ for $[\lambda]$ a.e. $t \in \mathbb{R}$.

By assumption (5), this holds true except on $E = \{t \in [a,b] : \mu(\partial\{f \geq t\}) > 0\}$.