# Convergence

> *Since there are some problems about mathjax, I also upload the pdf version /Large Sample / 7 Convergence.*

## Basic Notation

- $M_n(\theta)=\frac{1}{n}\sum_{i=1}^{n} l_\theta\left(X_i\right)=P_n l_\theta$
- $M(\theta)=\mathbb{P} l_\theta$
- $\Delta_n(\theta)=\left\{M_n(\theta)-M(\theta)\right\}-\left\{M_n\left(\theta_0\right)-M\left(\theta_0\right)\right\}=\left(P_n-\mathbb{P}\right)\left(l_\theta-l_{\theta_0}\right)$
- $\mathcal{D}$: A metric space.

## Basic Definition

- **Modulus of Continuity** :
  - $f: T \rightarrow \mathbb{R}$, $w_f(\delta)=\sup_{\rho(s,t)<\delta} \mid f(s)-f(t)$
  - $\mathbb{W}_n(\delta)=\sup_{d\left(\theta, \theta_0\right) \leq \delta}\left|\Delta_n(\theta)\right|$
- **Weak Convergence** $X_n \stackrel{d}{\rightarrow} X$: $\mathbb{E}\left\{f\left(X_n\right)\right\} \rightarrow \mathbb{E}\{f(X)\}$ for all bounded and (Lipschitz) continuous function $f$;
- **Tightness**: $X: \Omega \rightarrow \mathcal{D}$ is tight if $\forall \epsilon>0, \exists K \subset \mathcal{D}$ compact s.t. $\mathbb{P}(X \in K)>1-\epsilon$ ;
- **Asymptotically Tight**: $\left\{X_n\right\}: \Omega \rightarrow \mathcal{D}$ is asymptotically tight if $\forall \epsilon>0, \exists K \subset \mathcal{D}$, s.t. $\limsup_{n \rightarrow \infty} \mathbb{P}\left(X_n \notin K^\delta\right)<\epsilon$ where $K^\delta=K+\delta B(0,1)$.
- $L^{\infty}(T)=\left\{f:\|f\|_{\infty}<\infty\right\}$ where $T$ is compact;
- $UC(T, \rho)=\{f: T \rightarrow \mathbb{R}$ is uniformly continuous$\}$;
- **Equicontinuous**: $\mathcal{F}=\{f\}$ is equicontinuous if $\lim_{\delta \downarrow 0} \sup_{f \in \mathcal{F}} w_f(\delta)=0$;
- **Asymptotically Equicontinuous**: $\left\{X_n\right\}$ is asymptotically equicontinuous if $\forall \epsilon, \eta>0, \exists$ finite partition of $T$, $\left\{T_1, \cdots, T_k\right\}$ s.t. $\limsup_{n \rightarrow \infty} \mathbb{P}\left(\max_{1 \leq i \leq k} \sup_{s, t \in T_i}\left|X_{n, s}-X_{n, t}\right| \geq \epsilon\right) \leq \eta$

# Theorems

## Prohorov Theorem

- If $X_n \xrightarrow{d} X$, $X$ is tight $\Leftrightarrow$ $X_n$ is asymptotically tight;
- If $X_n$ is asymptotically, then $\exists\{X_{n_k}\}\&X$ tight s.t. $X_{n_k} \xrightarrow{d} X$.

## Arzela-Ascoli Theorem

For compact metric space $(T, \rho)$

- $\mathcal{F} \subseteq UC(T, \rho)$ is relatively compact w.r.t supremum norm

  $\Leftrightarrow$
- $\mathcal{F}$ is uniformly equicontinuous and $\exists t_0 \in T$ s.t. $\sup_{f \in \mathcal{F}} |f(t_0)| < \infty$

## Weak Convergence in $L^\infty(T)$

For $X_n \in L^\infty(T)$, for $X \in UC(T, \rho)$ is tight we have $X_n \xrightarrow{d} X \Leftrightarrow$

- Finite Dimensional Convergence: $\forall k < \infty, t_1, \ldots, t_k \in T$

$$(X_{n,t_1}, \ldots, X_{n,t_k}) \xrightarrow{d} (X_{t_1}, \ldots, X_{t_k})$$

- $X_n$ is asymptotically equicontinuous.

# Donsker Class

$\mathcal{F} \subset L^\infty(\mathcal{F})$ is called $\mathbb{P}$ -**Donsker** if

$$(\sqrt{n}(P_n - \mathbb{P}) f_1, \cdots, \sqrt{n}(P_n - \mathbb{P}) f_k) \to \left(G_{f_1}, \ldots, G_{f_k}\right)$$

where $G$ is a gaussian process and $\operatorname{cov}\left(G_{f_i}, G_{f_k}\right) = \operatorname{cov}\{f_i(X), f_j(X)\}$, where $X \sim \mathbb{P}$.

## Some sufficient Conditions

- Let $\mathcal{F} = \{f : \mathcal{X} \to \mathbb{R}\}$ with an envelop function $F$ that $\mathbb{P}F^2 < \infty$ and

$$\int_0^\infty \sup_Q \sqrt{\log N\left(\|F\|_{L_2(Q)}\epsilon, \mathcal{F}, L_2(Q)\right)}\mathrm{d}\epsilon < \infty$$

Then $\mathcal{F}$ is a $\mathbb{P}$-Donsker.

- When $\mathcal{F}$ is a VC-subgraph-class of functions with envelop function $F$ that $\mathbb{P}F^2 < \infty$, then $\mathcal{F}$ is a $\mathbb{P}$-Donsker.

### Example (Kolmogorov-Smirnoff test)

Suppose $X_1, \ldots, X_n \overset{\text{i.i.d}}{\sim} F$ (continuous) and aim to test $\mathcal{H}_0 : F = F_0$ with test statistics $KS_n = \sqrt{n}\sup_{t\in\mathbb{R}} |F_n(t) - F_0(t)|$. Under $\mathcal{H}_0$,
$KS_n = \sqrt{n}\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| = \sup_{f\in\mathcal{F}} |\sqrt{n}\left(P_n - \mathbb{P}\right) f|$ where $\mathcal{F} = \{\mathbf{1}[\cdot \le t], t \in \mathbb{R}\}$.
Then, $\mathcal{F}$ is a Donsker class and then we can use gaussian process to conduct hypothesis test.

# Reference

---


- Convergence Rate and Convergence in Law