We focus on the emprical process:

$\mathbb{G}_n f = \sqrt{n}(P_n f - Pf)$, $f \in \mathcal{F}$.

Note that, for fixed $f \in \mathcal{F}$

$$\sqrt{n}(P_n f - Pf) \to_d N(0, Pf^2 - (Pf)^2)$$

And define

$$K(f,g) = \mathrm{Cov}(f(X), g(X)) = Pfg - Pf\,Pg$$

Let $\mathbb{G}_P$ be the zero mean gaussion process defined on $\mathcal{F}$. Then

$$\mathbb{G}_P : \omega \mapsto \mathbb{G}_P(\omega) \in \ell^\infty(\mathcal{F})$$

Then $\mathbb{G}_P$ is a tight random element of $\ell^\infty(\mathcal{F})$, and $\mathbb{G}_n$ is a random element.

Def If $\mathbb{G}_n \to_d \mathbb{G}_P$, then we call $\mathcal{F}$ is P-Donsker class.

Remark: In order to show $\mathbb{G}_n \to_d \mathbb{G}$, it's sufficient to show

$\forall \varepsilon, \eta > 0$, $\exists$ finite $\{T_i\}$ s.t. $\bigcup_i T_i = \mathcal{F}$

s.t.

$$\overline{\lim_n}\ \mathbb{P}^*(\sup_i \sup_{f,g \in T_i} |\mathbb{G}_n f - \mathbb{G}_n g| > \varepsilon) \leq \eta$$

For a random process $\{X_f\}_{f\in\mathcal{F}}$, we want to bound $\mathbb{E}\sup_{f,g\in\mathcal{F}}|X_f - X_g|$

① One step discretization

We assume that $\mathcal{F}$ is totally bounded, then $\forall f, g\in\mathcal{F}$

$\exists\ \mathcal{F}_n\subset\mathcal{F}$, $|\mathcal{F}_n|<\infty$, s.t. $\exists\ g_n, f_n\in\mathcal{F}_n$

$$|X_f - X_g| \le |X_f - X_{f_n}| + |X_g - X_{g_n}| + |X_{f_n} - X_{g_n}|$$

If we have $\|X_f - X_g\|_\psi \lesssim d(f,g)$, then $\forall\varepsilon>0$, $\exists\ \delta>0$, $\mathbb{E}\sup_{f,g\in\mathcal{F}}|X_f - X_g| \le \mathbb{E}\sup_{f,g\in\mathcal{F}_n}|X_f - X_g| + 2\mathbb{E}\sup_{d(f,g)\le\delta}|X_f - X_g| \le \mathbb{E}\sup_{f,g\in\mathcal{F}_n}|X_f - X_g| + \varepsilon$

② Chaining

Define a chain for $\forall f\in\mathcal{F}_n$,

Let $f_n = f$, $\varepsilon_n\le\varepsilon_{n-1}\le\cdots\le\varepsilon_0 \lesssim K = \mathrm{diam}(\mathcal{F})/2<\infty$

$\forall f_k$, let $\mathcal{F}_k$ be the ε_k-packing center of $\mathcal{F}_n$, and $\exists\ f_{k-1}\in\mathcal{F}_{k-1}$ s.t. $d(f_k, f_{k-1})\le\varepsilon_k$

Note that f_{k-1} is unique since if $\exists\ f'_{k-1}\in\mathcal{F}_{k-1}$ s.t. $d(f_k, f'_{k-1})\le\varepsilon_k$

$\Rightarrow d(f_{k-1}, f'_{k-1})\le 2\varepsilon_k$, contraction.

Then
$$\begin{array}{ccccccc} f_n & \rightarrow & f_{n-1} & \rightarrow & \cdots & \rightarrow & f_0 \\ \mathcal{F}_n & & \mathcal{F}_{n-1} & & & & \mathcal{F}_0 \end{array}$$

Dudley's Inequality

Lemma If $\{X_f\}_{f\in\mathcal{F}}$ is mean zero process s.t. $\|X_f - X_g\|_\psi \lesssim d(f,g)$, and we assume that $|\mathcal{F}| < \infty$, then

$$\mathbb{E}\sup_{f,g}|X_f - X_g| \lesssim \int_0^K \sqrt{\log N(\varepsilon,\mathcal{F},d)}\,d\varepsilon \triangleq J(\mathcal{F},d)$$

Pf: $|X_f - X_g| \le \sum_{1\le k\le n} |X_{f_k} - X_{k-1}| + \sum_{1\le k\le n} |X_{g_k} - X_{g_{k-1}}| + |X_{f_0} - X_{g_0}|$

$\Rightarrow \mathbb{E}\sup_{f,g\in\mathcal{F}}|X_f - X_g| \le 2\sum_{1\le k\le n}\mathbb{E}\sup_{(f,g)\in(\mathcal{F}_k,\mathcal{F}_{k-1})}|X_f - X_g|$

$+ \mathbb{E}\sup_{f,g\in\mathcal{F}_0}|X_f - X_g|$

(Claim) $\lesssim \sum_{1\le k\le n}\varepsilon_k\sqrt{\log P(\varepsilon_k,\mathcal{F},d)\,P(\varepsilon_{k-1},\mathcal{F},d)}$

$+ \varepsilon_0\sqrt{\log P(\varepsilon_0,\mathcal{F},d)}$, Let $\varepsilon_0 = K/2$, $\varepsilon_k = \frac{1}{2}\varepsilon_{k-1}$

$\lesssim \sum_{0\le k\le n}\varepsilon_k\sqrt{\log N(\varepsilon_k,\mathcal{F},d)} \asymp \int_0^K \sqrt{\log(\varepsilon,\mathcal{F},d)}\,d\varepsilon.$ $\square$

Remark:

Claim: $\mathbb{E}\max_{i\le n}|X_i| \lesssim \sqrt{\log n}\,\max_{i\le n}\|X_i\|_\psi$, $n>1$

Pf:

$$\mathbb{E}\max_i\frac{|X_i|}{\sqrt{\log(i+1)}} = \int_0^{+\infty}\mathbb{P}\left(\max_i\frac{|X_i|}{\sqrt{\log(i+1)}} > t\right)dt$$

$$\le \sum_i 2\int_0^{+\infty} e^{-t^2\log(i+1)/\|X_i\|_\psi^2}\,dt$$

$$= \sum_i \sqrt{2\pi/\log(i+1)}\cdot\|X_i\|_\psi \lesssim \max_i\|X_i\|_\psi$$

$$\Rightarrow \mathbb{E}\max_i|X_i|/\sqrt{\log n} \lesssim \max_i\|X_i\|_\psi. \quad \square$$

Lemma $\sqrt{n}P_n^{\sigma}f = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sigma_i f(X_i) \triangleq X_f$, then

$$\|X_f - X_g\|_{\psi} \lesssim d^{(2)}(f,g)$$

Pf:

$$\|X_f - X_g\|_{\psi} = \frac{1}{\sqrt{n}}\left\|\sum_{i=1}^{n}\sigma_i(f(X_i)-g(X_i))\right\|_{\psi}$$

By Hoeffding inequality:

$$\mathbb{P}\left(\frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n}\sigma_i(f(X_i)-g(X_i))\right| \geq t \,\middle|\, X_i = x_i\right)$$

$$\leq 2e^{-\frac{t^2}{2\frac{1}{n}\sum_{i=1}^{n}(f(x_i)-g(x_i))^2}}$$

$$\Rightarrow \mathbb{P}\left(\sum_{i=1}^{n}\sigma_i(f(X_i)-g(X_i)) \geq t\right) \leq 2e^{-\frac{t^2}{2P|f-g|^2}}$$

$$\Rightarrow \|X_f - X_g\|_{\psi} \lesssim d^{(2)}(f,g) \qquad \square$$

Cor $PF^2 < \infty$, $\mathbb{E}\sup_{f,g\in\mathcal{F}}\sqrt{n}|P_n^{\sigma}f - P_n^{\sigma}g| \lesssim J(\mathcal{F}, d^{(2)})$

Donsker theorem

If $J(\mathcal{F}, d^{(2)}) < \infty$, $PF^2 < \infty$, then $\mathcal{F}$ is P-Donsker class.

Pf: $\forall \varepsilon, \eta > 0$,

$$\mathbb{P}(\sup_i \sup_{f,g \in T_i} |\mathbb{G}_n f - \mathbb{G}_n g| > \varepsilon)$$
$$\leq \mathbb{E} \sup_i \sup_{f,g\in T_i} |\mathbb{G}_n f - \mathbb{G}_n g| / \varepsilon$$

(Let $\delta = \max_i \operatorname{diam}(T_i)$, then)

$$\lesssim \mathbb{E} \sup_{\{d(f,g)\leq \delta\}} |\mathbb{G}_n f - \mathbb{G}_n g| / \varepsilon$$
$$\lesssim \mathbb{E} \sup_{\{d(f,g)\leq\delta\}} \sqrt{n}\, |P_n^0 f - P_n^0 g| / \varepsilon$$
$$\lesssim J(A(\delta), d)/\varepsilon, \quad A(\delta) = \{f, g \in \mathcal{F};\ d(f,g) < \delta\}$$

Let δ small enough, then $J(A(\delta), d) < \eta$ $\square$

Example If $\mathcal{F}$ is VC-class, then $\mathcal{F}$ is P-donsker class, $\forall P$.

Remark: Note that $N(\varepsilon, \mathcal{F}, d^{(2)}) \lesssim (\frac{1}{\varepsilon})^{Cv}$, $v = VCD) < \infty$, and $(PF^2)^{\frac{1}{2}} < \infty$

$$J(\mathcal{F}, d) \leq \int_0^{(PF^2)^{\frac{1}{2}}} \sqrt{Cv \log \frac{1}{\varepsilon}}\, d\varepsilon \lesssim \sqrt{v}$$

Example $\mathcal{F} = \{ \mathbb{1}_{(-\infty, t]}(x),\ t \in \mathbb{R} \}$, $K(t,s) = F(\min\{t,s\}) - F(t)F(s)$. Then $\sqrt{n}(\hat{F}_n - F) \to_d W_P$, $\operatorname{Cov}(W_P(t), W_P(s)) = K(t,s)$

Brownian Bridge

Define $B_t = W_P(F^{-1}(t))$, $t \in [0,1]$ is called the Brownian Bridge and $K_B(s,t) = \min\{s,t\} - st$.

Define $\sqrt{n}\,\|\hat{F}_n - F\|_\infty$, $\sqrt{n}\int (\hat{F}_n - F)^2 dF$, then

$$\sqrt{n}\,\|\hat{F}_n - F\|_\infty \to_d \|W_P\|_\infty$$

$$\sqrt{n}\int (\hat{F} - F)^2 dF \to_d \int W_P^2(t)\, dF(t)$$

And $\|W_P\|_\infty = \sup_{t\in\mathbb{R}} |W_P(t)|$

$$= \sup_{u\in[0,1]} |W_P(F^{-1}(u))|$$
$$= \sup_{u\in[0,1]} |B_u|$$

$$\int W_P^2(t)\, dF(t) = \int B_t^2\, dt$$

Example $\mathcal{F} = \{\psi_\theta ; \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^d$. $\exists L$ s.t. $|\psi_{\theta_1}(x) - \psi_{\theta_2}(x)| \le L(x)\|\theta_1 - \theta_2\|$ $\forall x$.
and $PL^2 < \infty$

Note that $d^{(2)}(\psi_{\theta_1}, \psi_{\theta_2}) \le (PL^2)^{\frac{1}{2}} \|\theta_1 - \theta_2\|$
$\lesssim \|\theta_1 - \theta_2\|$

And $\mathrm{Vol}(\Theta) \le \dfrac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)} r^d$, among $r = \mathrm{diam}(\Theta)/2$

$\Rightarrow N(\varepsilon, \mathcal{F}, d^{(2)}) \lesssim (\frac{1}{\varepsilon})^d$

Theorem $PF^2 < \infty$, $\hat{f}_n, f_0 \in \mathcal{F}$ and $\hat{f}_n$ is a random function dependent on $X_1, \cdots, X_n$, If $d^{(2)}(\hat{f}_n, f) = \int (\hat{f}_n(x) - f(x))^2 d(F(x)) \to_p 0$ and $\mathcal{F}$ is P-D, $G_n \hat{f}_n - G_p f = o_p(1)$

Pf: $\forall \varepsilon > 0$,

$$
\begin{aligned}
&\mathbb{P}(|G_n \hat{f}_n - G_p f| > \varepsilon) \\
&\le \mathbb{P}(|G_n \hat{f}_n - G_p f| > \varepsilon, d^{(2)}(\hat{f}_n, f_0) \le \delta) + o(1) \\
&\le \mathbb{P}(\sup_{d^{(2)}(f,g) < \delta} |G_n f - G_p g| > \varepsilon) + o(1) \\
&\le \mathbb{P}(\sup_{d^{(2)}(f,g) \le \delta} |G_n f - G_p f| + |G_p f - G_p g| \\
&\quad > \varepsilon) + o(1) \\
&\le \mathbb{P}(\|G_n - G_p\|_{\mathcal{F}} \ge \frac{\varepsilon}{2}, \sup_{d^{(2)}(f,g) \le \delta} |G_p(f-g)| \ge \frac{\varepsilon}{2}) \\
&\quad + o(1) \\
&= \mathbb{P}(\sup_{d(f,g) \le \delta} G_p(f-g) \ge \frac{\varepsilon}{2}) + o(1)
\end{aligned}
$$

Note that $\|G_p f - G_p g\|_4^2 = \mathrm{Var}(G_p f - G_p g)$

$$
\begin{aligned}
&= K(f,f) + K(g,g) - 2K(f,g) \\
&= Pf^2 - |Pf|^2 + Pg^2 - (Pg)^2 - 2Pfg - 2PfPg \\
&= P(f-g)^2 - (Pf - Pg)^2 \le P|f-g|^2
\end{aligned}
$$

$\Rightarrow$ $\mathbb{E} \sup_{d^{(2)}(f,g) \le \delta} G_p f - G_p g \lesssim J(A_\delta, d^{(2)})$

$A_\delta = \{ f, g \in \mathcal{F} ; d^{(2)}(f,g) \le \delta \}$, Let $\delta \to 0$, $J(A_\delta, d^{(2)}) \to 0$, means that $\forall \varepsilon_0 > 0$

$\mathbb{P}(|G_n \hat{f}_n - G_p f| > \varepsilon) < \varepsilon_0$. $\square$

Example (Sobolev class)

$W^{k,p}(\mathcal{X}) = \{ f ; D^{(k)} f$ exists and $\|D^{(k)} f\|_p < \infty \}$, we add a convex constraint $\|D f^{(i)}\|_p \leq C_i$ and assume that $\mathcal{X} \subset \mathbb{R}^d$ is open, convex and bound

Remark: Since $\mathcal{X}$ is totally bounded, then $\exists$ ε-covering center: $x_1, \dots, x_m$, $m \lesssim \frac{1}{\varepsilon^d}$

Note that $\forall x \in \mathcal{X}$, $f, g \in W^{k,p}(\mathcal{X})$

$$f(x) - g(x) = \sum_{i \leq k-1} D^{(i)}(f-g)(x_j) \frac{(x-x_j)^{(i)}}{(i)!} + \sum_{i=k} D^{(i)}(f-g)(\tilde{x}_{(i)j}) \frac{(x-x_j)^{(i)}}{(i)!}$$

, then $\forall x \in \mathcal{X}$, $\exists x_j$ s.t.

$$|f(x) - g(x)| \lesssim \sum_{i<k} |D^{(i)}(f-g)(x_j)| \frac{\varepsilon^i}{(i)!} + \varepsilon^k$$

$$\Rightarrow \|f-g\|_\infty \lesssim \sup_j \sum_{i<k} |D^{(i)}(f-g)(x_j)| \frac{\varepsilon^i}{(i)!} + \varepsilon^k$$

$$\lesssim \sup_j \sum_{i<k} |A_{i,j}(f) - A_{i,j}(g)| \varepsilon^i + \varepsilon^k \quad ①$$

$A_{i,j}(f) = D^{(i)} f(x_j)$, $Af = (A_{i,j}(f))$

$\mathcal{A} = \{ Af, f \in W^{k,p}(\mathcal{X}) \}$

In order to let $① < \delta$,

$$A_{i,j}(f-g) \lesssim \varepsilon^{k-i}, \text{ and } \varepsilon^k = \delta$$

which means that

$$N(\delta, W^{k,p}(\mathcal{X}), \|\cdot\|_\infty) \lesssim \prod_{(i)=0}^{(k)} \left(\frac{2C_i}{\varepsilon^{k-i}} + 1\right)^{\frac{c}{\varepsilon^d}}$$

$$= e^{\sum_{(i)=0}^{(k)} \frac{c}{\varepsilon^d} \ln\left(\frac{C_i}{\varepsilon^{k-i}} + 1\right)} \lesssim e^{\frac{c}{\varepsilon^d}} \asymp \left(\frac{1}{\delta}\right)^{\frac{d}{k}} \quad \square$$