STAT 620: Asymptotic Statistics **Spring 2022**

Lecture: Mar 29

Lecturer: Xianyang Zhang

1 Stochastic process

Let $\{X_t\}_{t\in\mathcal{T}}$ be a collection of real valued random variables. This is a stochastic process indexed by $\mathcal{T}$.

2 Sub-Gaussian process

Let $(\mathcal{T},\rho)$ be a metric space. We say $\{X_t\}_{t\in\mathcal{T}}$ is a sub-Gaussian process if:

$$\mathbb{E}[\exp(\lambda(X_s - X_t))] \le \exp\left(\frac{\lambda^2\rho(s,t)^2}{2}\right) \quad \forall\lambda > 0, s, t \in \mathcal{T}.$$

2.1 Example 1

A Gaussian process is an example of a sub-Gaussian process. To see this, let $\mathcal{T} = \mathbb{R}^d$ and $Z \sim N(0, \sigma^2 I_d)$. Define $X_t = t^\top Z$. Note that $X_t - X_s = (t-s)^\top Z$ has a normal distribution with mean zero and variance $\|t-s\|^2\sigma^2$. Therefore, we have

$$E[e^{\lambda(X_t - X_s)}] \le e^{\lambda^2\sigma^2\|t-s\|^2/2}.$$

2.2 Example 2

Let T be a vector space equipped with a norm $\|\cdot\|$, and X_i be some random variables taking values in $\mathcal{X}$. Suppose $l : T \times \mathcal{X} \to \mathbb{R}$ is Lipschitz in its first argument, that is

$$|l(t,x) - l(s,x)| \le \|t-s\|$$

for all $x \in \mathcal{X}$ and $s, t \in T$. Then for a sequence of i.i.d Rademacher random variables ϵ_i, as $\epsilon_i(l(t,X_i) - l(s,X_i))$ is bounded between $-\|t-s\|$ and $\|t-s\|$, we have

$$\begin{aligned}
E\left[\exp\left(\lambda\sum_{i=1}^n \epsilon_i(l(t,X_i) - l(s,X_i))\right)\right] =& E\left[E\left[\exp\left(\lambda\sum_{i=1}^n \epsilon_i(l(t,X_i) - l(s,X_i))\right)\middle| X_1,\dots,X_n\right]\right] \\
\le& E\left[E\left[\exp\left(\frac{\lambda^2}{2}\sum_{i=1}^n (l(t,X_i) - l(s,X_i))^2\right)\middle| X_1,\dots,X_n\right]\right] \\
\le& \exp\left(\frac{\lambda^2 n\|t-s\|^2}{2}\right).
\end{aligned}$$

Therefore, $Z_t = \sum_{i=1}^n \epsilon_i l(t,X_i)$ is a sub-Gaussian process with $\rho(s,t) = \sqrt{n}\|s-t\|$.

3 Dudley's integral entropy

Let $\{X_t\}_{t\in\mathcal{T}}$ be a ρ-sub-Gaussian separable and mean-zero process. Our goal is to derive an upper bound for the quantity

$$E[\sup_{t\in\mathcal{T}} X_t].$$

The key technique we shall use is the chaining argument.

3.1 Chaining argument

Let $\epsilon_k = 2^{-k}D$, where $D = \sup_{s,t\in\mathcal{T}} \rho(s,t)$ is the diameter of $\mathcal{T}$. Let $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \ldots$ be a sequence of covers of $\mathcal{T}$, where $\mathcal{T}_k$ is an ϵ_k-net of $\mathcal{T}$ and $\mathcal{T}_0 = \{t_0\}$ for some $t_0 \in \mathcal{T}$.

Define $\pi_k : \mathcal{T} \to \mathcal{T}_k$ the projetion of the points in $\mathcal{T}$ onto $\mathcal{T}_k$ so that

$$\rho(r, \pi_k(r)) \leq \epsilon_k$$

for all k and $r \in \mathcal{T}$.

For any $t \in \mathcal{T}_k$, we note that $\pi_k(t) = t$ and $\pi_0(t) = t_0$. Consider the following decomposition

$$X_t - X_{t_0} = \sum_{i=1}^{k} \left(X_{\pi_i \circ \cdots \circ \pi_k(t)} - X_{\pi_{i-1} \circ \pi_i \cdots \circ \pi_k(t)} \right)$$

for any $t \in \mathcal{T}_k$. Notice that

$$\max_{t\in\mathcal{T}_k} (X_t - X_{t_0}) \leq \sum_{i=1}^{k} \max_{t\in\mathcal{T}_k} \left(X_{\pi_i \circ \cdots \circ \pi_k(t)} - X_{\pi_{i-1} \circ \pi_i \cdots \circ \pi_k(t)} \right) \leq \sum_{i=1}^{k} \max_{t\in\mathcal{T}_i} \left(X_t - X_{\pi_{i-1}(t)} \right).$$

Here $\max_{t\in\mathcal{T}_i} \left(X_t - X_{\pi_{i-1}(t)} \right)$ is a finite maximum of $(2^{1-i}D)^2$-Sub-Gaussian variables. because $\rho(t, \pi_{i-1}(t)) \leq 2^{1-i}D$. Recall that for N σ^2-sub-Gaussian variables, we have

$$E[\max_{1\leq i\leq N} Y_i] \leq \sqrt{2\sigma^2 \log N}.$$

Thus we have

$$E[\max_{t\in\mathcal{T}_i} \left(X_t - X_{\pi_{i-1}(t)} \right)] \leq \sqrt{24^{1-i} D^2 \log |\mathcal{T}_i|}$$

where $|\mathcal{T}_i|$ denotes the cardinality of the set $\mathcal{T}_i$. Therefore, we get

$$
\begin{aligned}
E[\max_{t\in\mathcal{T}_k}(X_t - X_{t_0})] \leq & \sum_{i=1}^{k} E[\max_{t\in\mathcal{T}_i}\left(X_t - X_{\pi_{i-1}(t)}\right)] \\
\leq & \sum_{i=1}^{k} \sqrt{24^{1-i}D^2 \log|\mathcal{T}_i|} \\
= & \sum_{i=1}^{k} \sqrt{2}2D2^{-i}\sqrt{\log N(\mathcal{T},\rho,D2^{-i})} \\
\leq & 4\sqrt{2}\sum_{i=1}^{k}\int_{D2^{-(i+1)}}^{D2^{-i}} \sqrt{\log N(\mathcal{T},\rho,u)}du \\
= & 4\sqrt{2}\int_{D2^{-(k+1)}}^{D/2} \sqrt{\log N(\mathcal{T},\rho,u)}du \\
\leq & 4\sqrt{2}\int_{0}^{D/2} \sqrt{\log N(\mathcal{T},\rho,u)}du.
\end{aligned}
$$

We further get

$$
\begin{aligned}
E[\sup_{t\in\mathcal{T}} X_t] = & E[\liminf_{k} \sup_{t\in\mathcal{T}_k\cup\mathcal{T}_0} (X_t - X_{t_0}) + X_{t_0}] \\
\leq & \liminf_{k} E[\sup_{t\in\mathcal{T}_k\cup\mathcal{T}_0} (X_t - X_{t_0})] \\
\leq & 4\sqrt{2}\int_{0}^{D/2} \sqrt{\log N(\mathcal{T},\rho,u)}du.
\end{aligned}
$$

where the first equality follows the separability and the first inequality follows from Fatou's Lemma. To sum up, we obtain

$$
E[\sup_{t\in\mathcal{T}} X_t] \leq 4\sqrt{2}\int_{0}^{D/2} \sqrt{\log N(\mathcal{T},\rho,u)}du.
$$