MATH524 – Spring 2025

Problem Set: Week 6

1. (**Subgaussian Chaining and Dudley's Entropy Integral**)

Let $(\mathcal{T}, d)$ be a non-empty metric space and X_t be real-valued random variables for $t \in \mathcal{T}$. Suppose $\mathbb{E}[X_t] = 0$ for all $t \in \mathcal{T}$ and assume the subgaussian increments condition

$$\mathbb{E}\big[\exp\big(\lambda(X_t - X_s)\big)\big] \le \exp\left(\frac{\lambda^2 d(s,t)^2}{2}\right)$$

for all $s, t \in \mathcal{T}$. Assume that there is a countable subset $\mathcal{T}' \subseteq \mathcal{T}$ with $\sup_{t\in\mathcal{T}} X_t = \sup_{t'\in\mathcal{T}'} X_{t'}$ almost surely.

(a) Show that if $t_1, \ldots, t_N \in \mathcal{T}$ and $s_1, \ldots, s_N \in \mathcal{T}$ satisfy $\max_{1\le j\le N} d(s_j, t_j) \le \delta$ then

$$\mathbb{E}\left[\max_{1\le j\le N}\left(X_{t_j} - X_{s_j}\right)\right] \le \delta\sqrt{2\log N}.$$

Hint: *you may wish to show first that if $\lambda > 0$ then $\max_{1\le j\le N} x_j \le \frac{1}{\lambda}\log\sum_{j=1}^N \exp(\lambda x_j)$ for any $x_j \in \mathbb{R}$.*

Solution: The hint can be proven directly as

$$\max_{1\le j\le N} x_j = \frac{1}{\lambda}\log\exp\left(\lambda \max_{1\le j\le N} x_j\right) = \frac{1}{\lambda}\log \max_{1\le j\le N}\exp(\lambda x_j) \le \frac{1}{\lambda}\log\sum_{j=1}^N \exp(\lambda x_j).$$

So by Jensen's inequality, assuming $\delta > 0$ and $N \ge 2$,

$$\begin{aligned}\mathbb{E}\left[\max_{1\le j\le N}\left(X_{t_j} - X_{s_j}\right)\right] &\le \mathbb{E}\left[\frac{1}{\lambda}\log\sum_{j=1}^N \exp\Big(\lambda\big(X_{t_j} - X_{s_j}\big)\Big)\right] \le \frac{1}{\lambda}\log\sum_{j=1}^N \mathbb{E}\left[\exp\Big(\lambda\big(X_{t_j} - X_{s_j}\big)\Big)\right] \\ &\le \frac{1}{\lambda}\log\left(N\exp\left(\frac{\lambda^2\delta^2}{2}\right)\right) = \frac{\log N}{\lambda} + \frac{\lambda\delta^2}{2} \le \delta\sqrt{2\log N},\end{aligned}$$

where the last inequality follows by setting $\lambda = \frac{\sqrt{2\log N}}{\delta}$. Now if $\delta = 0$ or $N = 1$ then it is easy to see that $X_{t_j} - X_{s_j} = X_{t_{j'}} - X_{s_{j'}}$ almost surely for all $1 \le j, j' \le N$ and the result holds as $\mathbb{E}[X_{t_j} - X_{s_j}] = 0$.

(b) For $\varepsilon > 0$ we say a subset $\mathcal{T}_\varepsilon \subseteq \mathcal{T}$ is an ε-cover of $(\mathcal{T}, d)$ if $\sup_{t\in\mathcal{T}}\inf_{s\in\mathcal{T}_\varepsilon} d(s,t) \le \varepsilon$. The covering number $N(\varepsilon, \mathcal{T}, d)$ is the cardinality of the smallest ε-cover. Show that if $\varepsilon \le \varepsilon'$ then $N(\varepsilon, \mathcal{T}, d) \ge N(\varepsilon', \mathcal{T}, d)$.

Solution: This follows because every ε-cover is an ε'-cover, so the smallest ε-cover must be at least as large as the smallest ε'-cover.

(c) Define the diameter $D = \sup_{s,t\in\mathcal{T}} d(s,t)$. Show that if $D = \infty$ then $N(\varepsilon, \mathcal{T}, d) = \infty$ for all $\varepsilon > 0$.

Hint: *You may choose to prove this result by contradiction.*

Solution: Suppose for contradiction that $N(\varepsilon, \mathcal{T}, d) = N < \infty$ for some $\varepsilon > 0$. Let $\mathcal{T}_N = \{t_1, \ldots, t_N\}$ be a ε-cover and take $s, t \in \mathcal{T}$. Without loss of generality we can order the points so that $d(s, t_1) \le \varepsilon$ and $d(t, t_N) \le \varepsilon$. Then by the triangle inequality, $d(s,t) \le d(s,t_1) + \sum_{j=1}^{N-1} d(t_j, t_{j+1}) + d(t, t_N) \le 2\varepsilon + (N-1)\max_{1\le j,j'\le N-1} d(t_j, t_{j'})$. But then taking a supremum over $s, t \in \mathcal{T}$ shows that $D < \infty$.

For the remainder of this question we will assume $D < \infty$.

(d) For $k \geq 0$ let $\mathcal{T}_k$ be a $(2^{-k}D)$-cover of $(\mathcal{T}, d)$ with cardinality $N_k = N\big(2^{-k}D, \mathcal{T}, d\big)$ and for all $t \in \mathcal{T}$ let $\pi_k(t) \in \mathcal{T}_k$ satisfy $d(t, \pi_k(t)) \leq 2^{-k}D$. Using the result from part (a), show that $d(\pi_{k+1}(t), \pi_k(t)) \leq 3 \cdot 2^{-k-1}D$ and deduce that

$$\mathbb{E}\left[\sup_{t\in\mathcal{T}}\big(X_{\pi_{k+1}(t)} - X_{\pi_k(t)}\big)\right] \leq 3 \cdot 2^{-k}D\sqrt{\log N_{k+1}}.$$

Solution: By definition of π_k we have $d(\pi_{k+1}(t), \pi_k(t)) \leq d(\pi_{k+1}(t), t) + d(t, \pi_k(t)) \leq 2^{-k-1}D + 2^{-k}D = 3 \cdot 2^{-k-1}D$. Now apply part 1a to $X_{\pi_{k+1}(t)} - X_{\pi_k(t)}$. There are at most $N_k N_{k+1}$ such variables so

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in\mathcal{T}}\big(X_{\pi_{k+1}(t)} - X_{\pi_k(t)}\big)\right] &\leq 3 \cdot 2^{-k-1}D\sqrt{2\log(N_k N_{k+1})} \leq 3 \cdot 2^{-k-1}D\sqrt{2\log(N_{k+1}^2)} \\ &\leq 3 \cdot 2^{-k}D\sqrt{\log N_{k+1}}.\end{aligned}$$

(e) Show that $N_0 = 1$ and by considering the chain $t_{\pi_0(t)} \to t_{\pi_1(t)} \to \cdots \to t$ show that

$$\mathbb{E}\left[\sup_{t\in\mathcal{T}} X_t\right] \leq 6\sum_{k=1}^{\infty} 2^{-k}D\sqrt{\log N_k}.$$

Hint: *first assume that $\mathcal{T}$ is finite, then extend to the countable setting, and finally to a general set $\mathcal{T}$.*

Solution: $N_0 = 1$ because any singleton subset of $\mathcal{T}$ is a D-cover. Suppose that $\mathcal{T}$ is finite and take $\mathcal{T}_0 = \{t_0\}$. Let $2^{-K}D < \inf\{d(s,t) : d(s,t) \neq 0\}$ so that $\mathcal{T}_K = \mathcal{T}$ and note that since $\mathbb{E}[X_{t_0}] = 0$,

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in\mathcal{T}} X_t\right] &= \mathbb{E}\left[\sup_{t\in\mathcal{T}}\big(X_t - X_{t_0}\big)\right] = \mathbb{E}\left[\sup_{t\in\mathcal{T}}\big(X_{\pi_K(t)} - X_{\pi_0(t)}\big)\right] = \mathbb{E}\left[\sup_{t\in\mathcal{T}}\sum_{k=0}^{K-1}\big(X_{\pi_{k+1}(t)} - X_{\pi_k(t)}\big)\right] \\ &\leq \sum_{k=0}^{K-1}\mathbb{E}\left[\sup_{t\in\mathcal{T}}\big(X_{\pi_{k+1}(t)} - X_{\pi_k(t)}\big)\right] \leq 3\sum_{k=0}^{K-1} 2^{-k}D\sqrt{\log N_{k+1}} \leq 6\sum_{k=1}^{\infty} 2^{-k}D\sqrt{\log N_k}.\end{aligned}$$

(f) By considering Riemann sums, prove Dudley's entropy integral:

$$\mathbb{E}\left[\sup_{t\in\mathcal{T}} X_t\right] \leq 12\int_0^{\infty}\sqrt{\log N(\varepsilon, \mathcal{T}, d)}\,\mathrm{d}\varepsilon.$$

Solution:

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in\mathcal{T}} X_t\right] &\leq 6\sum_{k=1}^{\infty} 2^{-k}D\sqrt{\log N(2^{-k}D, \mathcal{T}, d)} = 12\sum_{k=1}^{\infty}\int_{2^{-k-1}D}^{2^{-k}D}\sqrt{\log N(2^{-k}D, \mathcal{T}, d)}\,\mathrm{d}\varepsilon \\ &\leq 12\int_0^{\infty}\sqrt{\log N(\varepsilon, \mathcal{T}, d)}\,\mathrm{d}\varepsilon.\end{aligned}$$

Since this does not depend on the cardinality of $\mathcal{T}$, it must also hold if $\mathcal{T}$ is countable by the monotone convergence theorem. By separability it holds for general $\mathcal{T}$ since $\sup_{t\in\mathcal{T}} X_t = \sup_{t'\in\mathcal{T}'} X_{t'}$ almost surely for some countable $\mathcal{T}'$.