In Lecture 15, we proved the following *Generalized VC inequality*

$$\mathbb{P}\left(\forall f \in \mathcal{F},\ \mathbb{E}f - \frac{1}{n}\sum_{i=1}^{n} f(x_i) \le \frac{2^{9/2}}{\sqrt{n}}\mathbb{E}_{x'}\int_0^{d(0,f)} \log^{1/2}\mathcal{D}(\mathcal{F},\varepsilon,d)d\varepsilon + 2^{7/2}\sqrt{\frac{\mathbb{E}_{x'}d(0,f)^2 t}{n}}\right) \ge 1-e^{-t}$$

$$d(f,g) = \left(\frac{1}{n}\sum_{i=1}^{n}\left(f(x_i) - f(x_i') - g(x_i) + g(x_i')\right)^2\right)^{1/2}$$

Definition 16.1. *We say that $\mathcal{F}$ satisfies uniform entropy condition if*

$$\forall n,\ \forall (x_1,\ldots,x_n),\ \mathcal{D}(\mathcal{F},\varepsilon,d_x) \le \mathcal{D}(\mathcal{F},\varepsilon)$$

where $d_x(f,g) = \left(\frac{1}{n}\sum_{i=1}^{n}\left(f(x_i) - g(x_i)\right)^2\right)^{1/2}$

Lemma 16.1. *If $\mathcal{F}$ satisfies uniform entropy condition, then*

$$\mathbb{E}_{x'}\int_0^{d(0,f)} \log^{1/2}\mathcal{D}(\mathcal{F},\varepsilon,d)d\varepsilon \le \int_0^{\sqrt{\mathbb{E}_{x'}d(0,f)^2}} \log^{1/2}\mathcal{D}(\mathcal{F},\varepsilon/2)d\varepsilon$$

Proof. Using inequality $(a+b)^2 \le 2(a^2+b^2)$,

$$\begin{aligned}
d(f,g) &= \left(\frac{1}{n}\sum_{i=1}^{n}\left(f(x_i) - g(x_i) + g(x_i') - f(x_i')\right)^2\right)^{1/2} \\
&\le \left(\frac{2}{n}\sum_{i=1}^{n}\left((f(x_i) - g(x_i))^2 + (g(x_i') - f(x_i'))^2\right)\right)^{1/2} \\
&= 2\left(\frac{1}{2n}\sum_{i=1}^{n}\left((f(x_i) - g(x_i))^2 + (g(x_i') - f(x_i'))^2\right)\right)^{1/2} \\
&= 2d_{x,x'}(f,g)
\end{aligned}$$

Since $d(f,g) \le 2d_{x,x'}(f,g)$, we also have

$$\mathcal{D}(\mathcal{F},\varepsilon,d) \le \mathcal{D}(\mathcal{F},\varepsilon/2,d_{x,x'}).$$

Indeed, let $f_1,\ldots,f_N$ be optimal ε-packing w.r.t. distance d. Then

$$\varepsilon \le d(f_i,f_j) \le 2d_{x,x'}(f_i,f_j)$$

and, hence,

$$\varepsilon/2 \le d_{x,x'}(f_i,f_j).$$

So, $f_1, ..., f_N$ is $\varepsilon/2$-packing w.r.t. $d_{x,x'}$. Therefore, can pack at least N and so $\mathcal{D}(\mathcal{F}, \varepsilon, d) \leq \mathcal{D}(\mathcal{F}, \varepsilon/2, d_{x,x'})$.

$$\begin{aligned}\mathbb{E}_{x'} \int_0^{d(0,f)} \log^{1/2} \mathcal{D}(\mathcal{F}, \varepsilon, d) d\varepsilon &\leq \mathbb{E}_{x'} \int_0^{d(0,f)} \log^{1/2} \mathcal{D}(\mathcal{F}, \varepsilon/2, d_{x,x'}) d\varepsilon \\ &\leq \int_0^{d(0,f)} \log^{1/2} \mathcal{D}(\mathcal{F}, \varepsilon/2) d\varepsilon\end{aligned}$$

Let $\phi(x) = \int_0^x \log^{1/2} \mathcal{D}(\mathcal{F}, \varepsilon) d\varepsilon$. It is concave because $\phi'(x) = \log^{1/2} \mathcal{D}(\mathcal{F}, \varepsilon/2)$ is decreasing when x is increasing (can pack less with larger balls). Hence, by Jensen's inequality,

$$\mathbb{E}_{x'} \phi(d(0,f)) \leq \phi(\mathbb{E}_{x'} d(0,f)) = \phi(\mathbb{E}_{x'} \sqrt{d(0,f)^2}) \leq \phi(\sqrt{\mathbb{E}_{x'} d(0,f)^2}).$$

□

Lemma 16.2. *If* $\mathcal{F} = \{f \colon \mathcal{X} \to [0,1]\}$, *then*

$$\mathbb{E}_{x'} d(0,f)^2 \leq 2 \max\left(\mathbb{E}f, \frac{1}{n}\sum_{i=1}^n f(x_i)\right)$$

Proof.

$$\begin{aligned}\mathbb{E}_{x'} d(0,f)^2 &= \mathbb{E}_{x'} \frac{1}{n}\sum_{i=1}^n \left(f(x_i) - f(x_i')\right)^2 \\ &= \frac{1}{n}\sum_{i=1}^n \left(f^2(x_i) - 2f(x_i)\mathbb{E}f + \mathbb{E}f^2\right) \\ &\leq \frac{1}{n}\sum_{i=1}^n (f^2(x_i) + \mathbb{E}f^2) \leq \frac{1}{n}\sum_{i=1}^n f(x_i) + \mathbb{E}f \\ &\leq 2\max\left(\mathbb{E}f, \frac{1}{n}\sum_{i=1}^n f(x_i)\right)\end{aligned}$$

□

Theorem 16.1. *If* $\mathcal{F}$ *satisfies Uniform Entropy Condition and* $\mathcal{F} = \{f \colon \mathcal{X} \to [0,1]\}$. *Then*

$$\mathbb{P}\left(\forall f \in \mathcal{F},\ \mathbb{E}f - \frac{1}{n}\sum_{i=1}^n f(x_i) \leq \frac{2^{9/2}}{\sqrt{n}} \int_0^{\sqrt{2\mathbb{E}f}} \log^{1/2} \mathcal{D}(\mathcal{F}, \varepsilon/2) d\varepsilon + 2^{7/2}\sqrt{\frac{2\mathbb{E}f \cdot t}{n}}\right) \geq 1 - e^{-t}.$$

Proof. If $\mathbb{E}f \geq \frac{1}{n}\sum_{i=1}^n f(x_i)$, then

$$2\max\left(\mathbb{E}f, \frac{1}{n}\sum_{i=1}^n f(x_i)\right) = 2\mathbb{E}f.$$

If $\mathbb{E}f \leq \frac{1}{n}\sum_{i=1}^n f(x_i)$,

$$\mathbb{E}f - \frac{1}{n}\sum_{i=1}^n f(x_i) \leq 0$$

and the bound trivially holds. □

Another result:

$$\mathbb{P}\left(\forall f \in \mathcal{F},\ \frac{1}{n}\sum_{i=1}^n f(x_i) - \mathbb{E}f \leq \frac{2^{9/2}}{\sqrt{n}}\int_0^{\sqrt{2\frac{1}{n}\sum_{i=1}^n f(x_i)}} \log^{1/2}\mathcal{D}(\mathcal{F},\varepsilon/2)d\varepsilon \right.$$
$$\left. +2^{7/2}\sqrt{\frac{2(\frac{1}{n}\sum_{i=1}^n f(x_i))t}{n}}\right) \geq 1 - e^{-t}.$$

Example 1 (VC-type entropy condition)**.**

$$\log\mathcal{D}(\mathcal{F},\varepsilon) \leq \alpha\log\frac{2}{\varepsilon}.$$

For VC-subgraph classes, entropy condition is satisfied. Indeed, in Lecture 13, we proved that $\mathcal{D}(\mathcal{F},\varepsilon,d) \leq \left(\frac{8e}{\varepsilon}\log\frac{7}{\varepsilon}\right)^V$ *for a VC-subgraph class* $\mathcal{F}$ *with* $VC(\mathcal{F}) = V$*, where* $d(f,g) = d_1(f,g) = \frac{1}{n}\sum_{i=1}^n |f(x_i) - g(x_i)|$*. Note that if* $f, g : \mathcal{X} \mapsto [0,1]$*, then*

$$d_2(f,g) = \left(\frac{1}{n}\sum_{i=1}^n (f(x_i) - g(x_i))^2\right)^{1/2} \leq \left(\frac{1}{n}\sum_{i=1}^n |f(x_i) - g(x_i)|\right)^{1/2}.$$

Hence, $\varepsilon < d_2(f,g) \leq \sqrt{d_1(f,g)}$ *implies*

$$\mathcal{D}(\mathcal{F},\varepsilon,d_2) \leq \mathcal{D}(\mathcal{F},\varepsilon^2,d_1) \leq \left(\frac{8e}{\varepsilon^2}\log\frac{7}{\varepsilon^2}\right)^V = \mathcal{D}(\mathcal{F},\varepsilon).$$

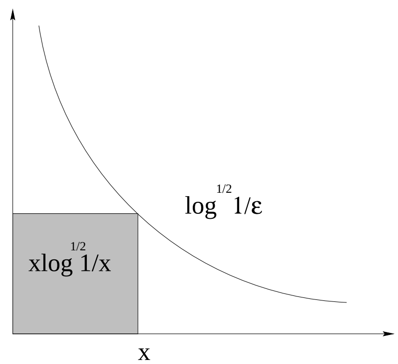
The entropy is

$$\log\mathcal{D}(\mathcal{F},\varepsilon) \leq \log\left(\frac{8e}{\varepsilon^2}\log\frac{7}{\varepsilon^2}\right)^V = V\log\left(\frac{8e}{\varepsilon^2}\log\frac{7}{\varepsilon^2}\right) \leq K\cdot V\log\frac{2}{\varepsilon},$$

where K *is an absolute constant.*

We now give an upper bound on the Dudley integral for VC-type entropy condition.

$$\int_0^x \sqrt{\log \frac{1}{\varepsilon}} d\varepsilon \le \begin{cases} 2x \log^{1/2} \frac{1}{x} & , \quad x \le \frac{1}{e} \\ 2x & , \quad x \ge \frac{1}{e} \end{cases} .$$



Proof. First, check the inequality for $x \le 1/e$. Taking derivatives,

$$\sqrt{\log \frac{1}{x}} \le 2\sqrt{\log \frac{1}{x}} + \frac{x}{\sqrt{\log \frac{1}{x}}} \left(-\frac{1}{x}\right)$$

$$\log \frac{1}{x} \le 2 \log \frac{1}{x} - 1$$

$$1 \le \log \frac{1}{x}$$

$$x \le 1/e$$

Now, check for $x \ge 1/e$.

$$\begin{aligned} \int_0^x \sqrt{\log \frac{1}{\varepsilon}} d\varepsilon &= \int_0^{\frac{1}{e}} \sqrt{\log \frac{1}{\varepsilon}} d\varepsilon + \int_{\frac{1}{e}}^x \sqrt{\log \frac{1}{\varepsilon}} d\varepsilon \\ &\le \frac{2}{e} + \int_{\frac{1}{e}}^x 1 dx \\ &= \frac{2}{e} + x - \frac{1}{e} = x + \frac{1}{e} \le 2x \end{aligned}$$

□

Using the above result, we get

$$\mathbb{P}\left(\forall f \in \mathcal{F},\ \mathbb{E}f - \frac{1}{n}\sum_{i=1}^n f(x_i) \le K\sqrt{\frac{\alpha}{n}\mathbb{E}f \log \frac{1}{\mathbb{E}f}} + K\sqrt{\frac{t\mathbb{E}f}{n}}\right) \ge 1 - e^{-t}.$$

Without loss of generality, we can assume $\mathbb{E}f \geq \frac{1}{n}$, and, therefore, $\log \frac{1}{\mathbb{E}f} \leq \log n$. Hence,

$$\mathbb{P}\left(\forall f \in \mathcal{F},\ \frac{\mathbb{E}f - \frac{1}{n}\sum_{i=1}^{n} f(x_i)}{\sqrt{\mathbb{E}f}} \leq K\sqrt{\frac{\alpha \log n}{n}} + K\sqrt{\frac{t}{n}}\right) \geq 1 - e^{-t}.$$