Problem 1

Solution

(a) Because $\forall i$, f_i is postive non-decreasing, $\prod_{i=1}^{k'} f_i$, $\forall 1 \le k' \le k$ is also non-decreasing. Thus,

$$\mathbb{E}\left[\prod_{i=1}^{k} f_i(\mathbf{X})\right] \le \mathbb{E}\left[\prod_{i=1}^{k-1} f_i(\mathbf{X})\right]\mathbb{E}\left[f_k(\mathbf{X})\right] \le \cdots \le \prod_{i=1}^{k}\mathbb{E}\left[f_i(\mathbf{X})\right].$$

(b) NA r.v.s $\{X_1, \dots, X_n\}$ may not be independent. If we can prove that Chernoff's trick also holds for NA r.v.s, then Hoeffding's inequality holds too.

$$P(\sum_{i=1}^{n} X_i \ge a) \le \frac{\mathbb{E}\left[\prod_{i=1}^{n} e^{\theta X_i}\right]}{e^{\theta a}} \le \frac{\prod_{i=1}^{n}\mathbb{E}\left[e^{\theta X_i}\right]}{e^{\theta a}}$$

where the last derivation uses the result in (a).

Problem 2

Solution

(a) Lemmas: 1. zero-one principle. 2. union of independent groups.

(b) Let Y_i be an indicator random variable to indicate if the i-th bin is non-empty. We have $Y_i \sim Ber(1 - (\frac{n-1}{n})^m)$, and $\mathbb{E}[Y_i] = 1 - (\frac{n-1}{n})^m$. Lemma: mapping. Thus, $\{Y_1, \dots, Y_n\}$ is NA. With the results in Problem 1 (b), we have $\forall o > 0$,

$$P\left(O - \mathbb{E}[O] = \sum_{i=1}^{n}(Y_i - \mathbb{E}[Y_i]) \ge o\right) \le \exp\left(\sum_{i=1}^{n}\frac{(b_i - a_i)^2}{8}\theta^2 - \theta o\right) = \exp\left(\frac{n}{8}\theta^2 - \theta o\right)$$
$$\le e^{\frac{-2o^2}{n}}$$

Problem 3

Solution

Let $Z_i \in \{1, 2, 3, \dots, n\}$ be the ID of the bin into which the i-th ball is put. Let f be a function such that $f(Z_1, Z_2, Z_3, \dots, Z_m)$ is the number of non-empty bins. Obviously, Z_i's are independent, and $|f(Z_1, \dots, Z_i, \dots, Z_m) - f(Z_1, \dots, Z_i', \dots, Z_m)| < 1$. Thus, McDiarmid's inequality applys:

$$P(O - \mathbb{E}[O] = f(Z_1, Z_2, Z_3, \dots, Z_m) - \mathbb{E}[f(Z_1, Z_2, Z_3, \dots, Z_m)] \ge o) \le e^{\frac{-2o^2}{n}},$$

which is the same as the Hoeffding bound.

Problem 4

Solution

(a) We have $\mathbb{E}[\exp(-\theta X)] \leq \exp(\nu^2\theta^2/2)$. Thus for any $\theta > 0$,

$$P(|X| > t) = P(X > t) + P(-X > t) = P(e^{\theta X} > e^{\theta t}) + P(e^{-\theta X} > e^{\theta t})$$
$$\leq \frac{\mathbb{E}\left[e^{\theta X}\right]}{e^{\theta t}} + \frac{\mathbb{E}\left[e^{-\theta X}\right]}{e^{\theta t}} \leq 2\exp(\nu^2\theta^2/2 - \theta t) \leq 2\exp(-\frac{t^2}{2\nu^2})$$

(b) Let the CDF of $|X|$ be F. Then, $F(x) = 0$, $\forall x \leq 0$. Because $g(x) = x^k$ is an increasing function, we have

$$\mathbb{E}\left[|X|^k\right] = \int_0^\infty g'(x)P(|X| > x)dx \leq 2\int_0^\infty kx^{k-1}\exp(-\frac{x^2}{2\nu^2})dx.$$

Subtitute $\frac{x^2}{2\nu^2}$ with y, and we have

$$\mathbb{E}\left[|X|^k\right] \leq 2k\int_0^\infty \nu^2(2\nu^2 y)^{k/2-1}\exp(-y)dy = k(2\nu^2)^{k/2}\Gamma(\frac{k}{2}).$$

Problem 5

Solution

First, we partation the indices into K index group such that $G_k = \{j : X_j \in A_k\}$. We also have for any $f \in F$ and j_1, $j_2 \in G_i$, $f(X_{j_1}) = f(X_{j_2})$. Thus,

$$R_n(F(X^n)) = \mathbb{E}\left[\sup_{f\in F}\left|\frac{1}{n}\sum_{i=1}^n \sigma_i f(X_i)\right|\right] = \mathbb{E}\left[\sup_{f\in F}\left|\frac{1}{n}\sum_{k=1}^K\sum_{i\in G_k}\sigma_i f(X_i)\right|\right] = \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^K\left|\sum_{i\in G_k}\sigma_i\right|\right]$$
$$\leq \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^K\sqrt{\sum_{i\in G_k}\sigma_i^2}\right] = \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^K\sqrt{n_k}\right] = \frac{1}{n}\sum_{k=1}^K\sqrt{n_k}$$

Problem 6

Solution

For $n = 3$, we can choose $z_2 = z_1 + 0.55$ and $z_3 = z_2 + 0.55$. It is easy to verify z_1, z_2, and z_3 can be shattered by F. When $n = 4$, no matter what z_1, z_2, z_3, and z_4 are, they cannot be assigned values 1, 0, 1, 0. Thus $VC(F) = 3$.

Problem 7

Solution

For n points, we can choose $x_i = 2^i\pi$, $i = 1, 2, 3, \ldots, n$. Then x^n can be shattered by F. Actually, we can find 2^n intervals $(1, \frac{2^n}{2^n-1})$, $(\frac{2^n}{2^n-1}, \frac{2^n}{2^n-2})$, $\ldots$, $(\frac{2^n}{1}, \infty)$ such that when θ belongs to different intervals, f assigns different values to x^n.