Homework 4

1. **Property of σ-Fields.** Let Ω be a sample space. Let $\mathcal{F}$ be a σ-field on Ω. Our objective is to prove that the sets $\mathcal{F}(x)$ partitions the set Ω.

   For any two elements $x, y \in \Omega$, if $\mathcal{F}(x) \neq \mathcal{F}(y)$ then prove that $\mathcal{F}(x) \cap \mathcal{F}(y) = \emptyset$.

   Solution.

2. **Independent Bounded Difference Inequality.** (20 points) Let $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$. Let $\mathbb{X}_1, \dots, \mathbb{X}_n$ be independent random variables over the sample spaces $\Omega_1, \dots, \Omega_n$. Let $f \colon \Omega \to \mathbb{R}$ be a function. Suppose there exists $c_1, \dots, c_n$ such that the following holds true. For every $1 \leqslant i \leqslant n$, and $x, y \in \Omega$ such that x and y are identical everywhere except at the i-th coordinate, then we have $f(x) - f(y) \leqslant c_i$. Let $\mu = \mathbb{E}\left[f(\mathbb{X}_1, \dots, \mathbb{X}_n)\right]$.

   Using Azuma's inequality, prove the following concentration inequality.

$$\mathbb{P}\left[f(\mathbb{X}_1, \dots, \mathbb{X}_n) - \mu \geqslant E\right] \leqslant \exp\left(-2E^2 / \sum_{i=1}^{n} c_i^2\right)$$

   Solution.

3. **Pólya's Urn.** (20 points) Consider the following experiment. There is an urn with R red balls and B blue balls at time $t = 0$. At every time step $t = 1, \dots, n$ you sample a ball $\mathbb{X}_t$ from the urn. Next, you replace the ball that you sample and introduce another ball that has the same color as $\mathbb{X}_t$ into the urn. After n times steps, let $\mathbb{S}_n$ represent the total number of red balls that you had sampled. Formally, $\mathbb{S}_n = \sum_{i=1}^{n} \mathbf{1}_{\{\mathbb{X}_i = R\}}$.

   (a) (5 points) What is $\mathbb{E}[\mathbb{S}_n]$?

   (b) (15 points) State and prove a concentration bound for $\mathbb{S}_n$ around its expected value using Azuma's inequality.

   Solution.

4. **Never too Far from Home!** (20 points) Suppose you start at your home at time $t = 0$. At every time step, you either take one step north or one step south uniformly and independently at random. Chernoff bound states that the probability that you "are far from your home at $t = n$ is small."

   In this problem, we want to claim that "we never go far from home at any time $t \leqslant n$."

   Let me elaborate the difference. Suppose $\mathbb{X}_i = +1$ represents a north-step at time $t = i$, and $\mathbb{X}_i = -1$ represents a south-step at time $t = i$. Let $\mathbb{S}_i = \mathbb{X}_1 + \cdots + \mathbb{X}_i$ represent our position at time $t = i$. Chernoff bound states that the following probability is small.

$$\mathbb{P}\left[\mathbb{S}_n \geqslant E\right]$$

   We want to show that the following probability is small.

$$\mathbb{P}\left[\max_{1 \leqslant i \leqslant n} \mathbb{S}_i \geqslant E\right]$$

   The second event is much harsher than the first event. For example, if you stray away from home even once but, later, return back close to home at time $t = n$ then the first event forgives you but the second event does not!

   Prove the following concentration bound using Azuma's inequality.

$$\mathbb{P}\left[\max_{1 \leqslant i \leqslant n} \mathbb{S}_i \geqslant E\right] \leqslant \exp(-E^2/2n)$$

   Solution.

5. **Largest Convex Subset.** (20 points) Suppose we pick n points $\mathbb{X}_1, \dots, \mathbb{X}_n$ uniformly at random from the unit square $[0,1]^2$. A set of points is said to be in *convex position* if no point in this set can be written as the convex linear combination of other points in the set. Let $\mathbb{S}$ represent the size of the largest subset of $\{\mathbb{X}_1, \dots, \mathbb{X}_n\}$ that lie in convex position. Use the Talagrand inequality to prove a concentration of the random variable $\mathbb{S}$ around its median.

   Solution.

Collaborators :