Problem Set 2

- This problem set is due on **February 24, 2020**.
- Each problem carries 10 points.
- Collaboration is **not permitted**. Each student must submit his/her own work.

1. **(Probabilistic Method)** *Probabilistic method* is a simple yet powerful tool to prove existential results. Consider the following general problem statement - we are given a set of real numbers $S = \{v_1, v_2, \ldots, v_N\}$ and a threshold λ. We want to show that there exists an element $v_i \in S$ such that $v_i \geq \lambda$. In order to solve this problem via probabilistic method, we first define a *suitable* probability distribution $\boldsymbol{p} = (p_1, p_2, \ldots, p_N)$ over the set S. Next, we compute the expectation $E_{\boldsymbol{p}} = \sum_{i=1}^{N} p_i v_i$. If $E_{\boldsymbol{p}} \geq \lambda$, the existence of the desired element follows immediately (convince yourself!). The trick here is to pick a suitable distribution $\boldsymbol{p}$ depending on the problem at hand. Solve the following problems using the above principle:

   (a) Let $G = (V, E)$ be a simple graph and let d_v denote the degree of of the vertex v. An *independent set* is a set of vertices no pair of which is joined by an edge. Let $\alpha(G)$ be the size of the largest independent set of G. Show that

   $$\alpha(G) \geq \sum_{v \in V} \frac{1}{d_v + 1}.$$

   (b) Let A be an arbitrary binary $n \times n$ matrix (*i.e.,* $A_{i,j} \in \{0, 1\}, \forall i, j.$) Show that there exists a column vector $b \in \{-1, +1\}^n$ so that $||Ab||_\infty = O(\sqrt{n \log n})$[1].

   HINT: Let the elements of b be i.i.d. uniform in $\{\pm 1\}$. Bound the MGF of $||Ab||_\infty$ by the sum of MGFs of $|A_i b|, \forall i$, where A_i denotes the i^{th} row of A. Finally, optimize the argument of the MGF to get the best bound possible.

   ---

2. **(Random Walk)** Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables with

   $$\mathbb{P}(X_1 = 1) = p, \mathbb{P}(X_1 = -1) = 1 - p,$$

   where $0 \leq p \leq 1$ and $p \neq \frac{1}{2}$. Let $S_0 = 0$ and $S_n = S_{n-1} + X_n, n \geq 1$. For each $n \geq 1$, define the event $A_n = \{\omega : S_n(\omega) = 0\}$ and let $A = \limsup A_n$. Let $\mathcal{F}_\infty$ be the tail σ-algebra corresponding to the sequence of r.v.s $\{X_n\}_{n\geq 1}$. Show that

[1] Recall, for a vector x, $||x||_\infty := \max_i |x_i|$.

(a) $A \notin \mathcal{F}_\infty$.

(b) Nonetheless, $\mathbb{P}(A) \in \{0, 1\}$.

HINT: Use Stirling's approximation.

---

3. **($\mathcal{L}^2$ vs. $\mathcal{L}^1$)**

   (a) Let X_1, X_2 be i.i.d. non-negative random variables such that $X_1 \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ but $X_1 \notin \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Let $Y = \min\{X_1, X_2\}$. Show that $Y \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

   (b) (**Generalization**) Let $X_1, X_2, \ldots, X_n$ be a collection of n i.i.d. non-negative random variables such that $X_1 \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ but $X_1 \notin \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Let Y be the *Second largest* random variable in the above collection [2]. Show that $Y \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

4. **(Conditional Expectation)**

   (a) Let the random variables $\{Z_n\}_{n\geq 1}$ be independent, each with finite mean. Let $X_0 = a$, and $X_n = a + Z_1 + Z_2 + \ldots + Z_n$ for $n \geq 1$. Prove that

   $$\mathbb{E}\big(X_{n+1} | \sigma(X_1, X_2, \ldots, X_n)\big) = X_n + \mathbb{E}(Z_{n+1}).$$

   (b) Suppose that $X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ such that

   $$\mathbb{E}(X|\sigma(Y)) = Y, \mathbb{E}(Y|\sigma(X)) = X \quad \text{a.s.}$$

   Show that $X = Y$ almost surely.

   (c) **(Linear Estimation)** Let $X_1, X_2, \ldots, X_n$ be random variables with zero expectations and covariance matrix $\boldsymbol{V}$ [3]. Using the orthogonality principle, find the linear map $h(\cdot)$ of $\{X_i\}_{i=1}^n$ which minimizes the mean squared error $\mathbb{E}\{(Y - h(X_1, X_2, \ldots, X_n))^2\}$.

5. **(Kolmogorov-Hajek-Renyi inequality)** Let $\{Z_n\}_{n\geq 0}$ be a Martingale sequence with $\mathbb{E}(Z_0) = 0$ and let $\{v_j\}_{j\geq 0}$ be a sequence of non-decreasing constants with $v_0 = 0$. Prove that

   $$\mathbb{P}(|Z_j| \leq v_j, \quad \forall 1 \leq j \leq n) \geq 1 - \sum_{j=1}^{n} \mathbb{E}(Z_j - Z_{j-1})^2 / v_j^2.$$

---

[2] In order-statistic notation $Y = X_{(n-1)}$.

[3] This means that $V_{ij} = \mathbb{E}(X_i X_j), 1 \leq i, j \leq n$.