MathTools HW 1

Do not be alarmed by the length of this problem set; the exercises, while numerous, are all very short.

1. Let X and Y be two discrete random variables, taking values on (for simplicity, finite) sets $\mathcal{X}$ and $\mathcal{Y}$ respectively. Recall that they are **independent** if for every $x \in \mathcal{X}, y \in \mathcal{Y}$,

$$\Pr(X = x \text{ and } Y = y) = \Pr(X = x) \Pr(Y = y) .$$

Prove that the following property is *equivalent* (implies and is implied by - there are two directions to show here) to independence: for every functions $f : \mathcal{X} \to \mathbb{R}$ and $g : \mathcal{Y} \to \mathbb{R}$,

$$\mathbb{E}\left[f(X)g(Y)\right] = \mathbb{E}\left[f(X)\right] \mathbb{E}\left[g(Y)\right] .$$

2. (a) Let X, Y be real-valued random variables. Show that for any $a, t \in \mathbb{R}$,

$$\Pr\left(X + Y \geq t\right) \leq \Pr\left(X \geq a\right) + \Pr\left(Y \geq t - a\right) .$$

Hint: Use the following properties of probability measures:

- *Monotonicty*: If E_1 and E_2 are events such that $E_1 \subset E_2$, then $\Pr(E_1) \leq \Pr(E_2)$.
- *Union bound*: For any events E_1, E_2, $\Pr(E_1 \cup E_2) \leq \Pr(E_1) + \Pr(E_2)$.

(b) Assume now that $X, Y, a, t > 0$. Show that

$$\Pr(XY \geq t) \leq \Pr(X \geq a) + \Pr(Y \geq t/a) .$$

3. Suppose that X is a random variable that takes values in $\{0, 1, 2, \ldots\}$.

(a) Show that

$$\Pr(X > 0) \leq \mathbb{E}[X] .$$

(b) Assume that $\mathbb{E}[X] > 0$. Show that

$$\Pr(X = 0) \leq \frac{\mathrm{Var(X)}}{\mathbb{E}[X]^2} .$$

4. **Cauchy-Schwartz inequality.** Let X, Y be random variables. Prove that

$$|\mathbb{E}[XY]| \leq \left(\mathbb{E}(X^2)\right)^{1/2} \left(\mathbb{E}(Y^2)\right)^{1/2}.$$

Hint: Compute $\min_{\alpha \in \mathbb{R}} \mathbb{E}[(X - \alpha Y)^2]$.

Remark: Compare this with the Cauchy-Schwartz inequality you've seen in your linear algebra courses. Indeed, $\langle X, Y \rangle = \mathbb{E}[XY]$ is (almost) an inner product, on the vector space of random variables. In this sense, the quantity $\min_{\alpha \in \mathbb{R}} \mathbb{E}[(X - \alpha Y)^2]$ is the (squared) distance between X and the subspace spanned by Y (where distance is with respect to the corresponding norm, $\|X\|^2 = \mathbb{E}[X^2]$. Don't worry if you didn't follow this explanation - we will talk about inner product spaces in more detail later in the course.

Remark 2: An immediate consequence of this inequality is that $|\mathrm{Cov}(X, Y)| \leq (\mathrm{Var}(X))^{1/2} (\mathrm{Var}(Y))^{1/2}$ (make sure you see why!). The quantity

$$\rho = \frac{\mathrm{Cov}(X, Y)}{(\mathrm{Var}(X))^{1/2} (\mathrm{Var}(Y))^{1/2}} \in [-1, 1]$$

is sometimes called the correlation coefficient between X and Y.

5. **Paley-Zygmund inequality.** Let $X \geq 0$ be a random variable. Prove that for any $0 \leq \theta \leq 1$:

$$\Pr\left(X > \theta \mathbb{E}[X]\right) \geq (1 - \theta)^2 \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}.$$

Hint: Write $X = X\mathbb{1}_{\{X \leq \theta \mathbb{E}[X]\}} + X\mathbb{1}_{\{X > \theta \mathbb{E}[X]\}}$, and take the expectation. The first term, $\mathbb{E}\left[X\mathbb{1}_{\{X \leq \theta \mathbb{E}[X]\}}\right]$ can be bounded readily from above (it is the expectation of a bounded random variable). To bound the second term, use the Cauchy-Shwartz inequality.

6. A random variable X taking value in $\mathbb{R}$ is, clearly, determined completely by its tail probabilities $\Pr(X \geq t)$ for all t. In this exercise we shall show how to recover the expectation $\mathbb{E}[X]$ from the tail, assuming that X is non-negative.

   (a) Assume that X takes values in $0, 1, 2, \ldots$. Prove (directly from the definition) that

   $$\mathbb{E}[X] = \sum_{k=1}^{\infty} \Pr(X \geq k).$$

   (b) Assume that $X \geq 0$, but otherwise not constrained to be discrete. Prove that

   $$\mathbb{E}[X] = \int_0^{\infty} \Pr(X \geq t) dt.$$

   Make sure you see why (a) follows from this!

   Hint: Since X is non-negative, we can write $X = \int_0^X dt = \int_0^\infty \mathbb{1}_{t \leq X} dt$. Take the expectation, and exchange the order of the integral and expectation (no need to justify why this is legit).

7. Let $X_1, \ldots, X_n$ be RVs, **not necessarily independent**, with uniformly bounded p-th absolute moment (here $p > 0$):

$$\mathbb{E}\left[|X_i|^p\right] \le m_p\,, \quad \forall i = 1, \ldots, n\,.$$

(a) Show that for any $t > 0$,

$$\Pr\left(|X_i| \ge t\right) \le \frac{m_p}{t^p}\,.$$

Hint: Mimic the proof of Chebyshev's inequality.

(b) Let a_n be any sequence such that $a_n \to \infty$. Prove that

$$\lim_{n\to\infty} \Pr\left(\max_{1\le i\le n} |X_i| \ge a_n \cdot n^{1/p}\right) = 0\,.$$

Hint: Use a union bound.