Problem Set 2

- This problem set is due on **September 14, 2021** in the class.
- Each problem carries 10 points.
- Collaboration is **strictly prohibited**. Each student must submit their own work.

1. **(Independent Normals)** Let X_1, X_2 be two independent Gaussian variables such that $\text{Var}(X_1) = \text{Var}(X_2)$. Show that $X_1 + X_2 \perp X_1 - X_2$.

2. **(Bivariate Normals)** Let the r.v.s X and Y be jointly Gaussian with zero mean, unit variances and $\mathbb{E}(XY) = \rho$. Find $\mathbb{P}(X \geq 0, Y \geq 0)$.

   `Hint:` Transform to polar coordinates.

3. **(Convergence of R.V.s)** Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables such that $X_1 = 0$ and for all $n \geq 2$:

$$\begin{aligned} \mathbb{P}(X_n = n) &= \frac{1}{2n\log n}, \\ \mathbb{P}(X_n = -n) &= \frac{1}{2n\log n}, \\ \mathbb{P}(X_n = 0) &= 1 - \frac{1}{n\log n}. \end{aligned}$$

   Let $S_n = \sum_{i=1}^{n} X_i$. Show that for any $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}(|\frac{S_n}{n}| > \epsilon) = 0, \quad \text{but} \quad \frac{S_n}{n} \nrightarrow 0 \text{ a.s.}.$$

Hint: For the first part, the following elementary result on limits might be useful.

Theorem 0.1 (Stolz-Cesàro Theorem) *Let $\{b_n\}_{n\geq 1}$ be a sequence of positive real numbers, such that $\sum_{n=1}^{\infty} b_n = \infty$. Let $\{a_n\}_{n\geq 1}$ be another sequence of real numbers such that $\lim_{n\to\infty} \frac{a_n}{b_n}$ exists. Then*

$$\lim_{n\to\infty} \frac{a_1 + a_2 + \ldots + a_n}{b_1 + b_2 + \ldots + b_n} = \lim_{n\to\infty} \frac{a_n}{b_n}.$$

For the second part, use the second Borel-Cantelli Lemma.

4. **(Paley-Zygmund's inequality)** (a) Let X be a random variable such that $\mathbb{E}X > 0$ and $\mathbb{E}X^2 < \infty$. Using the Cauchy-Schwartz inequality with the random variables X and $Y = 1_{\{X>t\}}$ or otherwise, show that

$$\mathbb{P}(X > t) \geq \frac{(\mathbb{E}(X) - t)^2}{\mathbb{E}X^2}, \quad 0 \leq t \leq \mathbb{E}X.$$

(b) Using the above result or otherwise, show that for a r.v. X with $\mathbb{E}X = \mu > 0$ and variance σ^2 satisfying $0 \leq \sigma \leq \mu$, the following inequality holds:

$$\mathbb{P}(X > \mu - \sigma) \geq \frac{\sigma^2}{\sigma^2 + \mu^2}.$$

(c) **(Anti-concentration)** Let $X_1, X_2, \ldots, X_n$ be independent and uniformly distributed random variables (*i.e.,* each X_i is uniformly distributed in $\{+1, -1\}$). Let $S_n = \sum_{i=1}^n X_i$. Show that for $\alpha \in (0, 1)$:

$$\mathbb{P}(|S_n| > \alpha\sqrt{n}) \geq \frac{(1 - \alpha^2)^2}{3}.$$

`Hint:` In computing $\mathbb{E}S_n^4$, note that all odd-powered terms vanish due to symmetry.

5. **(Darmois-Skitovic's Theorem)** This problem proves a converse result to Problem 1. Let X_1, X_2 be two independent random variables with finite variances such that $X_1 + X_2$ and $X_1 - X_2$ are also independent random variables. Assume that log-MGFs of the random variables X_1 and X_2 are twice continuously differentiable. Using the following steps, show that X_1 and X_2 are jointly Gaussian random variables such that $\texttt{Var}(X_1) = \texttt{Var}(X_2)$.

   (a) Let $f_1 = \log \phi_{X_1}$ and $f_2 = \log \phi_{X_2}$ be the log-MGFs. Show that there exist functions g_1 and g_2 satisfying

   $$f_1(t_1 + t_2) + f_2(t_1 - t_2) = g_1(t_1) + g_2(t_2), \quad \forall t_1, t_2 \in \mathbb{R}. \tag{1}$$

   (b) If f_1, f_2 are twice continuously differentiable and there exist functions g_1 and g_2 satisfying (1), show that f_1 and f_2 are polynomials of degree at most 2.

   (c) Using the above, conclude the result.