Probability Theory 2, Spring 2024 - Homework 1
Due one hour before lecture on 2/11 (Gradescope)

Your submitted solutions to assignments should be your own work. While discussing homework problems with peers is permitted, the final work and implementation of any discussed ideas must be executed solely by you. Acknowledge any source you consult.

Problem 1 - How bad can atoms be?

Let $\mu \in (\mathbb{R})$ be a probability measure and $F(x) := \mu((-\infty, x])$ be its associated distribution function. Show that F has at most countably many discontinuities. Further, let $\nu \in (\mathbb{R})$, show that if $X \sim \mu$ and $Y \sim \nu$ be independent, then

$$\mathbb{P}(X + Y = 0) = \sum_y \mu(\{-y\})\nu(\{y\}).$$

Problem 2 - Levy's Inversion Formula for atoms

Let $\mu \in (\mathbb{R})$ be a probability measure and let φ be its characteristic function. In class we covered Levy's Inversion formula for intervals. Here you will develop one for atoms. Show that the following facts hold true.

(a) For any $a \in \mathbb{R}$ we have

$$\mu(\{a\}) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} e^{itx} \varphi(t) dt.$$

(b) Let $X \sim \mu$, if $\mathbb{P}(X \in h\mathbb{Z}) = 1$ where $h > 0$ and $\mathbb{Z}$ denotes the integer numbers. Then, $\varphi(2\pi/h + t) = \varphi(t)$ and so

$$\mathbb{P}(X = x) = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} e^{itx} \varphi(t) dt \quad \text{for any } x \in h\mathbb{Z}.$$

(c) Let X and Y be i.i.d. random variables with distribution μ. Then,

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} |\varphi(t)|^2 dt = \mathbb{P}(X - Y = 0) = \sum_x \mu(\{x\})^2.$$

Conclude that if $\varphi(t) \to 0$ as $t \to \infty$, then μ has no point masses.

Problem 2 - Central Limit Theorem

For this question you will use what we proved in class to establish a fairly general CLT for i.i.d. random variables. The strategy is simple: leverage Levy's Convergence Theorem by proving the characteristic function of normalized averages converges to the characteristic function of a normal distribution. For any n, define the Taylor residual as

$$R_n(x) := \exp(ix) - \sum_{k=0}^{n} \frac{(ix)^n}{k!}.$$

(a) Notice that $R_0(x) = \exp(ix) - 1 = \int_0^x i\exp(iy)dy$. Show that $|R_0(x)| \leq \min\{2, |x|\}$ and use the fact that $R_n(x) = \int_0^x iR_{n-1}(y)dy$ to derive that for any n

$$|R_n(x)| \leq \min\left\{\frac{2|x|^n}{n}, \frac{|x|^{n+1}}{(n+1)!}\right\}.$$

This inequality right here is the reason why our CLT will only need second moments.

(b) Let $X \in \mathcal{L}^2(\mathbb{R})$ be a random variable with zero mean and φ be its characteristic function. Prove that for any θ

$$\varphi(\theta) = 1 - \frac{1}{2}\sigma^2\theta^2 + o(\theta^2)$$

where $o(\theta^2)$ is a term such that $o(\theta^2)/\theta^2 \to 0$ as $\theta \to 0$.

(c) Let $X_1, \ldots, X_n \in \mathcal{L}^2(\mathbb{R})$ be zero-mean i.i.d. random variables with $\sigma^2 = \mathrm{Var}(X_1)$, and define $S_n := \frac{1}{n}\sum_{k=1}^n X_k$ and $G_n = \frac{\sqrt{n}S_n}{\sigma}$. Show that for any $x \in \mathbb{R}$ and as $n \to \infty$, we have that

$$\mathbb{P}\left(G_n \leq x\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x \exp(-y^2/2)dy.$$

Hint: Consider analyzing the limit of $\log \varphi_{Gn}(\theta)$ (You can use the bonus question without proving it).

(d) (Bonus question) Let $z \in$ be such that $|z| \leq 1/2$, show that

$$|\log(1+z) - z| \leq |z|^2.$$

Problem 3 - Coupon collecting

Let $X_1, X_2, \ldots$ be independent and uniformly drawn from $1, 2, \ldots, n\}$. You can imagine these arn n "coupons." Let

$$\tau_k^n = \min\{t\colon |\{X_1, \ldots, X_t\}| = k\}.$$

That is, the first time when k distinct coupons are collected.

(a) Show that increment $\Delta_k^n = \tau_k^n - \tau_{k-1}^n$ has a geometric distribution with parameter $1 - (k-1)/n$ and it is independent of Δ_j^n for all $1 \leq j < k$.

(b) Suppose that $n \to \infty$ and $\frac{k}{\sqrt{n}} \to \lambda \in (0, \infty)$, show that

$$\tau_k^n - k \xrightarrow{w} Z \qquad Z \sim \mathrm{Poisson}(\lambda^2/2).$$

Problem 4 - Convergence in practice

For this exercise we will empirically observe the convergence of the CLT and the Law of rare events.

(a) Repeat the following exercise with every possible parameter configuration of $n \in \{100, 500, 1000\}$ and $\lambda \in \{1/2, 1, 2\}$: Code a Python function that given n and λ generates a random draw of τ_k^n from Problem 3 where we set $k = \lceil \lambda\sqrt{n} \rceil$. Simulate $N = 1000$ i.i.d. draws of $\tau_k^n - k$, plot a histogram of the N draws and the density of $\mathrm{Poisson}(\lambda^2/2)$ (try to plot them on top of each other to ease the comparison).

(b) Repeat the following exercise with every possible parameter configuration of $n \in \{100, 500, 1000\}$: Code a Python function that given n generates a random draw of $S_n = \frac{1}{n}\sum_{k=1}^{n} X_i$ where the X_i's are i.i.d. Rademacher random variables. Simulate $N = 1000$ i.i.d. draws of $G_n = \sqrt{n}S_n$, plot a histogram of the N draws and the density of $N(0, 1)$.