Preliminary Exam: Probability

Time: 10:00am - 3:00pm, Thursday, August 20, 2020.

Your goal should be to demonstrate mastery of probability theory and maturity of thought. Your arguments should be clear, careful, and complete.

The exam consists of six main problems, each with several steps designed to help you in the overall solution.

Important: If you cannot solve a certain part of a problem, you still may use its conclusion in a later part!

Please make sure to apply the following guidelines:

1. On each page you turn in, write your assigned code number and the page number (n/m). Don't write your name on any page.
2. Start each problem on a new page.

1. Let $Z_1, \dots, Z_n$ be independent and identically distributed (i.i.d.) random variables with $Z_1 \sim N(0,1)$. Throughout this exam, $N(0,1)$ denotes a standard normal distribution. Let

$$\overline{Z} = \frac{\sum_{k=1}^n Z_k}{n}.$$

(1a). Find the characteristic function of $\overline{Z}$.

(1b). Find functions $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ so that for all $s_1, \dots, s_n, t \in \mathbb{R}$ we have

$$\mathbb{E}\left(e^{i[t\overline{Z} + \sum_{k=1}^n s_k(Z_k - \overline{Z})]}\right) = f(s_1, \dots, s_n) \cdot g(t).$$

[Hint: You may find $A_1, \dots, A_n$ so that $t\overline{Z} + \sum_{k=1}^n s_k(Z_k - \overline{Z}) = \sum_{k=1}^n A_k Z_k$. Also, to streamline notations, it may be helpful to use $n\bar{s} = \sum_{k=1}^n s_k$.]

(1c). It follows from part (1b) that the random vector $(Z_k - \overline{Z},\ k = 1, \dots, n)$ and the random variable $\overline{Z}$ are independent. You can use this fact without a proof.

Let $X_1, \dots, X_n$ be i.i.d. random variables with $X_1 \sim N(\mu, \sigma^2)$ and let $\overline{X} = \frac{\sum_{k=1}^n X_k}{n}$.

(i). Prove that the random variables $(X_k - \overline{X}, k = 1, \dots, n)$ and the random variable $\overline{X}$ are independent.

(ii). Let $S^2 = \frac{\sum_{k=1}^n (X_k - \overline{X})^2}{n-1}$. Prove that S^2 and $\overline{X}$ are independent.

2. (2a). It is proved in the textbook that if Y is a random variable with density

$$f_Y(y) = \frac{1}{2}e^{-|y|}, \qquad y \in \mathbb{R},$$

then its characteristic function is given by $\varphi_Y(t) = \frac{1}{1+t^2}$, $t \in \mathbb{R}$.

(i). Prove that $\int_{\mathbb{R}} |\varphi_Y(t)| dt < \infty$.

(ii). Quote the inversion formula for an integrable characteristic function and use it to calculate the density function of a random variable X whose characteristic function is given by $\varphi_X(t) = e^{-|t|}$, $t \in \mathbb{R}$.

(2b). Let $\{X, X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables and let $S_n = \sum_{k=1}^{n} X_k$ for all $n \geq 1$. Assume that the characteristic function of X is given by

$$\varphi_X(t) = e^{-|t|}, \qquad t \in \mathbb{R}.$$

Prove that, for every $n \geq 1$, the characteristic functions of $\frac{S_n}{n}$ and X are identical.

(2c). We continue with the notations of part (2b). Let $\varphi_{\frac{S_n}{\sqrt{n}}}(t)$ be the characteristic function of $\frac{S_n}{\sqrt{n}}$.

(i). Calculate $\lim_{n\to\infty} \varphi_{\frac{S_n}{\sqrt{n}}}(t)$ for all $t \in \mathbb{R}$. Is the limit a continuous function in $t \in \mathbb{R}$? Is the limit a characteristic function?

(ii). Does $\frac{S_n}{\sqrt{n}}$ converge in distribution?

3. Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables with

$$\mathbb{P}(X_n = 2) = \frac{1}{n}, \qquad \mathbb{P}(X_n = -1) = 1 - \frac{1}{n}.$$

(3a) Show that the variance $\mathrm{Var}(S_n)$ of $S_n = \sum_{m=1}^{n} X_m$ satisfies

$$\lim_{n\to\infty} \frac{\mathrm{Var}(S_n)}{9\ln n} = 1.$$

[Hint: You may use the fact that $\left|\sum_{m=1}^{n} \frac{1}{m} - \ln n\right| \leq 1$ for all $n \geq 1$.]

(3b). For any $n \geq 2$ and $1 \leq m \leq n$, define

$$X_{m,n} = \frac{X_m - \mathbb{E}(X_m)}{3\sqrt{\ln n}}.$$

Prove that as $n \to \infty$,

$$\sum_{m=1}^{n} X_{m,n} \Longrightarrow N(0,1).$$

(3c). Prove that as $n \to \infty$,

$$\frac{S_n + n - 3\ln n}{3\sqrt{\ln n}} \Longrightarrow N(0,1).$$

4. Let $\{S_n, n \geq 0\}$ be a simple random walk on $\mathbb{Z}$ starting at 0. Namely, $S_0 = 0$ and

$$S_n = X_1 + \cdots + X_n, \qquad n \geq 1,$$

where $\{X_n\}_{n \geq 1}$ are i.i.d. with $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$.

(4a). Compute $\mathbb{P}(S_n = 2)$.

(4b). Show that $\sum_{n=1}^{\infty} \mathbb{P}(S_n = 2) = \infty$.

[Hint: You may use Stirling's formula $n! \sim e^{-n} n^n \sqrt{2\pi n}$ as $n \to \infty$.]

(4c). Let N be the first time that $\{S_n, n \geq 0\}$ hits 1:

$$N = \inf\{n \geq 1 : S_n = 1\}.$$

Compute $\mathbb{P}(S_{N+n} = 3)$.

(4d). For any $n \geq 0$, let $\mathcal{F}_n = \sigma(S_k,\, 0 \leq k \leq n)$ be the σ-algebra generated by $(S_k,\, 0 \leq k \leq n)$. Find a sequence of constants $\{a_n\}_{n \geq 0}$ such that $\{S_n^2 - a_n\}_{n \geq 1}$ is a martingale with respect to $\{\mathcal{F}_n\}_{n \geq 0}$.

5. Let $\{\xi_i^{(n)}\}_{i,n\geq 1}$ be a sequence of i.i.d. non-negative integer-valued random variables. The corresponding Galton-Watson process $\{Z_n\}_{n\geq 0}$ is defined by $Z_0 = 1$ and

$$Z_{n+1} = \begin{cases} \sum_{i=1}^{Z_n} \xi_i^{(n+1)}, & \text{if } Z_n > 0; \\ 0, & \text{if } Z_n = 0. \end{cases}$$

Let $\mu = \mathbb{E}(\xi_1^{(1)})$ and $X_n = \frac{Z_n}{\mu^n}$ for all $n \geq 0$. Then $\{X_n\}_{n\geq 0}$ is a martingale with respect to the filtration $\{\mathcal{F}_n\}_{n\geq 0}$, where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma(\xi_i^{(m)} : 1 \leq m \leq n, i \geq 1)$ for all $n \geq 1$. It is known that $\{X_n\}_{n\geq 0}$ converges almost surely.

We assume $\sigma^2 = \mathrm{Var}(\xi_1^{(1)}) < \infty$. This implies $\mathbb{E}(X_n^2) < \infty$ for all $n \geq 1$. You can use this fact without a proof.

(5a). Write $X_n = X_{n-1} + (X_n - X_{n-1})$ and prove

$$\begin{aligned} \mathbb{E}\big(X_n^2|\mathcal{F}_{n-1}\big) &= X_{n-1}^2 + \mathbb{E}\big[(X_n - X_{n-1})^2|\mathcal{F}_{n-1}\big] \\ &= X_{n-1}^2 + \mu^{-2n}\mathbb{E}\big[(Z_n - \mu Z_{n-1})^2|\mathcal{F}_{n-1}\big]. \end{aligned}$$

(5b). Show that $\mathbb{E}\big[(Z_n - \mu Z_{n-1})^2|\mathcal{F}_{n-1}\big] = Z_{n-1}\sigma^2$.

(5c). Show that for all integers $n \geq 1$, we have

$$\mathbb{E}(X_n^2) = 1 + \sigma^2 \sum_{j=2}^{n+1} \mu^{-j}.$$

(5d). Assume $\mu > 1$. Show that $\{X_n\}_{n\geq 0}$ converges in $L^2(\mathbb{P})$.

(5e). Assume $\mu > 1$. Show that $\{X_n\}_{n\geq 0}$ is uniformly integrable and converges in $L^1(\mathbb{P})$.

6. Let $\{B(t), t \geq 0\}$ be a real-valued Brownian motion starting from 0 and let $T > 0$ be a fixed constant. For any partition Π of $[0, T]$, where

$$\Pi = \big\{(t_0, t_1, \ldots, t_m) : 0 = t_0 < t_1 < \cdots < t_{m-1} < t_m = T\big\},$$

we define $\Delta(\Pi) = \max\limits_{1 \leq i \leq m} |t_i - t_{i-1}|$ and

$$Q(\Pi) = \sum_{i=1}^{m} |B(t_i) - B(t_{i-1})|^2.$$

(6a). Compute $\mathbb{E}\big[Q(\Pi)\big]$ and $\mathbb{E}\big[(Q(\Pi) - T)^2\big]$.

(6b). Let $\{\Pi_n, n \geq 1\}$ be a sequence of partitions of $[0, T]$, where

$$\Pi_n = \big\{(t_0^n, t_1^n, \ldots, t_m^n) : 0 = t_0^n < t_1^n < \cdots < t_{m-1}^n < t_m^n = T\big\},$$

such that $\lim\limits_{n\to\infty} \Delta(\Pi_n) = 0$. Prove $\lim\limits_{n\to\infty} \mathbb{E}\big[(Q(\Pi_n) - T)^2\big] = 0$.

(6c). Assume $\limsup\limits_{n\to\infty} n^2 \Delta(\Pi_n) < \infty$. Prove $\lim\limits_{n\to\infty} Q(\Pi_n) = T$ almost surely.

(6d). For each partition Π_n, consider the variation

$$V(\Pi_n) = \sum_{i=1}^{m} |B(t_i^n) - B(t_{i-1}^n)|.$$

Assume that $\lim\limits_{n\to\infty} \Delta(\Pi_n) = 0$. Prove that if $\lim\limits_{n\to\infty} Q(\Pi_n) = T$ almost surely then $\lim\limits_{n\to\infty} V(\Pi_n) = \infty$ almost surely.