Preliminary Exam: Probability.

Time: 9:00am - 2:00pm, Friday, January 5, 2018.

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!**

<u>Please make sure to apply the following guidelines:</u>

1. On each page you turn in, write your assigned code number. Don't write your name on any page.

2. Start each problem on a new page.

Problem 1.

Let $\varphi_X(t), t \in \mathcal{R}$ denote the characteristic function (c.f.) of the random variable X. In what follows $\frac{\sin(0)}{0} = 1$.

a. Prove for each $T > 0$: $\frac{1}{2T}\int_{-T}^{T} \varphi_X(t)dt = E(\frac{\sin(TX)}{TX})$.

b. (i) Prove that $\frac{\sin(TX)}{TX}$ converges a.s. when $T \to \infty$ and identify its limit.

(ii). Prove that $\frac{1}{2T}\int_{-T}^{T} \varphi_X(t)dt$ converges when $T \to \infty$ and identify its limit.
Hint: $\frac{\sin(u)}{u}$, $u \in \mathcal{R}$ is a bounded function.

c. Find a formula that calculates $P(X = a), a \in \mathcal{R}$ from $\varphi_X(t), t \in \mathcal{R}$.
Hint: Work with $Y = X - a$. Express $\varphi_Y(t)$ in terms of $\varphi_X(t), t \in \mathcal{R}$.

Problem 2. Let $\{X_n\}, \{Y_n\}$, $n = 1, 2, \dots$ be 2 sequences of random variables defined on the same probability space. In what follows the symbol " $\Longrightarrow$ " is used for convergence in distribution.

a. Prove that if $X_n - Y_n \to 0$ in probability as $n \to \infty$ then for every **uniformly** continuous and bounded function $f: \mathcal{R} \to \mathcal{R}$ we have:
$$E\big(f(X_n)\big) - E\big(f(Y_n)\big) \to 0 \text{ as } n \to \infty.$$
Hint. Recall: f is uniformly continuous if for each $\epsilon > 0$ there is $\delta > 0$ so that $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$.

b. Prove: If $X_n \Longrightarrow X$ and $X_n - Y_n \to 0$ in probability, then $Y_n \Longrightarrow X$.
Hint: To prove convergence in distribution it is enough to work with uniformly continuous and bounded functions.

c. Let $\{X_{n,k}\}_{k=1}^{n}$, $n = 1, 2, \dots$ be a triangular array of random variables that are row-wise independent and identically distributed with
$P\big(X_{n,k} = 0\big) = 1 - \frac{1}{n} - \frac{1}{n^{1.5}}$, $P\big(X_{n,k} = 1\big) = \frac{1}{n}$ and $P\big(X_{n,k} = 2\big) = \frac{1}{n^{1.5}}$
for $1 \le k \le n$.

Let $S_n = \sum_{k=1}^{n} X_{n,k}$ and $T_n = \sum_{k=1}^{n} X_{n,k} \cdot 1_{\{X_{n,k} < 2\}}$.

(i) Prove that $S_n - T_n \to 0$ in probability as $n \to \infty$.

(ii) Prove that $S_n \Longrightarrow S$ as $n \to \infty$ and identify the distribution of S.

Problem 3.

Let $\{X_k\}, k = 1, 2, \ldots$ be a sequence of independent random variables.

a. Prove that if $\sum_{k=1}^{\infty} P(|X_k| > 1) = \infty$ then $\sum_{k=1}^{\infty} X_k$ **diverges** a.s.

For the rest of the problem we assume that for each $k = 1, 2, \ldots$ $|X_k| \le 1$, a.s.

b. Assume that $\sum_{k=1}^{\infty} \operatorname{var}(X_k) = \infty$.

(i) Prove: $\frac{\sum_{k=1}^{n} X_k - E(X_k)}{\sqrt{\sum_{k=1}^{n} \operatorname{var}(X_k)}}$ converges in distribution to standard normal distribution.

(ii) Prove that if in addition to the assumption $\sum_{k=1}^{\infty} \operatorname{var}(X_k) = \infty$ we also assume that $\sum_{k=1}^{\infty} X_k$ converges a.s. then $\frac{-\sum_{k=1}^{n} E(X_k)}{\sqrt{\sum_{k=1}^{n} \operatorname{var}(X_k)}}$ converges in distribution to a standard normal distribution. Why does this lead to the conclusion that if $\sum_{k=1}^{\infty} X_k$ converges a.s. then $\sum_{k=1}^{\infty} \operatorname{var}(X_k) < \infty$?

c. Prove, by using part b, that if $\sum_{k=1}^{\infty} X_k$ converges a.s. then $\sum_{k=1}^{\infty} X_k - E(X_k)$ converges a.s. and $\sum_{k=1}^{\infty} E(X_k)$ converges.

Problem 4.

Let $\{X, X_n\}, n = 1, 2, \ldots$ be a sequence of random variables whose values are in $\mathcal{Z}$ (the integers). Recall that with the notation $a^+ = \max\{a, 0\}$, $a^- = \max\{-a, 0\}$, we get

$a = a^+ - a^-, |a| = a^+ + a^-$, $a \in \mathcal{R}$.

a. Prove: (i) $\sum_{k\in\mathcal{Z}} [P(X = k) - P(X_n = k)]^+ = \sum_{k\in\mathcal{Z}} [P(X = k) - P(X_n = k)]^-$

(ii) $\sum_{k\in\mathcal{Z}} |P(X = k) - P(X_n = k)| = 2 \cdot \sum_{k\in\mathcal{Z}} [P(X = k) - P(X_n = k)]^+$

b. Find a sequence $a_k \ge 0, k \in \mathcal{Z}$ so that the following holds:

(1) $a_k \ge [P(X = k) - P(X_n = k)]^+, k \in \mathcal{Z}$, and

(2) $\sum_{k\in\mathcal{Z}} a_k < \infty$.

c. Prove that if $X_n \Longrightarrow X$ (converge in distribution as $n \to \infty$) then $\sum_{k\in\mathcal{Z}} |P(X = k) - P(X_n = k)| \to 0$, as $n \to \infty$.

Hint. Prove first that $X_n \Longrightarrow X$ implies $P(X_n = k) \to P(X = k), k \in \mathcal{Z}$.

Problem 5.

In this problem $n = 2,3,\dots$ Let (G_n) denote an integer-valued sequence of random variables defined on a probability space with $G_2 \equiv 1$. In what follows $G_{n+1} \in \{k, k+1\}$ if $G_n = k, k \in \mathcal{Z}$ which leads to $G_n \in \{1, \dots, n-1\}$. Using the notations:

$$X_n \equiv \frac{G_n}{n}, \qquad \mathcal{F}_n \equiv \sigma\{G_k, k = 2, \dots, n\}$$

and letting $P_{\mathcal{F}_n}$ represent conditional probability given $\mathcal{F}_n$, we assume:

$$P_{\mathcal{F}_n}(G_{n+1} = G_n + 1) = X_n \text{ and } P_{\mathcal{F}_n}(G_{n+1} = G_n) = 1 - X_n.$$

a. (i) Prove that $(X_n, \mathcal{F}_n)$ is a martingale.
(ii) Prove that X_n converges a.s. as $n \to \infty$. Denote the limit random variable by X. Prove also that $E(|X_n - X|^p) \to 0,\ 1 \le p < \infty$.

b. Prove that G_n is uniformly distributed on $\{1, \dots, n-1\}$, namely:
$P(G_n = k) = \frac{1}{n-1}, k = 1, \dots, n-1$. Hint: use mathematical induction.

c. Identify the distribution of X. (observe: X is defined in part a(ii))

Problem 6.

Let $\{B_t : t \ge 0\}$ denote a standard Brownian motion. Let $T_a = \inf\{t:\ B_t = a\}, a > 0$.

a. (i) Prove: $P(T_a < t) = 2P\left(Z < \frac{-a}{\sqrt{t}}\right), a > 0$, where $Z \sim N(0,1)$.
(ii) Find explicitly $f_{T_a}(t), t > 0$ (the density of T_a) by differentiating the formula in (i).

b. Prove by using part a) that $T_a = a^2 T_1$ in distribution.

c. (i) It is proved in the book that $T_2 = T_1 + \tilde{T}_1$ in distribution, where $\tilde{T}_1$ is an independent copy of T_1. Also, it follows from part b) that $T_2 = 4T_1$ in distribution. What can we conclude about $E(T_1)$ from those two results?
(ii) What can be said about the relationship between the quartiles of T_1 and T_2.