2009 Final for Advanced Probability for Communications

Each problem consists of 5 subproblems, and weights 9 points. You are required to answer at least 3 self-selected subproblems among the 5 in order to obtain 9 points for each problem. In total, you may achieve 135 points. Good luck.

1. (Section 6)

   (a) For the example on slide 6-6, compute

   $$\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\bigcap_{i=k}^{\infty} A_i.$$

   (Hint: Let $B_k = \bigcap_{i=k}^{\infty} A_i$. Determine $\limsup_{n\to\infty} B_n$ based on the observations that $B_1 \subset B_2 \subset B_3 \subset \cdots$.)

   (b) Write down the representation of a set that consists all the elements that occurs only finitely many times in $\{A_n\}_{n=1}^{\infty}$. (Hint: Use set subtraction operation in addition to set intersection and set union operations.)

   (c) Is it possible that $P(\limsup_{n\to\infty} A_n) = 0.2$? Justify your answer by creating an example of sequences $A_1, A_2, A_3, \ldots$ such that the above statement is true. (Hint: Make these sets extremely dependent.)

   (d) Give that the sequence of sets $A_1, A_2, A_3, \cdots$ and probability measure P satisfy the conditions in the second Borel-Cantelli lemma. Create a new sequence based on $A_1, A_2, A_3, \cdots$ such that they do not satisfy the conditions in the second Borel-Cantelli lemma, but they satisfy the conditions in the refined second Borel-Cantelli lemma. You need to justify your answer. (Hint: Modify the sequence $A_1, A_2, \ldots$, which satisfy the conditions in the second Borel-Cantelli lemma, into a new sequence such that the new sequence is not i.i.d. For example, add one or two sets at the beginning. Note that it is possible that $A_1 = A_2 = A_3 = \cdots = \Omega$, namely, the sample space.)

   (e) For a sequence of binary 0-1 digits, $b_1, b_2, b_3, \cdots$, what can we say about the limit inferior and limit superior based on binary-logical operations:

$$\bigoplus_{n=1}^{\infty} \bigotimes_{k=n}^{\infty} b_k \quad \text{and} \quad \bigotimes_{n=1}^{\infty} \bigoplus_{k=n}^{\infty} b_k,$$

where $\bigotimes$ and $\bigoplus$ are the AND and OR Boolean operations, respectively. Which one is larger? (Hint: Define $a_n = \bigotimes_{k=n}^{\infty} b_k$ and $c_n = \bigoplus_{k=n}^{\infty} b_k$. Then, notice that $a_1, a_2, a_3, \cdots$ is monotonically nondecreasing and $c_1, c_2, c_3, \cdots$ is monotonically nonincreasing.)

Solution.

(a) $\limsup_{n\to\infty} B_n$ consists of all the ω's that occur i.o. in B_n. Since $B_1 \subset B_2 \subset B_3 \subset \cdots$, an element ω occurs i.o. in $\{B_n\}_{n=1}^{\infty}$ if, and only if, $\omega \in B_k = \bigcap_{i=k}^{\infty} A_i$ for a given k. So, $\omega \in \limsup_{n\to\infty} B_n$ implies that $\omega \in A_k$ and $\omega \in A_{k+1}$ and $\omega \in A_{k+2}$ and $\cdots$ for some k, i.e., $\omega = .d_1 d_2 \cdots d_{k-1}000000000\cdots$. As a result,

$$\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \bigcap_{i=k}^{\infty} A_i = \left\{ \omega \in [0,1) : \omega = \frac{M}{2^j} \text{ for some integers } M \text{ and } j \right\}.$$

(b) $\bigcup_{n=1}^{\infty} A_n - \limsup_{n\to\infty} A_n$.

(c) By the first Borel-Cantelli lemmas, we should have

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = \infty.$$

By the second Borel-Cantelli lemmas, we should have

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) < \infty$$

or $\{A_n\}_{n=1}^{\infty}$ are not independent. This concludes that the only situation that makes valid

$$P(\limsup_{n\to\infty} A_n) = 0.2$$

is that $\{A_n\}_{n=1}^{\infty}$ are dependent with

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = \infty.$$

An example is that $A_1 = A_2 = \cdots = A_n = \cdots$ and $P(A_1) = 0.2$. Note that A_1 cannot be independent of A_1 because its probability is not either 1 or 0.

(d) Create a new sequence as $B_1, B_2, B_3, \ldots$ such that $B_j = A_{j-2}$ for $j = 3$, and $B_1 = B_2 = U$, and $P(U) = 1/2$. Since $P(B_1 \cap B_2) \neq P(B_1)P(B_2)$, $B_1, B_2, B_3, \ldots$ cannot be independent events because

$$\frac{1}{2} = P(U) = P(B_1 \cap B_2) \neq P(B_1)P(B_2) = P^2(U) = \frac{1}{4}.$$

(Note: You cannot define a sequence like $B_1, A_1, \ldots$ because it is possible that $A_1 = A_2 = A_3 = \cdots = \Omega$, namely, the sample space. In such case, $B_1, A_1, A_2, \ldots$ are still independent events.) Then, we compute by letting $u_n = \sum_{k=1}^{n-2} P(A_k)$,

$$\begin{aligned}
\sum_{j=1}^{n}\sum_{k=1}^{n} P(B_j \cap B_k) &= P(B_1 \cap B_1) + P(B_1 \cap B_2) + P(B_2 \cap B_1) \\
&+ P(B_2 \cap B_2) + 2\sum_{j=1}^{n} P(B_1 \cap A_j) + 2\sum_{j=1}^{n} P(B_2 \cap A_j) \\
&+ \sum_{j=1}^{n}\sum_{k=1}^{n} P(A_j \cap A_k) \\
&= 4P(U) + 4\sum_{j=1}^{n-2} P(U \cap A_j) + \sum_{j=1}^{n-2}\sum_{k=1}^{n-2} P(A_j)P(A_k) \\
&\leq 4P(U) + 4\sum_{j=1}^{n-2} P(A_j) + \sum_{j=1}^{n-2}\sum_{k=1}^{n-2} P(A_j)P(A_k) \\
&= 2 + 2u_n + u_n^2,
\end{aligned}$$

and

$$\begin{aligned}
\left(\sum_{k=1}^{n} P(B_k)\right)^2 &= \left(2P(U) + \sum_{k=1}^{n-2} P(A_k)\right)^2 \\
&= 4P^2(U) + 4P(U)\sum_{k=1}^{n-2} P(A_k) + \sum_{j=1}^{n-2}\sum_{k=1}^{n-2} P(A_j)P(A_k) \\
&= 1 + 2u_n + u_n^2.
\end{aligned}$$

Since

$$\liminf_{n\to} \frac{\sum_{j=1}^{n}\sum_{k=1}^{n} P(B_j \cap B_k)}{\left(\sum_{k=1}^{n} P(B_k)\right)^2} \le \liminf_{n\to\infty} \frac{2+2u_n+u_n^2}{1+2u_n+u_n^2} = 1,$$

and

$$\sum_{k=1}^{n} P(B_k) = P(B_1) + P(B_2) + \sum_{k=1}^{n-2} P(A_k) = \infty,$$

the conditions in the refined second Borel-Cantelli lemma are satisfied.

(e) The sequence $a_1, a_2, a_3, \ldots$ is equal to

$$\begin{cases} \text{all zero sequence}, & \text{if 0 occurs i.o. in } b_1, b_2, b_3, \ldots \\ \underbrace{00\cdots00}_{m \text{ zeros}}\underbrace{111\cdots}_{\text{all ones}} & \text{otherwise} \end{cases}$$

where m is the largest integer such that $b_m = 0$. Hence,

$$\bigoplus_{n=1}^{\infty}\bigotimes_{k=n}^{\infty} b_k \quad = \quad \begin{cases} 0, & \text{if 0 occurs i.o. in } b_1, b_2, b_3, \ldots \\ 1, & \text{otherwise} \end{cases}$$

The sequence $c_1, c_2, c_3, \ldots$ is equal to

$$\begin{cases} \text{all one sequence}, & \text{if 1 occurs i.o. in } b_1, b_2, b_3, \ldots \\ \underbrace{11\cdots11}_{m \text{ ones}}\underbrace{000\cdots}_{\text{all zeros}} & \text{otherwise} \end{cases}$$

where m is the largest integer such that $b_m = 1$. Hence,

$$\bigotimes_{n=1}^{\infty}\bigoplus_{k=n}^{\infty} b_k \quad = \quad \begin{cases} 1, & \text{if 1 occurs i.o. in } b_1, b_2, b_3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

Considering the three possible cases, i.e., (i) both 0 and 1 occur i.o., (ii) only 0 occurs i.o., (iii) only 1 occurs i.o., (notably, at least one of 0 and 1 should occur i.o.), we obtain that

$$\left(\limsup_{n\to\infty} b_k = \right) \bigotimes_{n=1}^{\infty}\bigoplus_{k=n}^{\infty} b_k \ge \bigoplus_{n=1}^{\infty}\bigotimes_{k=n}^{\infty} b_k \left(= \liminf_{n\to\infty} b_n \right).$$

□

2. (Section 9)

   (a) For a random variable Y and a parameterized non-negative function $f_t(y)$, where both Y and y take values in $\Re$, prove that

$$\Pr[Y \in \mathcal{A}] = E[f_t(Y)] \int_{\mathcal{A}} f_t^{-1}(y) dP_{Y^{(t)}}(y),$$

   where $Y^{(t)}$ is the random variable having distribution

$$dP_{Y^{(t)}}(y) \triangleq \frac{f_t(y) dP_Y(y)}{E[f_t(Y)]},$$

   and $\mathcal{A}$ is a subset of $\Re$.

   (b) By following (a), suppose we can find a function $f_t(\cdot)$ such that $E[f_t(Y)] = 10^{-5}$ (for which the first moment of a function of Y is possibly easier to obtain than the probability $\Pr[Y \in \mathcal{A}]$.) Can you give a law-of-large-numbers-like formula to estimate $\Pr[Y \in \mathcal{A}]$ based on the samples $y_1, y_2, y_3, \ldots, y_n$ that are drawn according to distribution $P_{Y^{(t)}}$? Comment on this indirect approach to estimate $\Pr[Y \in \mathcal{Y}]$.

   (c) Let

$$f_n(X) = \sum_{\substack{ik+j\ell=n,\ ik\geq j\ell \\ i\geq j\geq 0,\ k,\ell\geq 1}} a_{i,k,j,\ell} E^k[X^i] E^\ell[X^j]$$

   be a function of random variable X with $a_{n,1,0,1} = 1$. Find the coefficients such that $f_2(X) = f_2(X - x)$ for every $x \in \Re$ and every random variable X. Is variance function of X, i.e., $f_2(X) = E[X^2] - E^2[X]$, the only second-order function that remains unchanged with a time shift x?

   (d) One of the coefficients that make $f_3(X) = f_3(X - x)$ for every X and x is

$$f_3(X) = E[X^3] - 3E[X^2]E[X] + 2E^3[X].$$

   Can you find another set of coefficients with $a_{3,1,0,1} = 1$, satisfying $f_3(X) = f_3(X - x)$ for every X and x? Justify your answer. (Hint: Since this condition must be true for every X and x, choose specific simple X and x to verify the possibility. For example, binary X.)

(e) Prove that for $E[U]=0$, $E[U^2]>0$, $p_1>1$ and integer $k\geq 1$,

$$\Pr[U\geq 0]\geq \frac{E^{p_1/(p_1-1)}[U^{2k}]}{2^{p_1/(p_1-1)}E^{1/(p_1-1)}\left[U^{2kp_1}\right]},$$

provided kp_1 is an integer.

Solution.

(a)

$$\begin{aligned}\Pr\left[Y\in\mathcal{A}\right] &= \int_{\mathcal{A}} dP_Y(y)\\ &= \int_{\mathcal{A}} \frac{dP_Y(y)}{dP_{Y^{(t)}}(y)} dP_{Y^{(t)}(y)}\\ &= E[f_t(Y)]\int_{\mathcal{A}} f_t^{-1}(y)dP_{Y^{(t)}}(y).\end{aligned}$$

(b) By (a), we can estimate $\Pr[Y\in\mathcal{A}]$ by:

$$\frac{f_t^{-1}(y_1)\cdot\mathbf{1}\{y_1\in\mathcal{A}\}+\cdots+f_t^{-1}(y_n)\cdot\mathbf{1}\{y_n\in\mathcal{A}\}}{n}\times 10^{-5}.$$

Since $f_t(\cdot)$ is a known, chosen function, we can accurate obtain its value for any given y. So, as long as $\Pr[Y^{(t)}\in\mathcal{A}]$ is sufficiently enlarged, there will be a lot of samples lies in $\mathcal{A}$ (anticipatedly, more often than the samples drawn according to distribution P_Y). This will accordingly reduce the number of the Monte Carlo simulation runs for the estimate of $\Pr[Y\in\mathcal{A}]$.

(c)

$$\begin{aligned}&a_{2,1,0,1}E[X^2]+(a_{1,2,0,1}+a_{1,1,1,1})E^2[X]\\ &\quad = a_{2,1,0,1}E[(X-x)^2]+(a_{1,2,0,1}+a_{1,1,1,1})E^2[X-x]\end{aligned}$$

implies

$$a_{2,1,0,1}(E[X^2]-E[(X-x)^2]+(a_{1,2,0,1}+a_{1,1,1,1})(E^2[X]-E^2[X-x])=0,$$

which in turns implies

$$(a_{2,1,0,1}+a_{1,2,0,1}+a_{1,1,1,1})(2xE[X]-x^2)=0.$$

Hence, the coefficients that make $f_2(X) = f_2(X-x)$ for every x and X is $a_{2,1,0,1} = 1$ and $a_{1,2,0,1} + a_{1,1,1,1} = -1$. This concludes that the function satisfies $f_2(X) = f_2(X-x)$ for every x and $a_{n,1,0,1} = 1$ is unique.

(d) Give $f_3(X) = E[X^3] + a_{2,1,1,1}E[X^2]E[X] + a_{1,2,1,1}E^2[X]E[X]$. We need to find $a_{2,1,1,1}$ and $a_{1,2,1,1}$ such that $f_3(X) = f_3(X-x)$ for every random variable X and constant x. So, it must be true for X binary, equal probable over $\{-1,1\}$, and $x = -1,-2$. This random variable X gives

$$E[X^3] = 0, E[X^2] = 1, E[X] = 0,$$

$$E[(X+1)^3] = 4, E[(X+1)^2] = 2, E[X+1] = 1,$$

and

$$E[(X+2)^3] = 14, E[(X+2)^2] = 5, E[X+2] = 2,$$

This results in:

$$0 = 4 + 2a_{2,1,1,1} + a_{1,2,1,1} = 14 + 10a_{2,1,1,1} + 8a_{1,2,1,1}.$$

As a result, $a_{2,1,1,1} = -3$ and $a_{1,2,1,1} = 2$. This means that there exists no other set of coefficients with $f_3(X) = f_3(X-x)$ for every X and x.

(e) Let $U^+ = U \cdot I_{[U \geq 0]}$ and $U^- = -U \cdot I_{[U<0]}$, where $I_{[\cdot]}$ is an indicator random variable that equals 1 when the event is true, and 0 when the event is false. By their definitions, U^+ and U^- are both non-negative. Then by Hölder's inequality,

$$\begin{aligned} E\left[(U^+)^{2k}\right] &= E\left[U^{2k} \cdot I_{[U\geq 0]}\right] \leq E^{1/p_1}\left[U^{2kp_1}\right] E^{1/q_1}\left[I^{q_1}_{[U\geq 0]}\right] \\ &= E^{1/p_1}\left[U^{2kp_1}\right] E^{1/q_1}\left[I_{[U\geq 0]}\right], \end{aligned}$$

and

$$\begin{aligned} E\left[(U^-)^{2k}\right] &= E\left[\left|U^-\right|^{1/p_2}\left|U^-\right|^{2k-1/p_2}\right] \\ &\leq E^{1/p_2}\left[\left(\left|U^-\right|^{1/p_2}\right)^{p_2}\right] E^{1/q_2}\left[\left(\left|U^-\right|^{2k-1/p_2}\right)^{q_2}\right] \\ &= E^{1/p_2}\left[\left|U^-\right|\right] E^{1/q_2}\left[\left|U^-\right|^{(2k-1/p_2)q_2}\right] \\ &= E^{1/p_2}\left[U^-\right] E^{1/q_2}\left[U^{(2k-1/p_2)q_2}\right]. \end{aligned}$$

Also by Hölder's inequality, together with $E[U] = 0$ and $U = U^+ - U^-$,

$$\begin{aligned} E[U^-] &= E[U^+] \\ &= E[U \cdot I_{[U\geq 0]}] \\ &= E[|U \cdot I_{[U\geq 0]}|] \\ &\leq E^{1/p_3}[U^{p_3}]\, E^{1/q_3}[I^{q_3}_{[U\geq 0]}] \\ &= E^{1/p_3}[U^{p_3}]\, E^{1/q_3}[I_{[U\geq 0]}]. \end{aligned}$$

Hence,

$$\begin{aligned} E[U^{2k}] &= E[(U^+)^{2k}] + E[(U^-)^{2k}] \\ &\leq E^{1/p_1}\left[U^{2kp_1}\right] E^{1/q_1}\left[I_{[U\geq 0]}\right] + E^{1/p_2}\left[U^-\right] E^{1/q_2}\left[U^{(2k-1/p_2)q_2}\right] \\ &\leq E^{1/p_1}\left[U^{2kp_1}\right] E^{1/q_1}\left[I_{[U\geq 0]}\right] \\ &\quad + \left(E^{1/p_3}[U^{p_3}]\, E^{1/q_3}[I_{[U\geq 0]}]\right)^{1/p_2} E^{1/q_2}\left[U^{(2k-1/p_2)q_2}\right] \\ &= E^{1/p_1}\left[U^{2kp_1}\right] E^{1/q_1}\left[I_{[U\geq 0]}\right] \\ &\quad + E^{1/(p_2p_3)}\left[U^{p_3}\right] E^{1/q_2}\left[U^{(2k-1/p_2)q_2}\right] E^{1/(p_2q_3)}\left[I_{[U\geq 0]}\right] \\ &= \left(E^{1/p_1}\left[U^{2kp_1}\right] + E^{1/(p_2p_3)}\left[U^{p_3}\right] E^{1/q_2}\left[U^{(2k-1/p_2)q_2}\right]\right) E^{1/q_1}\left[I_{[U\geq 0]}\right], \end{aligned}$$

where we will choose

$$q_1 = p_2q_3 \text{ and } 2kp_1 = p_3 = (2k - 1/p_2)q_2. \qquad (1)$$

Accordingly,

$$\Pr[U \geq 0] \geq \frac{E^{p_1/(p_1-1)}[U^{2k}]}{2^{p_1/(p_1-1)} E^{1/(p_1-1)}\left[U^{2kp_1}\right]},$$

where $p_2 = (p_1 - 1/(2k))/(p_1 - 1)$ and $p_3 = 2kp_1$ will satisfy (1). □

3. (Section 20)

   (a) For a subset $\mathcal{X}$ of $[0, 1)$, and a real number $c \in (0, 1]$, define $\mathcal{X} \circledcirc c$ as $\{y \in (0, 1] : y = (x + c) \bmod 1 \text{ for some } x \in \mathcal{X}\}$. Prove that if $\mathcal{X}$ is a Borel set, then $\mathcal{X} \circledcirc c$ is also a Borel set. (Hint: $\mathcal{X} = \bigcup_k I_k$, where $\{I_k\}$ are disjoint intervals.)

(b) Vitali shows that $[0,1) = \cap_{r\in\mathbb{Q}}\mathcal{H} \circledcirc r$ and each $\mathcal{H} \circledcirc r$ is disjoint. So, for any measure λ satisfying "countable additivity", we have:

$$\lambda([0,1)] = \sum_{r\in\mathbb{Q}} \lambda(\mathcal{H} \circledcirc r).$$

What will $\lambda([0,1))$ be if λ is invariant to "$\circledcirc$" operation, i.e., $\lambda(\mathcal{A}) = \lambda(\mathcal{A} \circledcirc x)$ for any real set $\mathcal{A} \subset [0,1)$ and any number $x \in [0,1)$. (Note: Please replace $(0,1]$ by $[0,1)$ in Vitali's example in the lecture slides since by convention, $(1+1) \bmod 1 = 0$.)

(c) Provide an example of a non-$\mathcal{B}/\mathcal{B}$-measurable function.

(d) Complete the assignment of P such that X_1 and X_2 are independent in the example in slide 20-22, provided $P(\{\blacktriangle, \square\}) = (\{\blacksquare, \blacklozenge\}) = 1/4$.

(e) For a random process $\{X_t, t \in [0,1)\}$ defined over the probability space $(\Omega = [0,1), \mathcal{F} = \mathcal{B}([0,1)), \lambda = \text{Lebesque measure})$, is the set $\{\omega \in \Omega : \min_{t\in[0,1)} X_t(\omega) < 1\}$ an event, where $\mathcal{B}([0,1))$ consists of all the Borel sets in $[0,1)$? Justify your answer.

Solution.

(a) $(I_k + c)$mod 1 remains an interval if the right margin plus c is less than 1. But, $(I_k + c)$mod 1 becomes two intervals if the right margin plus c is larger than or equal to 1. Also, if I_i and I_j are disjoint intervals, then, $I_i \circledcirc c$ and $I_j \circledcirc c$ are also disjoint. Thus, $\mathcal{X} \circledcirc c$ is still a Borel set.

(b) $\lambda([0,1))$ is sum of countably many equal quantities. So, it can only be either zero, $-\infty$, or $+\infty$.

(c) For a given $\mathcal{H}$ in Vitali's example, define $f : [0,1) \to [0,1)$ such that $f(x) = r$, if $x \in \mathcal{H} \circledcirc r$. This is a function but not $\mathcal{B}/\mathcal{B}$-measurable.

(d) The sets of $\{\omega \in \Omega : X_1(\omega) = i \text{ and } X_2(\omega) = j\}$ can be listed as follows:

	$X_2 = 1$	$X_2 = 2$
$X_1 = 1$	$\{\blacktriangle, \square\}$	$\{\lozenge\}$
$X_1 = 2$	$\{\blacktriangledown\}$	$\{\blacksquare, \blacklozenge\}$

Hence,

	$X_2 = 1$	$X_2 = 2$
$X_1 = 1$	$1/4$	$P(\{\Diamond\})$
$X_1 = 2$	$P(\{\blacktriangledown\})$	$1/4$

By independence, we can obtain that $P(\{\Diamond\}) = P(\{\blacktriangledown\}) = 1/4$.

(e) Define $X_t(\omega) = 0$ if $t \in \mathcal{H}$ and $\omega = t$, and $X_t(\omega) = 1$, otherwise, where $\mathcal{H}$ is the Vitali set. Then,

$$\begin{aligned}
&\{\omega \in \Omega : \min_{t\in[0,1)} X_t(\omega) < 1\} \\
&= \cup_{t\in[0,1)}\{\omega \in \Omega : X_t(\omega) < 1\} \\
&= \cup_{t\in\mathcal{H}}\{\omega \in \Omega : X_t(\omega) < 1\}, \text{ since } \{\omega \in \Omega : X_t(\omega) < 1\} = \emptyset \text{ for } t \notin \mathcal{H} \\
&= \mathcal{H}.
\end{aligned}$$

Therefore, $\{\omega \in \Omega : \min_{t\in[0,1)} X_t(\omega) < 1\}$ is not an event because if it were, it should be Lebesque measurable (by definition of probability space). □

4. (Section 21)

(a) Prove that for a *simple* nonnegative random variable X,

$$E[X] = \int_0^\infty \Pr[X > t]dt.$$

(b) Based on (a) and Theorem 13.5a (cf. Slide 20-19), prove that for a nonegative random variable X (possibly non-simple, and possibly having no density),

$$E[X] = \int_0^\infty \Pr[X > t]dt.$$

(Hint: Use monotone convergence theorem such as "if $X_n(\omega) \uparrow X(\omega)$ for every $\omega \in \Omega$, then $\lim_{n\to\infty} \int_\Omega X_n(\omega)dP(\omega) = \int_\Omega X(\omega)dP(\omega)$." Or "if $f_n(x) \uparrow f(x)$ for every $x \geq 0$, then $\lim_{n\to\infty} \int_0^\infty f_n(x)dx = \int_0^\infty f(x)dx$.")

(c) Why $\int_\mathcal{X} \Pr[X < t]dt = \int_\mathcal{X} \Pr[X \leq t]dt$ for any linear Borel set $\mathcal{X}$? In other words, why $\int_\mathcal{X} \Pr[X = t]dt = 0$?

(d) Represent $E[X^{2k}]$ for integer $k \geq 1$ in terms of the cdf of X, i.e., $\Pr[X \leq t]$. Also, represent the moment generating function of X in terms of the cdf of X.

(e) Give the sufficient and necessary condition under which the generalized Hölder's inequality, i.e., for $p_j > 1$ with $\sum_{j=1}^{n} 1/p_j = 1$,

$$E\left[\left|\prod_{j=1}^{n} X_j\right|\right] \le \prod_{j=1}^{n} E^{1/p_j}\left[|X_j|^{p_j}\right]$$

becomes equality, provided that $\{X_j\}_{j=1}^{n}$ are pair-wise independent non-negative random variables with identical marginal distribution and $E[X^p] > 0$ for $p > 1$.

Solution.

(a) For a simple random variable that takes values from $\{x_1, x_2, \ldots, x_m\}$ with probabilities $\Pr[X = x_j] = p_j$, its mean is equal to $E[X] = \sum_{j=1}^{m} p_j x_j$. Also, $\Pr[X > t] = \sum_{j=1}^{m} p_j \cdot I_{[x_j > t]}$, where $I_{[\cdot]}$ is the set indicator function. Hence,

$$\begin{aligned}
\int_0^\infty \Pr[X > t] dt &= \int_0^\infty \sum_{j=1}^{m} p_j \cdot I_{[x_j > t]} dt \\
&= \sum_{j=1}^{m} p_j \int_0^\infty I_{[x_j > t]} dt \\
&= \sum_{j=1}^{m} p_j \int_0^{x_j} dt \\
&= \sum_{j=1}^{m} p_j x_j = E[X].
\end{aligned}$$

(b) Let X be defined over the probability space $(\Omega, \mathcal{F}, P)$. By Theorem 13.5a, there exists a sequence of simple non-negative random variables $\{X_n\}_{n\ge 1}$ that are defined over the same probability space as X such that

$$X_n(\omega) \uparrow X(\omega)$$

for every $\omega \in \Omega$. Now, because $X_n(\omega) \le X_{n+1}(\omega)$,

$$\begin{aligned}
\Pr[X_n > t] &= P(\{\omega \in \Omega : X_n(\omega) > t\}) \\
&\le P(\{\omega \in \Omega : X_{n+1}(\omega) > t\}) \\
&= \Pr[X_{n+1} > t]
\end{aligned}$$

Then, by monotone convergence theorem,

$$\lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \int_0^\infty \Pr[X_n > t]dt = \int_0^\infty \Pr[X > t]dt$$

and

$$\lim_{n\to\infty} E[X_n] = \lim_{n\to\infty} \int_\Omega X_n(\omega)dP(\omega) = \int_0^\infty X(\omega)dP(\omega) = E[X].$$

Thus, the desired theorem is proved.

(c) Define $\mathcal{X}_0 = \{t \in \mathcal{X} : \Pr[X = t] = 0\}$. Then, the set $\mathcal{X} - \mathcal{X}_0$ consists of at most countable number of elements, whose probability mass is bounded above by 1. Hence, by $\int_{\mathcal{X}-\mathcal{X}_0} dt = 0$ and $\Pr[X = t] \geq 0$, we have:

$$\begin{aligned}
\int_{\mathcal{X}} \Pr[X = t]dt &= \int_{\mathcal{X}_0} \Pr[X = t]dt + \int_{\mathcal{X}-\mathcal{X}_0} \Pr[X = t]dt \\
&\leq \int_{\mathcal{X}_0} 0 \cdot dt + \int_{\mathcal{X}-\mathcal{X}_0} 1 \cdot dt \\
&= 0
\end{aligned}$$

and $\int_{\mathcal{X}} \Pr[X = t]dt \geq 0$. Accordingly, $\int_{\mathcal{X}} \Pr[X = t]dt = 0$.

(d)

$$\begin{aligned}
E[X^{2k}] &= \int_0^\infty \Pr[X^{2k} > t]dt \\
&= \int_0^\infty (\Pr[X > t^{1/(2k)}] + \Pr[X < -t^{1/(2k)}])dt \\
&= \int_0^\infty \Pr[X > t^{1/(2k)}]dt + \int_0^\infty \Pr[X < -t^{1/(2k)}]dt \\
&= \int_0^\infty (1 - \Pr[X \leq t^{1/(2k)}])dt + \int_0^\infty \Pr[X \leq -t^{1/(2k)}]dt
\end{aligned}$$

and

$$\begin{aligned}
E[e^{sX}] &= \int_0^\infty \Pr[e^{sX} > t]dt \\
&= \int_0^\infty \Pr[sX > \log(t)]dt \\
&= \begin{cases} \int_0^\infty \Pr[X > \frac{1}{s}\log(t)]dt, & s > 0 \\ \int_0^\infty \Pr[0 > \log(t)]dt, & s = 0 \\ \int_0^\infty \Pr[X < \frac{1}{s}\log(t)]dt, & s < 0 \end{cases} \\
&= \begin{cases} \int_0^\infty (1 - \Pr[X \leq \frac{1}{s}\log(t)])dt, & s > 0 \\ \int_0^1 dt, & s = 0 \\ \int_0^\infty \Pr[X \leq \frac{1}{s}\log(t)]dt, & s < 0 \end{cases} \\
&= \begin{cases} \int_0^\infty (1 - \Pr[X \leq \frac{1}{s}\log(t)])dt, & s > 0 \\ 1, & s = 0 \\ \int_0^\infty \Pr[X \leq \frac{1}{s}\log(t)]dt, & s < 0 \end{cases}
\end{aligned}$$

(e) In its most general form, equality holds for Hölder's inequality if, and only if,

$$\Pr\left[\frac{|X_j|^{p_j}}{E[|X_j|^{p_j}]} = \text{constant for all } j\right] = 1.$$

Hence, for any $i \neq j$, we derive by pair-wise independence, nonnegativity and identical marginal that

$$
\begin{aligned}
1 &= \Pr\left[\frac{X_i^{p_i}}{E[X_i^{p_i}]} = \frac{X_j^{p_j}}{E[X_j^{p_j}]}\right] \\
&= \int_{-\infty}^{\infty} \Pr\left[\frac{X_i^{p_i}}{E[X_i^{p_i}]} = \frac{x_j^{p_j}}{E[X_j^{p_j}]}\right] dP_{X_j}(x_j) \\
&= \int_{-\infty}^{\infty} \Pr\left[X_i = x_j^{p_j/p_i} \frac{E^{1/p_i}[X_i^{p_i}]}{E^{1/p_i}[X_j^{p_j}]}\right] dP_{X_j}(x_j) \\
&= \int_{\mathcal{X}_0} \Pr\left[X_i = x_j^{p_j/p_i} \frac{E^{1/p_i}[X_i^{p_i}]}{E^{1/p_i}[X_j^{p_j}]}\right] dP_{X_j}(x_j) \\
&\quad + \int_{\Re-\mathcal{X}_0} \Pr\left[X_i = x_j^{p_j/p_i} \frac{E^{1/p_i}[X_i^{p_i}]}{E^{1/p_i}[X_j^{p_j}]}\right] dP_{X_j}(x_j) \\
&= \int_{\Re-\mathcal{X}_0} \Pr\left[X_i = x_j^{p_j/p_i} \frac{E^{1/p_i}[X_i^{p_i}]}{E^{1/p_i}[X_j^{p_j}]}\right] dP_{X_j}(x_j),
\end{aligned}
$$

where

$$
\mathcal{X}_0 = \left\{ x_j \in \Re : \Pr\left[X_i = x_j^{p_j/p_i} \frac{E^{1/p_i}[X_i^{p_i}]}{E^{1/p_i}[X_j^{p_j}]}\right] = 0 \right\}.
$$

Hence, X_j must be a discrete random variable because $\Pr[X_j \in \Re - \mathcal{X}_0] = 1$ and $\Re - \mathcal{X}$ consists of at most countably many elements. As a result, by letting X having the same distribution as X_j to simplify the notation, we have:

$$
1 = \sum_{s \in \Re - \mathcal{X}_0} \Pr\left[X = s^{p_j/p_i} \frac{E^{1/p_i}[X^{p_i}]}{E^{1/p_i}[X^{p_j}]}\right] \Pr[X = s] \qquad (2)
$$

Now, for a random variable Y, which takes value $\Pr\left[X = s^{p_j/p_i} \frac{E^{1/p_i}[X^{p_i}]}{E^{1/p_i}[X^{p_j}]}\right]$ with probability $\Pr[X = s]$ (i.e.,

$$
\Pr\left(Y = \Pr\left[X = s^{p_j/p_i} \frac{E^{1/p_i}[X^{p_i}]}{E^{1/p_i}[X^{p_j}]}\right]\right) = \Pr[X = s]),
$$

we know that $0 \leq E[Y] \leq 1$ because $\Pr[0 \leq Y \leq 1] = 1$. (2) however tells us that $E[Y] = 1$. Therefore, $\Pr[Y = 1] = 1$. In other words, $\Pr[X = s] = 1$ for s satisfying $s^{p_j/p_i} \frac{E^{1/p_i}[X^{p_i}]}{E^{1/p_i}[X^{p_j}]} = s$ for any $i \neq j$, which is automatically true for deterministic X. Consequently, the desired condition is that X is deterministic with $\Pr[X = s] = 1$ for any $s > 0$. □

5. (Section 22)

   (a) If the mean of X_n does not exist for an i.i.d. sequence $\{X_n\}_{n\geq 1}$, then what can we say about

   $$\frac{X_1^+ + X_1^- + X_2^+ + X_2^- + \ldots + X_n^+ + X_n^-}{n}?$$

   (b) Prove that if $\{X_n\}_{n\geq 1}$ is i.i.d. with first and third moments being zero, then the third moment of $S_n = X_1 + X_2 + \ldots + X_n$ is also zero.

   (c) Suppose that $X_1, X_2, \ldots$ are independent with both first and third moments being zero and finite fourth moment (not necessarily identically distributed). Then for $\alpha > 0$,

   $$\Pr\left[\max_{1\leq k\leq n} |S_k| \geq \alpha\right] \leq \frac{1}{\alpha^4} E[S_n^4],$$

   where $S_n = X_1 + \cdots + X_n$.

   (d) Is it possible to construct a sequence of independent non-deterministic random variables $X_1, X_2, \ldots$ such that

   $$\Pr\left[\max_{1\leq k\leq n} |S_k| \geq 3\alpha\right] = 3 \max_{1\leq k\leq n} \Pr[|S_k| \geq \alpha]$$

   for every n and every $\alpha \geq 0$, where $S_n = X_1 + \cdots + X_n$? Justify your answer.

   (e) For a sequence of independent binary random variables taking values in $\{0, 1\}$, prove by means of Three-Series Theorem that $\sum_{n=1}^{\infty} X_n$ converges with probability 1 if, and only if, $\sum_{n=1}^{\infty} p_n < \infty$, where $p_n = \Pr[X_n = 1]$.

Solution.

(a) By the nonexistence of mean, we know that $E[X_n^+] = E[X_n^-] = \infty$. Hence, from the theorem on slide 22-11, both

$$\frac{X_1^+ + X_2^+ + \ldots + X_n^+}{n}$$

and

$$\frac{X_1^- + X_2^- + \ldots + X_n^-}{n}$$

converge with probability 1 to infinity. Hence, sum of the above two surely converges with probability 1 to infinity.

(b) For zero-mean independent random variables, Y and Z,

$$\begin{aligned} E[(Y+Z)^3] &= E[Y^3] + 3E[Y^2]E[Z] + 3E[Y]E[Z^2] + E[Z^3] \\ &= E[Y^3] + E[Z^3] \end{aligned}$$

Hence, the claim can be proved by induction using $E[S_n^3] = E[S_{n-1}^3] + E[X_n^3] = E[S_{n-1}^3]$.

(c) Define the event

$$A_k = \big[|S_1| < \alpha \wedge |S_2| < \alpha \wedge \cdots \wedge |S_{k-1}| < \alpha \wedge |S_k| \geq \alpha\big].$$

Since exactly one of $\{A_k\}_{k=1}^{\infty}$ is true,

$$
\begin{aligned}
E[S_n^4] &= E\left[S_n^4\left(I_{A_1}+I_{A_2}+\cdots+I_{A_n}+I_{A_{n+1}}+\cdots\right)\right] \\
&\geq E\left[S_n^4\left(I_{A_1}+I_{A_2}+\cdots+I_{A_n}\right)\right] \\
&= \sum_{k=1}^{n} E\left[S_n^4 I_{A_k}\right] \\
&= \sum_{k=1}^{n} E\left[(S_k+(S_n-S_k))^4 I_{A_k}\right] \\
&= \sum_{k=1}^{n} E\Big[\big(S_k^4+4S_k^3(S_n-S_k)+6S_k^2(S_n-S_k)^2 \\
&\quad +4S_k(S_n-S_k)^3+(S_n-S_k)^4\big) I_{A_k}\Big] \\
&\geq \sum_{k=1}^{n} E\left[\left(S_k^4+4S_k^3(S_n-S_k)+4S_k(S_n-S_k)^3\right) I_{A_k}\right] \\
&= \sum_{k=1}^{n} E\left[S_k^4 I_{A_k}+4S_k^3 I_{A_k}(S_n-S_k)+4S_k I_{A_k}(S_n-S_k)^3\right] \\
&= \sum_{k=1}^{n} \left(E\left[S_k^4 I_{A_k}\right]+4E\left[S_k^3 I_{A_k}(S_n-S_k)\right]+4E\left[S_k I_{A_k}(S_n-S_k)^3\right]\right) \\
&= \sum_{k=1}^{n} \left(E\left[S_k^4 I_{A_k}\right]+4E\left[S_k^3 I_{A_k}\right]E\left[S_n-S_k\right]+4E\left[S_k I_{A_k}\right]E\left[(S_n-S_k)^3\right]\right),
\end{aligned}
$$

where the last step follows from the independence between $S_k I_{A_k}$

and $S_n - S_k$. Continue the previous derivation:

$$\begin{aligned}
E[S_n^4] &\geq \sum_{k=1}^{n} \left(E\left[S_k^4 I_{A_k}\right] + 4E\left[S_k^3 I_{A_k}\right] E\left[S_n - S_k\right] + 4E\left[S_k I_{A_k}\right] E\left[(S_n - S_k)^3\right] \right) \\
&= \sum_{k=1}^{n} E\left[S_k^4 I_{A_k}\right] \\
&\geq \sum_{k=1}^{n} E\left[\alpha^4 I_{A_k}\right] \quad (I_{A_k} = 1 \text{ only when } |S_k| \geq \alpha) \\
&= \alpha^4 \sum_{k=1}^{n} \Pr[A_k] \\
&= \alpha^4 \Pr\left[\max_{1 \leq k \leq n} |S_k| \geq \alpha\right].
\end{aligned}$$

(d) The answer is no. Take $n = 1$ as an example. The condition gives that

$$\Pr\left[|X_1| \geq 3\alpha\right] = 3 \Pr[|X_1| \geq \alpha].$$

Then, by $\Pr[X_1 \geq 3\alpha] \leq 1$, we obtain that for every $\alpha \geq 0$,

$$\Pr[|X_1| \geq \alpha] \leq \frac{1}{3}.$$

Then, again, by $\Pr[X_1 \geq 3\alpha] \leq 1/3$, we obtain that for every $\alpha \geq 0$,

$$\Pr[|X_1| \geq \alpha] \leq \frac{1}{3^2}.$$

Repeating this process concludes that for everyt $\alpha \geq 0$

$$\Pr[|X_1| \geq \alpha] = 0,$$

which cannot be made valid by any random variable.

(e) For independent binary random variables, the three series become:

$$\sum_{n=1}^{\infty} \Pr[|X_n| > c] = \begin{cases} \sum_{n=1}^{\infty} p_n, & 0 < c < 1; \\ 0, & c \geq 1 \end{cases}$$

$$\sum_{n=1}^{\infty} E[X_n I_{[|X_n| \leq c]}] = \begin{cases} 0, & 0 < c < 1; \\ \sum_{i=1}^{\infty} p_n, & c \geq 1 \end{cases}$$

and

$$\sum_{n=1}^{\infty} \text{Var}[X_n I_{[|X_n|\le c]}] = \begin{cases} 0, & 0 < c < 1; \\ \sum_{n=1}^{\infty} p_n(1-p_n), & c \ge 1. \end{cases}$$

Thus, if $\sum_{n=1}^{\infty} p_n$ converges, then by Theorem 22.8, $\sum_{n=1}^{\infty} X_n$ converges with probability 1. On the contrary, if $\sum_{n=1}^{\infty} X_n$ converges with probability 1, $\sum_{n=1}^{\infty} p_n$ should converge for every $c > 0$; hence, $\sum_{n=1}^{\infty} p_n$ converges. □

6. (Section 25)

   (a) Prove that if two nonnegative random variables X and Y satisfies $\Pr[|X-Y| > \epsilon] = 0$ for a given ϵ, then $|E[X] - E[Y]| \le \epsilon$.

   (b) Regarding the theorem on slide 25-22, provide an example that $\Pr[X \in \mathcal{D}_h] > 0$ and $h(X_n) \not\Rightarrow h(X)$ even if $X_n \Rightarrow X$.

   (c) Regarding Theorem 25.8, give an unbounded continuous function h such that $\lim_{n\to\infty} \int_{\Re} f(x)dF_n(x) \ne \int_{\Re} f(x)dF(x)$ even if $F_n \Rightarrow F$.

   (d) Prove that

   $$\Pr[X_1 \le a] - \sum_{i=1}^{N} \Pr[|X_i - X_{i+1}| > b_i] \le \Pr\left[X_{N+1} \le a + \sum_{i=1}^{N} b_i\right].$$

   (e) Prove that a sequence of uniformly bounded random variables is always uniformly integrable.

Solution.

   (a) Since for any y,

   $$\Pr[X \le y-\epsilon] - \Pr[|X-Y| > \epsilon] \le \Pr[Y \le y] \quad \text{and} \quad \Pr[|X-Y| > \epsilon] = 0,$$

   we have

   $$\Pr[X > y - \epsilon] \ge \Pr[Y > y].$$

   By $E[Y] = \int_0^{\infty} \Pr[Y > y]dy$, we obtain:

   $$\int_0^{\infty} \Pr[X > y-\epsilon]dy = \epsilon + E[X] \ge E[Y].$$

   Interchanging the roles of X and Y yields $E[Y] \ge E[X] - \epsilon$.

(b) Let $\Pr[X_n = x_n] = 1$ and $\Pr[X = x_0] = 1$, and $x_n \uparrow x_0$. Also, let $h(x) = 0$ for $x < x_0$ and $h(x) = 1$ for $x \geq x_0$. Then, obvious $\Pr[h(X_n) = 0] = 1$ and $\Pr[h(X) = 1] = 1$. So, $h(X_n) \not\Rightarrow h(X)$.

(c) Define $h(x) = e^{x^2}$. Let X_n and X respectively have cdfs $1 - e^{-(1-\epsilon_n)x^2-x}$ and $1 - e^{-x^2-x}$ with support $[0, \infty)$, where $\epsilon_n \downarrow 0$. Then,

$$\int_0^\infty e^{x^2}(2(1-\epsilon_n)x+1)e^{-(1-\epsilon_n)x^2-x}dx = \int_0^\infty (2(1-\epsilon_n)x+1)e^{\epsilon_n x^2-x}dx = \infty$$

and

$$\int_0^\infty e^{x^2}(2x+1)e^{-x^2-x}dx = \int_0^\infty (2x+1)e^{-x}dx = 3 < \infty$$

(d)

$$\begin{aligned}
&\Pr[X_1 \leq a] - \Pr[|X_1 - X_2| > b_1] \leq \Pr[X_2 \leq a + b_1] \\
&\Pr[X_2 \leq a + b_1] - \Pr[|X_2 - X_3| > b_2] \leq \Pr[X_3 \leq a + b_1 + b_2] \\
&\Pr[X_3 \leq a + b_1 + b_2] - \Pr[|X_3 - X_4| > b_3] \leq \Pr[X_4 \leq a + b_1 + b_2 + b_3] \\
&\cdots
\end{aligned}$$

This implies

$$\Pr[X_1 \leq a] - \sum_{i=1}^N \Pr[|X_i - X_{i+1}| > b_i] \leq \Pr\left[X_{N+1} \leq a + \sum_{i=1}^N b_i\right].$$

(e) The proof can be completed by taking α (in the definition of uniformly integrability) larger than the uniform bound of these random variables. □

7. (Section 26)

(a) For any two random variables X_1 and X_2 respectively with densities f_1 and f_2, prove that

$$|\varphi_1(t) - \varphi_2(t)| \leq 2\int_{\mathcal{A}} (f_1(x) - f_2(x))dx,$$

where $\varphi_i(t)$ is the characteristic function of X_i, and

$$\mathcal{A} = \{x \in \Re : f_1(x) > f_2(x)\}.$$

(b) Prove that for a given t fixed,

$$\lim_{h \downarrow 0} \sup_{X \in \mathbb{X}} |\varphi_{X+h}(t) - \varphi_X(t)| = 0,$$

where $\mathbb{X}$ is the set of all possible random variables, and $\varphi_{X+h}(t)$ and $\varphi_X(t)$ are respectively characteristic functions of $X + h$ and X.

(c) In the lecture slides, we obtain that

$$e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} = \begin{cases} \dfrac{i^{n+1}}{n!} \displaystyle\int_0^x (x-t)^n e^{it} dt \\ \dfrac{i^n}{(n-1)!} \displaystyle\int_0^x (x-t)^{n-1}(e^{it} - 1) dt \end{cases}$$

Find another expression for $e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!}$ using $(x-t)^{n-2}$.

(d) Suppose $x_k = k\theta$ for some rational number θ other than 0. Is it possible that $x_1, x_2, x_3, \ldots,$ is uniformly distributed modulo 1? Justify your answer.

Solution.

(a)

$$\begin{aligned}
|\varphi_1(t)-\varphi_2(t)| &= \left|\int_{\Re} e^{itx} f_1(x)dx - \int_{\Re} e^{itx} f_2(x)dx\right| \\
&= \left|\int_{\Re} e^{itx}(f_1(x)-f_2(x))dx\right| \\
&\le \int_{\Re} \left|e^{itx}\right| |f_1(x)-f_2(x)|\, dx \\
&= \int_{\Re} |f_1(x)-f_2(x)|\, dx \\
&= \int_{\mathcal{A}} (f_1(x)-f_2(x))dx + \int_{\mathcal{A}^c} (f_2(x)-f_1(x))dx \\
&= \int_{\mathcal{A}} (f_1(x)-f_2(x))dx + \int_{\mathcal{A}^c} f_2(x)dx - \int_{\mathcal{A}^c} f_1(x)dx \\
&= \int_{\mathcal{A}} (f_1(x)-f_2(x))dx + 1 - \int_{\mathcal{A}} f_2(x)dx - 1 + \int_{\mathcal{A}} f_1(x)dx \\
&= 2\int_{\mathcal{A}} (f_1(x)-f_2(x))dx
\end{aligned}$$

(b) By

$$\begin{aligned}
|\varphi_{X+h}(t)-\varphi_X(t)| &= \left|e^{jth}\varphi_X(t)-\varphi_X(t)\right| \\
&\le \left|e^{jth}-1\right|,
\end{aligned}$$

for which the upper bound is a quantity that is nothing to do with X and that approaches 0 as $h \to 0$, we obtain the desired result.

(c) Integration by part again yields:

$$\begin{aligned}
&\int_0^x (x-t)^{n-2}(e^{it}-1-it)dt \\
&= -\frac{(x-t)^{n-1}}{n-1}(e^{it}-1-it)\Big|_0^x + \int_0^x \frac{(x-t)^{n-1}}{n-1} i\left(e^{it}-1\right)dt \\
&= \frac{i}{n-1}\int_0^x (x-t)^{n-1}(e^{it}-1)dt.
\end{aligned}$$

Hence,

$$\frac{i^n}{(n-1)!}\int_0^x (x-t)^{n-1}(e^{it}-1)dt = \frac{i^{n-1}}{(n-2)!}\int_0^x (x-t)^{n-2}(e^{it}-1-it)dt.$$

(d) Suppose $\theta = p/q$ with integers p and q. Then, when $m\theta$ is an integer,

$$\frac{1}{n}\sum_{k=1}^{n} e^{i2\pi m(x_k - \lfloor x_k \rfloor)} = \frac{1}{n}\sum_{k=1}^{n} e^{i2\pi mk\theta} = 1.$$

So, it cannot be uniformly distributed modulo 1. □

8. (Section 27)

(a) Suppose that $Y_k = Y_{k-1} + Z_{k-1}$ for $k \geq 1$, where $\{Z_k\}_{k=0}^{\infty}$ are i.i.d. with zero mean and unit variance, and $Y_0 = 0$. Find the limit law of

$$\frac{Y_1 + Y_2 + \cdots + Y_n}{\sqrt{\text{Var}[Y_1 + Y_2 + \cdots + Y_n]}}.$$

(b) Is it possible to have an independent sequence of random variables with zero-mean and finite second moment, for which

$$\frac{X_1 + X_2 + \cdots + X_n}{\sqrt{\text{Var}[X_1 + X_2 + \cdots + X_n]}}$$

does not converge to Gaussian? Justify your answer.

(c) Suppose $\{\boldsymbol{X}_k = (X_{k,1}, X_{k,2})\}_{k\geq 1}$ are i.i.d. random vectors, in which $X_{k,1}$ is generally dependent on $X_{k,2}$. The marginal means and marginal variances for $X_{k,1}$ and $X_{k,2}$ are respectively 0 and 1. Determine the limit law of random vector

$$\frac{\boldsymbol{X}_1 + \boldsymbol{X}_2 + \cdots + \boldsymbol{X}_n}{\sqrt{n}}.$$

(Hint: Use the fact that a random vector is Gaussian if, and only if, every linear combination of the vector components is one-dimensional Gaussian.)

(d) Can a sequence of deterministic random variables placing all mass respectively at $x_1, x_2, x_3, \cdots$, can the sum (if elaborately normalized, i.e., if divided by n^α by a properly chosen α) converge to non-degenerated Gaussian?

(e) For a sequence of serially connected linear filters with common non-negative impulse response h, satisfying $\int_0^\infty h(t)dt = 1$ and $h(t) = 0$ for $t \le 0$, determine the limit output signal as $n \to \infty$ if the input signal is $i(t) = \delta(t)/\sqrt{n}$, where n is the number of filters in sequence, and $\delta(\cdot)$ is the Dirac delta function. Note that the filter input i and output o are characterized by

$$o(t) = \int_{-\infty}^{\infty} h(\tau)i(t-\tau)d\tau.$$

Solution.

(a) By $Y_n = Z_0 + Z_1 + \cdots + Z_{n-1}$, we have:

$$S_n = Y_1 + Y_2 + \cdots + Y_n = nZ_0 + (n-1)Z_1 + (n-2)Z_2 + \cdots + Z_{n-1}.$$

Define $X_{n,k} = (n-k)Z_k$ for $k = 0, 1, \ldots, n-1$. Then,

$$\begin{aligned}
\lim_{n\to\infty} \sum_{k=0}^{n-1} \frac{1}{s_n^2} E\left[X_{n,k}^2 I_{[|X_{n,k}|\ge \epsilon s_n]}\right] &= \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=0}^{n-1} E\left[(n-k)^2 Z_k^2 I_{[|(n-k)Z_k|\ge \epsilon s_n]}\right] \\
&= \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=0}^{n-1} (n-k)^2 E\left[Z_k^2 I_{[|Z_k|\ge \epsilon s_n/(n-k)]}\right] \\
&\le \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=0}^{n-1} (n-k)^2 E\left[Z_k^2 I_{[|Z_k|\ge \epsilon s_n/n]}\right] \\
&= \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=0}^{n-1} (n-k)^2 E\left[Z_1^2 I_{[|Z_1|\ge \epsilon s_n/n]}\right] \\
&= \lim_{n\to\infty} E\left[Z_1^2 I_{[|Z_1|\ge \epsilon s_n/n]}\right] = 0,
\end{aligned}$$

where the last step follows from the dominated convergence theorem, and

$$s_n^2 = \sum_{k=0}^{n-1} E\left[X_{n,k}^2\right] = n^2 + (n-1)^2 + \cdots + 1 = \frac{1}{6}n(1+n)(1+2n).$$

Therefore,

$$\frac{S_n}{\sqrt{\mathrm{Var}[S_n]}} \Rightarrow N.$$

(b) Let $X_2, X_3, \cdots, X_n$ be i.i.d. sequence of Gaussian random variables with mean zero and unit variance, and let X_1 be equal-probable random variable taking value on $\{-\sqrt{n-1}, \sqrt{n-1}\}$. Hence, by independence, $X_2 + X_3 + \cdots + X_n$ is Guassian distributed with mean zero and variance $(n-1)$. We can also derive that X_1 has mean zero and variance $(n-1)$. Therefore,

$$\frac{X_1 + X_2 + \cdots + X_n}{\sqrt{\mathrm{Var}[X_1 + X_2 + \cdots + X_n]}} = \frac{X_1}{\sqrt{2(n-1)}} + \frac{X_2 + X_3 + \cdots + X_n}{\sqrt{2(n-1)}},$$

in which the first term is binary distributed, taking values in $\{-1/\sqrt{2}, 1/\sqrt{2}\}$, and the second term is Gaussian distributed with mean and variance $1/2$. Since the distributions of these two terms are nothing to do with n, the limit law is apparently non-Gaussian.

(c) $\{aX_{k,1}+bX_{k,2}\}_{k\geq 1}$ is apparently i.i.d. with mean zero and variance $E[(aX_{k,1}+bX_{k,2})^2] = a^2+b^2+2ab\rho$, where $\rho = E[X_{k,1}X_{k,2}]$. Hence, $\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(aX_{k,1}+bX_{k,2})$ converges in distribution to Gaussian with mean zero and variance $a^2 + b^2 + 2ab\rho$. Accordingly,

$$\frac{\boldsymbol{X}_1 + \boldsymbol{X}_2 + \cdots + \boldsymbol{X}_n}{\sqrt{n}}$$

converges in distribution to Gaussian vector with mean zero and covariance matrix $\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$.

(d)

$$\left\{a_n = \frac{x_1 + x_2 + \cdots + x_n}{n^\alpha}\right\}_{n\geq 1}$$

is a sequence of real numbers. Hence, it can only converge to a constant with probability 1 or does not converge.

(e) Let $\{X_k\}_{k\geq 1}$ be i.i.d. random variables with marginal pdf h. Then, $o(t)$ will be the density function of $(X_1 + X_2 + \cdots + X_n)/\sqrt{n}$ as $n \to \infty$. Since $E[X_k]$ is positive, the above sequence of normalized sum does not converge to Gaussian but to infinity with probability 1. In other words, $o(t)$ will blow up ultimately. □

9. (Section 28)

(a) If F is the limit law of $S_n = X_{n,1} + \cdots + X_{n,r_n}$, and has characteristic function

$$\varphi(t) = \exp\left\{\int_{\Re} \left(e^{itx} - 1 - itx\right) \frac{1}{x^2}\mu(dx)\right\}.$$

Suppose the cdf of $X_{n,k}$ is $F_{n,k}$, and $X_{n,k}$ has zero-mean and positive variance. Find the canonical measure formula $\mu(-\infty, x]$ corresponding to F using $\{F_{n,k}\}$. (Hint: The proof of Theorem 28.3.)

(b) Can a deterministic random variable with all the mass at origin be an infinitely divisible distribution? Justify your answer.

(c) Can a deterministic random variable with all the mass at a point other than the origin be an infinitely divisible distribution? Justify your answer.

(d) Show that the function $\left(e^{itx} - 1 - \frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}$ does not converge to zero as $|x| \to \infty$.

(e) Show that centered gamma distribution is infinitely divisible with canonical measure $\mu(dx) = uxe^{-x}$ for $0 < x < \infty$.

Solution.

(a) From the proof of Theorem 28.3,

$$\mu(-\infty, x] = \lim_{n\to} \mu_n(-\infty, x] = \lim_{n\to\infty} \sum_{k=1}^{r_n} \int_{-\infty}^{x} y^2 dF_{n,k}(y).$$

(b) The characteristic function of this deterministic random variable is

$$E[e^{jtX}] = 1.$$

If it is infinitely divisible, then by definition of infinite divisibility, there exists characteristic function $\varphi_n(t)$ such that

$$\varphi_n^n(t) = 1,$$

which is true because $\varphi_n(t)$ can be chosen to be 1.

(c) The characteristic function of this deterministic random variable is

$$E[e^{jtX}] = e^{ita},$$

where $\Pr[X = a] = 1$. If it is infinitely divisible, then by definition of infinite divisibility, there exists characteristic function $\varphi_n(t)$ such that

$$\varphi_n^n(t) = e^{ita},$$

which is true because $\varphi_n(t)$ can be chosen to be $e^{ita/n}$.

(d)

$$\left(e^{itx} - 1 - \frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2} \quad = \quad (e^{itx} - 1) + (e^{itx} - 1 - itx)\frac{1}{x^2}.$$

The second term converges to zero as $|x| \to \infty$ as has been shown in the lecture note. But the first term $|e^{itx} - 1| = 2|\sin(tx/2)|$ does not disappear at $|x|$ large.

(e) Straightforward; hence, no detail is provided.