STOCKHOLMS UNIVERSITET	MT7001
MATEMATISKA INSTITUTIONEN	TENTAMEN
Avd. Matematisk statistik	November 25, 2021

Second Exam Probability III

November 25, 2021 kl. 8–13

Examiner: Pieter Trapman, tel. 07 22928174, ptrapman "at" math.su.se
or Daniel De La Riva Massaad, tel. 07 38220230 daniel.delariva "at" math.su.se
Permissible tools: pen, paper and attached "cheat-sheet"

5 problems. Maximum of 60 points

	A	B	C	D	E
Needed points	50	45	40	35	30

Partial answers might be worth points, even if you cannot finish an answer! You are allowed to use results from the "cheat sheet" without proof, unless the proof is explicitly asked for in the question. You may also use other results discussed in the lectures or in the course material, such as the Borel-Cantelli Lemma's. If you use such a result refer to it by stating the theorem you are using or by referring to its proper name (e.g. Fatou's lemma), and explicitly check whether the conditions of the theorem are satisfied.

Problem 1

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let X be a random variable being defined on this space. Let $\mathcal{A}$ be a sub-σ-field of $\mathcal{F}$, generated by a finite partition $\mathcal{P} = \{A_1, A_2, \cdots, A_n\}$.

(**a**) Provide the definition of an $\mathcal{F}$-measurable random variable. (Just using Proposition 3 of the cheat-sheet is not enough). (4p)

For part (b) and (c), let $\hat{X} = \mathbb{E}[X|\mathcal{A}]$.

(**b**) Show that if Y is an $\mathcal{A}$-measurable random variable, then $\mathbb{E}[(X - \hat{X})Y] = 0$. (4p)

(**c**) Show that if Z is an $\mathcal{A}$-measurable random variable, then $\mathbb{E}[(X - Z)^2] \geq \mathbb{E}[(X - \hat{X})^2]$. (4p)

Problem 2

For $\lambda > 0$, let X_λ be Poisson distributed with parameter λ. That is

$$\mathbb{P}(X_\lambda = k) = \frac{\lambda^k}{k!}e^{-\lambda} \qquad \text{for } k \in \{0, 1, 2, \cdots\}$$

and $\mathbb{P}(X_\lambda = k) = 0$ for $k \notin \{0, 1, 2, \cdots\}$.

a) Compute $\psi_\lambda(t) = \mathbb{E}[e^{tX_\lambda}]$, for $t \in \mathbb{R}$. That is, compute the moment generating function of X_λ? (2p)

Define $Y_\lambda = (X_\lambda - \lambda)/\sqrt{\lambda}$.

b) Compute $\hat{\psi}_\lambda(t) = \mathbb{E}[e^{tY_\lambda}]$, for $t \in \mathbb{R}$. That is, compute the moment generating function of Y_λ? (2p)

Let Z be a standard normal distributed random variable. That is Z has density

$$f_Z(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2} \qquad \text{for } z \in \mathbb{R}.$$

c) Compute $\psi_Z(t) = \mathbb{E}[e^{tZ}]$, for $t \in \mathbb{R}$. That is, compute the moment generating function of Z? (4p)

In part d) you may use without proof that the characteristic function of Y_λ is given by $\hat{\psi}_\lambda(it)$ and the characteristic function of Z is given by $\psi_Z(it)$, where $i = \sqrt{-1}$.

d) Show that Y_λ converges in distribution to Z as $\lambda \to \infty$. (4p)

Problem 3

Let $X_1, X_2, \cdots$ be a sequence of random variables and X another random variable, all defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

a) Let $g : [0, \infty) \mapsto [0, \infty)$ be a strictly increasing function. Show that

$$\mathbb{P}(|X| > \epsilon) \leq \frac{\mathbb{E}[g(|X|)]}{g(\epsilon)}.$$

(3p)

b) Let $g(x) = x/(1+x)$. Show that $X_n \xrightarrow{\mathbb{P}} 0$ if and only if $g(|X_n|)$ converges in expectation to 0. (6p)

c) Show that if X_n converges in expectation to X then X_n converges in probability to X. (3p)

Problem 4

Consider a supercritical Galton-Watson Branching Process $\{Z_0, Z_1, Z_2, \cdots\}$ with $Z_0 = 1$. That is, let $\{X_{ij}\}_{i=0,1,2,\cdots;j=1,2,\cdots}$ be independent and identically distributed random variables with the same distribution as the non-negative integer valued random variable X. Define

$$Z_0 = 1 \qquad \text{and} \qquad Z_{k+1} = \sum_{j=1}^{Z_k} X_{kj} \qquad \text{for } k \geq 0.$$

Assume $\mathbb{P}(X \geq 1) = 1$, $\mathbb{E}[X] = m > 1$ and $Var[X] = \sigma^2 < \infty$.

a) Show that for all $n \geq 1$

$$\mathbb{E}[Z_n] = m^n \qquad \text{and} \qquad \mathbb{E}[(Z_n)^2] = m^2\mathbb{E}[(Z_{n-1})^2] + m^{n-1}\sigma^2.$$

Deduce from this (e.g. by induction) that for $n \geq 1$,

$$\mathbb{E}[(Z_n)^2] = m^{2n} + \sigma^2 \sum_{k=1}^{n} m^{n-k} m^{2(k-1)} = m^{2n} + \sigma^2 m^{n-1} \frac{m^n - 1}{m - 1}. \qquad \text{(4p)}$$

b) Show that $W_n = m^{-n} Z_n$ converges almost surely to a random variable W as $n \to \infty$. (4p)

c) Show that as $n \to \infty$,

$$\frac{\sum_{i=1}^{n} Z_i}{\sum_{j=0}^{n-1} Z_j} - m = \frac{\sum_{i=1}^{n} (Z_i - mZ_{i-1})}{\sum_{j=0}^{n-1} Z_j} \xrightarrow{a.s.} 0.$$

(4p)

Hint: Note that by $\mathbb{P}(X \geq 1) = 1$ we have $\mathbb{P}(Z_{n+1} \geq Z_n) = 1$ for all $n \in \mathbb{N}$ and therefore $\sum_{i=1}^{n} Z_i \to \infty$ as $n \to \infty$.

Problem 5

Let N be a strictly positive integer, $X_0 = 1$ and $X_1, X_2, \cdots$ be a sequence of dependent non-negative integer valued random variables and

$$\underline{\mathcal{F}} = \{\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \cdots\}$$

be a filtration generated by thes random variables.

For $m, n \in \{0, 1, \cdots, N\}$ and for $k \in \{0, 1, 2, \cdots\}$, set

$$\mathbb{P}(X_{k+1} = n|\mathcal{F}_k) = \mathbb{P}(X_{k+1} = n|X_k)$$

and

$$\mathbb{P}(X_{k+1} = n|X_k = m) = \binom{N}{n}\left(\frac{m}{N}\right)^n \left(1 - \frac{m}{N}\right)^{N-n}$$

Define $Y_k = \left(\frac{N}{N-1}\right)^k X_k(N - X_k)$ and let $T = \min\{k \geq 1; Y_k = 0\}$.

a) Show that $Y_0, Y_1, \cdots$ is a martingale with respect to $\underline{\mathcal{F}}$. (4p)

b) Show that

$$\frac{4(N-1)}{N^2}(1 - 1/N)^n \leq \mathbb{P}(T > n) \leq (1 - 1/N)^n.$$

(8p)

Hint: Note that if $Y_n \neq 0$ then $Y_n \geq \left(\frac{N}{N-1}\right)^n (N-1)$ and $Y_n \leq \left(\frac{N}{N-1}\right)^n N^2/4$.

Good Luck!

Reminder

σ-algebras, probability measures and expectation

Definition 1 *The Borel σ-algebra on $\mathbb{R}$, is the smallest σ-algebra generated by the open subsets of $\mathbb{R}$. This definition can be extended to $\mathbb{R}^d$ for $d \geq 1$.*

Definition 2

$$\limsup_{n\to\infty} A_n := \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m$$

$$\liminf_{n\to\infty} A_n := \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m$$

Proposition 3 *A random variable X is $\mathcal{F}$-measurable if and only if $\{X \leq x\} := \{\omega \in \Omega : X(\omega) \leq x\}$ belongs to $\mathcal{F}$.*

Definition 4 *The distribution measure μ_X of the random variable X is the probability measure on $(\mathbb{R}, \mathcal{B})$ defined by $\mu_X(B) = \mathbb{P}(X \in B)$ for Borel sets $B \in \mathcal{B}$, where $\mathcal{B}$ is the Borel σ-algebra.*

Proposition 5 *If the σ-algebra $\mathcal{A}$ is generated by a finite partition $\mathcal{P}$. Then the function Y is $\mathcal{A}$ measurable if and only if Y is constant on each element of $\mathcal{P}$.*

Lemma 6 *If X, Y satisfy $\min(\mathbb{E}(X^+), \mathbb{E}(X^-)) < \infty$, then*
(i) $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$ *(linearity)*
(ii) $\mathbb{E}(X) \leq \mathbb{E}(Y)$ *if* $X \leq Y$ *a.s.* *(monotonicity)*

Definition 7 *Let $\mathcal{P} = \{A_1, \cdots, A_n\}$ be a finite partition, which generates the σ-algebra $\mathcal{A} \subset \mathcal{F}$, then $\mathbb{E}(X|\mathcal{A})(\omega) = \sum_{i=1}^{n} \mathbb{E}(X|A_i)\mathbb{1}(\omega \in A_i)$ for $\omega \in \Omega$*

Lemma 8 (Jensen's inequality) *We have $\mathbb{E}(\phi(X)) \geq \phi(\mathbb{E}(X))$ for convex functions ϕ.*

Characteristic functions

Definition 9 *the Characteristic function of a random variable X is the function $\varphi : \mathbb{R} \to \mathbb{C}$, defined by $\varphi_X(t) = \mathbb{E}(e^{itX}) = \mathbb{E}(\cos[tX]) + i\mathbb{E}(\sin[tx])$ where $i = \sqrt{-1}$.*

Properties of φ_X:

- $\varphi_X(0) = 1$
- $|\varphi_X(t)| \leq 1$
- $\varphi_X(-t) = \overline{\varphi_X(t)}$
- If $a, b \in \mathbb{R}$ and $Y = aX + b$ then $\varphi_Y(t) = e^{itb}\varphi_X(at)$
- If the random variables X and Y are independent, then $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$
- φ_X is real if and only if X and $-X$ have the same distribution, (X is symmetric)

Theorem 10 *Let X be a random variable with distribution function F and characteristic function φ. If F is continuous in both a and b, then*

$$F(b) - F(a) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itb} - e^{-ita}}{-it} \varphi(t) dt$$

special cases:

- If $\int_{\mathbb{R}} |\varphi(t)| dt < \infty$, then X has a continuous distribution with density
$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt$$
- If the distribution of X is discrete, then
$$\mathbb{P}(X = x) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{-itx} \varphi(t) dt$$

Theorem 11 *Let $\varphi^{(k)}(\cdot)$ be the k-th complex derivative of φ.*

- *If $\varphi_X^{(k)}(0)$ exists then $\mathbb{E}(|X^k|) < \infty$ if k is even and $\mathbb{E}(|X^{k-1}|) < \infty$ if k is odd*
- *if $\mathbb{E}(|X^k|) < \infty$ then $\varphi_X(t) = \sum_{j=0}^{k} \frac{\mathbb{E}(X^j)}{j!}(it)^j + o(t^k)$, where $f(x) = o(x)$ if $f(x)/x \to 0$ for $x \to 0$*

Some useful results for convergence results

Chebychev's inequality: $\mathbb{P}(|X| > x) \leq \frac{\mathbb{E}(X^2)}{x^2}$

Markov inequality: $\mathbb{P}(|X| > x) \leq \frac{\mathbb{E}(|X|^r)}{x^r}$

Hölder's inequality: For $p, q > 1$ such that $1/p + 1/q = 1$ we have

$$\mathbb{E}(|XY|) \leq [\mathbb{E}(|X|^p)]^{1/p}[\mathbb{E}(|X|^q)]^{1/q}$$

Minkovski's inequality: For $r \geq 1$ we have

$$[\mathbb{E}(|X+Y|^r)]^{1/r} \leq [\mathbb{E}(|X|^r)]^{1/r} + [\mathbb{E}(|X|^r)]^{1/r}$$

Lemma 12 (Fatou's Lemma) *Let $X_1, X_2, \cdots$ be non-negative random variables, then $\mathbb{E}(\liminf X_n) \leq \liminf \mathbb{E}(X_n)$.*

Definition 13 (Tail events) *If $X_1, X_2, \cdots$ are random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{H}_n = \sigma(X_{n+1}, X_{n+2}, \cdots)$ is the smallest σ-algebra in which all random variables $X_{n+1}, X_{n+2}, \cdots$ are measurable, then $\mathcal{H}_\infty := \cap_n \mathcal{H}_n$ is called the tail σ-algebra, and events contained in it are tail events.*

Theorem 14 (Kolmogorov's zero-one law) *If $X_1, X_2, \cdots$ are independent, then all tail events $H \subset \mathcal{H}_\infty$ satisfy either $\mathbb{P}(H) = 1$ or $\mathbb{P}(H) = 0$*

Definition 15 (Uniform integrability) *A sequence of r.v. $X_1, X_2, \cdots$ is uniformly integrable if*

$$\sup_{n \geq 1} \mathbb{E}(|X_n| \mathbb{1}(|X_n| > a)) \to 0 \qquad \text{as } a \to \infty$$

Theorem 16 *Let X and $X_1, X_2, \cdots$ be random variables such that $X_n \xrightarrow{\mathbb{P}} X$ then the following statements are equivalent*

1. *$X_1, X_2, \cdots$ is uniformly integrable*
2. *$\mathbb{E}(|X_n|) < \infty$ for all n, $\mathbb{E}(|X|) < \infty$ and $X_n \xrightarrow{1} X$*
3. *$\mathbb{E}(|X_n|) < \infty$ and $\mathbb{E}(|X_n|) \to \mathbb{E}(|X|) < \infty$*

Martingales

Some Properties of martingales: Let $S_1, S_2, \cdots$ be a martingale with respect to $\underline{\mathcal{F}} = (\mathcal{F}_0, \mathcal{F}_1, \cdots)$.

- $\mathbb{E}(S_{n+m}|\mathcal{F}_n) = S_n$
- $\mathbb{E}(S_n) = \mathbb{E}(S_1)$
- $\mathbb{E}((S_n)^2)$ is non decreasing

Theorem 17 (Doob decomposition) *A $\underline{\mathcal{F}}$-submartingale $Y_0, Y_1, \cdots$ with finite means may be expressed in the form $Y_n = M_n + S_n$, where $M_1, M_2, \cdots$ is a $\underline{\mathcal{F}}$-martingale and S_n is $\mathcal{F}_{n-1}$ measurable for all n. This decomposition is unique.*

Lemma 18 (Doob-Kolmogorov inequality) *If $S_1, S_2, \cdots$ is a martingale with respect to $\underline{\mathcal{F}}$, then for all $\epsilon > 0$ we have $\mathbb{P}\left(\max_{1\leq k\leq n} |S_k| \geq \epsilon\right) \leq \epsilon^{-2}\mathbb{E}((S_n)^2)$.*

Theorem 19 (Martingale convergence theorem) *If $S_1, S_2, \cdots$ is a martingale with respect to $\underline{\mathcal{F}}$ and $\mathbb{E}((S_n)^2) \nearrow M < \infty$, then there exists a random variable S such that $S_n \overset{a.s.}{\to} S$.*

Definition 20 (Cauchy sequence) *A sequence of real numbers $x_1, x_2, \cdots$ is a Cauchy sequence if for all $\epsilon > 0$ there exists an N such that for all $n \geq m \geq N$, we have $|x_n - x_m| < \epsilon$.*
We know that a sequence is convergent if and only if it is a Cauchy sequence.

Theorem 21 *Let $S_0, S_1, \cdots$ be a martingale with respect to $\underline{\mathcal{F}}$ such that $S_0 = 0$ and $\mathbb{E}((S_n)^2) < \infty$ for all n. Define*

$$\langle S\rangle_n = \sum_{k=1}^{n} \mathbb{E}((S_k - S_{k-1})^2|\mathcal{F}_{k-1}) \qquad \textit{and} \qquad \langle S\rangle_\infty = \lim_{n\to\infty} \langle S\rangle_n.$$

Let $f \geq 1$ be a given increasing function satisfying $\int_0^\infty [f(x)]^{-2}dx < \infty$. Then,
(i) On $\{\omega : \langle S(\omega)\rangle_\infty < \infty\}$ $S_n \overset{a.s.}{\to} S$ for some random variable S
(ii) On $\{\omega : \langle S(\omega)\rangle_\infty = \infty\}$, $S_n/f(\langle S\rangle_n) \overset{a.s.}{\to} 0$

Theorem 22 (Strong Law of Large Numbers) *Let $X_1, X_2, \cdots$ be i.i.d. with $\mathbb{E}(X_1) = \mu$ and $Var(X_1) = \sigma^2 < \infty$ and define $S_0 = 0$ and $S_n = \sum_{k=1}^{n}(X_k - \mu)$ for $n \geq 1$. Then $\frac{S_n}{n} \overset{a.s.}{\to} 0$.*

Theorem 23 (Martingale Central Limit theorem) *$S_0, S_1, \cdots$ is a martingale with respect to $\underline{\mathcal{F}}$, with $S_0 = 0$ and $\mathbb{E}((S_n)^2) < \infty$ for all n. Assume that $n^{-1}\langle S\rangle_n \xrightarrow{\mathbb{P}} \sigma^2 > 0$ and for all $\epsilon > 0$*

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}((S_k - S_{k-1})^2 \mathbb{1}((S_k - S_{k-1})^2 > \epsilon n)) \to 0.$$

Then, $\frac{1}{\sqrt{n\sigma^2}} S_n \xrightarrow{d} \mathcal{N}(0,1)$

Theorem 24 (Optional stopping I) *Let $S_1, S_2, \cdots$ be a martingale with respect to $\underline{\mathcal{F}}$. If T is an a.s. bounded stopping time for $\underline{\mathcal{F}}$ (i.e. $\mathbb{P}(T \leq a) = 1$ for some $a \geq 0$), then $\mathbb{E}(S_T) = \mathbb{E}(S_1)$.*

Theorem 25 (Optional stopping II) *Let $S_1, S_2, \cdots$ be a martingale with respect to $\underline{\mathcal{F}}$ and T a stopping time for $\underline{\mathcal{F}}$. Then $\mathbb{E}(S_T) = \mathbb{E}(S_1)$, if the following conditions hold*

- $\mathbb{P}(T < \infty) = 1$,
- $\mathbb{E}(|S_T|) < \infty$,
- $\mathbb{E}(S_n \mathbb{1}(T > n)) \to 0$ *as* $n \to \infty$.

Theorem 26 (Optional Stopping III) *Let $S_1, S_2, \cdots$ be a martingale with respect to $\underline{\mathcal{F}}$ and T a stopping time for $\underline{\mathcal{F}}$. Then $\mathbb{E}(S_T) = \mathbb{E}(S_1)$, if the following conditions hold*

- $\mathbb{E}(T) < \infty$,
- $\mathbb{E}(|S_{n+1} - S_n||\mathcal{F}_n) \leq K$ *for all* $n < T$ *and some* $K > 0$

Wald's equation and identity: If $X_1, X_2, \cdots$ are i.i.d. random variables with $\mathbb{E}(X_1) = \mu < \infty$ and $S_n = \sum_{k=1}^{n} X_k$ and T is a stopping time satisfying $\mathbb{E}(T) < \infty$, then $\mathbb{E}(S_T) = \mu\mathbb{E}(T)$.
If in addition there exists a $h > 0$ such that $M(t) = \mathbb{E}(e^{tX_1}) < \infty$ for all $|t| < h$ and $M(t) > 1$ and $|S_n| < C$ for some constant $C > 0$ and all $n \leq T$, then $\mathbb{E}(e^{tS_T}[M(t)]^{-T}) = 1$.