UNIVERSITY OF OSLO

Faculty of mathematics and natural sciences

Deferred exam in:	STK-MAT3710/4710 — Probability Theory
Day of examination:	18th of January 2024
Examination hours:	15:00 – 19:00
This problem set consists of 8 pages.	
Appendices:	None
Permitted aids:	None

Please make sure that your copy of the problem set is complete before you attempt to answer anything.

This exam consists of four problems. The first one is of theoretical nature, while the three remaining problems are more applied. Make sure to be **precise** and **rigorous** when stating mathematical results and formulae. Please, **write clearly**, **orderly** and avoid scribbling.

Grading: The total score is 10 points. The grading scale is F [0,4), E [4,5), D [5, 6), C [6, 7), B [7, 8.5), A [8.5,10].

Problem 1 Theory (2 points)

Let X be an integrable random variable defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $\mathcal{F}, \mathcal{G}$ sub-σ-algebras such that $\mathcal{G} \subseteq \mathcal{F} \subseteq \mathcal{A}$.

(a) Prove that if $\mathcal{F}$ is independent of $\sigma(X)$, then $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$.

Solution: Let $Z = \mathbb{E}[X|\mathcal{F}]$, then Z is the $\mathbb{P}$-a.s. unique $\mathcal{F}$-measurable and integrable random variable such that

$$\int_F Z d\mathbb{P} = \int_F X d\mathbb{P}, \quad F \in \mathcal{F}.$$

Observe that $Z = \mathbb{E}[X]$ satisfies the requirement. It is a constant, so it is integrable and obviously $\mathcal{F}$-measurable. Finally,

$$\int_F \mathbb{E}[X] d\mathbb{P} = \mathbb{E}[X]\mathbb{P}[F] = \mathbb{E}[\mathbb{I}_F X] = \int_F X d\mathbb{P}$$

for all $F \in \mathcal{F}$. Hence, $Z = \mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$.

Points: Give 0.5 if the rigorous definition of conditional expectation is given. Give 0.75 if the exercise is solved for random variables. Subtract 0.2 per mistake and 0.5 per serious mistake.

(Continued on page 2.)

(b) Prove that $\mathbb{E}[\mathbb{E}[X|\mathcal{F}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$.

Solution: Let $Z = \mathbb{E}[X|\mathcal{F}]$, then $\mathbb{E}[\mathbb{E}[X|\mathcal{F}]|\mathcal{G}] = \mathbb{E}[Z|\mathcal{G}]$ is the object such that

$$\int_G \mathbb{E}[Z|\mathcal{G}]d\mathbb{P} = \int_G Z d\mathbb{P}$$

for all $G \in \mathcal{G}$. But, now $Z = \mathbb{E}[X|\mathcal{F}]$ is the object that satisfies

$$\int_F Z d\mathbb{P} = \int_F X d\mathbb{P}$$

for al $F \in \mathcal{F}$, in particular for $F \in \mathcal{G}$ since $\mathcal{G} \subseteq \mathcal{F}$. Hence,

$$\int_G Z d\mathbb{P} = \int_G X d\mathbb{P}$$

for all $G \in \mathcal{G}$. In a summary, we have proven

$$\int_G \mathbb{E}[Z|\mathcal{G}]d\mathbb{P} = \int_G X d\mathbb{P}$$

for all $G \in \mathcal{G}$, so

$$\mathbb{E}[Z|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$$

as we wanted to show.

Points: Give 0.5 if the first step is conducted. Give 0.5 for the rest. Subtract 0.2 per mistake and 0.5 per serious mistake.

Problem 2 Random variables (3 points)

Consider the sequence of random variables $\{X_n\}_{n=1}^{\infty}$ and Y on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ given by

$$X_n = \frac{(-1)^n Y}{n}, \quad n \geq 1,$$

where Y is exponentially distributed with parameter $\lambda > 0$, i.e. Y is absolutely continuous with density function $f_Y(y) = \lambda e^{-\lambda y}$, $y > 0$.

(a) Show that X_n, $n \geq 1$ are absolutely continuous with density functions given by

$$f_{X_n}(x) = \lambda n e^{-\lambda n x}, \quad x > 0,$$

if n is even and

$$f_{X_n}(x) = \lambda n e^{\lambda n x}, \quad x < 0,$$

if n is odd.

(Continued on page 3.)

Solution: Let us assume that n is even, then $X_n = \frac{Y}{n}$. Since $Y \in (0,\infty)$ so does $X_n = \frac{Y}{n}$ as well. Let $x \in (0,\infty)$, the distribution function of X_n is given by

$$\mathbb{P}[X_n \leq x] = \mathbb{P}\left[\frac{Y}{n} \leq x\right] = \mathbb{P}[Y \leq nx] = 1 - e^{-\lambda nx}.$$

We see that the distribution function is differentiable, hence the density is given by $f_{X_n}(x) = \lambda n e^{-\lambda nx}$, $x > 0$.
Now if n is odd then $X_n = \frac{-Y}{n}$. Since $Y \in (0,\infty)$ then $X_n = \frac{-Y}{n} \in (-\infty, 0)$. Let $x \in (-\infty, 0)$, the distribution function of X_n now is given by

$$\mathbb{P}[X_n \leq x] = \mathbb{P}\left[\frac{-Y}{n} \leq x\right] = \mathbb{P}[Y \geq -nx] = e^{\lambda nx}.$$

We see that the distribution function is differentiable, hence the density is given by $f_{X_n}(x) = \lambda n e^{\lambda nx}$, $x < 0$.

Points: Give 0.5 per each and 0.2 for an attempt to compute the distribution. Subtract 0.2 per mistake and 0.5 per serious mistake.

(b) Discuss the convergence of the sequence $\{X_n\}_{n=1}^{\infty}$ almost surely, in L^p for any $p \geq 1$, in probability and in distribution.

Solution: Let us try to show a.s. convergence, then we will automatically get convergence in probability and thereby convergence in distribution. From the distribution functions above we can directly see convergence to 0, in distribution. Hence, if there is a limit, it will be 0. Let us define the events $A_n = \{|X_n| > \varepsilon\}$ for an arbitrarily small $\varepsilon > 0$. Then

$$\mathbb{P}[A_n] = \mathbb{P}[|X_n| > \varepsilon] = \mathbb{P}[Y > n\varepsilon] = e^{-\lambda n\varepsilon}.$$

Now,

$$\sum_{n=1}^{\infty} \mathbb{P}[A_n] = \sum_{n=1}^{\infty} e^{-\lambda n\varepsilon} = \sum_{n=1}^{\infty} \left(e^{-\lambda\varepsilon}\right)^n < \infty$$

since $e^{-\lambda\varepsilon} < 1$ and we have a geometric series. Hence, by Borel-Cantelli's first lemma

$$\mathbb{P}[\limsup_n A_n] = 0$$

meaning that the events $\{|X_n| > \varepsilon\}$ do not happen infinitely often for any arbitrarily small $\varepsilon > 0$ and hence X_n converges almost surely to 0.
Another way to prove a.s. convergence is by direct definition:

$$\mathbb{P}\left[\lim_n X_n(\omega) = 0\right] = \mathbb{P}\left[Y(\omega)\lim_n \frac{(-1)^n}{n} = 0\right] = \mathbb{P}\left[\lim_n \frac{(-1)^n}{n} = 0\right] = 1.$$

(Continued on page 4.)

Almost sure convergence implies convergence in probability (which we also see in the fact that $\lim_n \mathbb{P}[A_n] = 0$) and convergence in distribution.
Finally, since

$$|X_n|^p \leq \frac{Y^p}{n^p} \leq Y^p$$

and $\mathbb{E}[Y^p] < \infty$. By Lebesgue's dominated convergence theorem

$$\lim_n \mathbb{E}[|X_n|^p] = \mathbb{E}[|\lim_n X_n|^p] = 0,$$

where we used the continuity of the function $|\cdot|^p$ and that $X_n \to 0$ almost surely.

Points: Give 0.2 per each type of convergence and 0.1 per just defining each type. Give 0.2 if the right Borel-Cantelli lemma is stated correctly. Subtract 0.2 per mistake and 0.5 per serious mistake.

(c) Let $\{a_n\}_{n=1}^{\infty} \subset \mathbb{R}$ be a sequence of non-negative real numbers such that $\sum_{n=1}^{\infty} \frac{n(n+1)}{2n+1} a_n < \infty$. We say that X_n and X_{n+1} are *friends* if their distance is at most a_n. Prove that the infinite sequence $X_1(\omega), X_2(\omega), \ldots$ has a finite number of friends.

Solution: We need to look at the occurrence of the events $B_n = \{|X_{n+1} - X_n| \leq a_n\}$. Since B_n, $n \geq 1$ are not independent we can again hope for Borel-Cantelli's first lemma. Observe that

$$|X_{n+1} - X_n| = \frac{2n+1}{n(n+1)} Y.$$

$$\mathbb{P}[A_n] = \mathbb{P}\left[\frac{2n+1}{n(n+1)} Y \leq a_n\right] = 1 - e^{-\lambda \frac{n(n+1)a_n}{2n+1}}.$$

Now, using that $1 - e^{-x} \leq x$ we get

$$\sum_{n=1}^{\infty} \mathbb{P}[A_n] \leq \lambda \sum_{n=1}^{\infty} \frac{n(n+1)a_n}{2n+1} < \infty,$$

by assumption of a_n. Then, by the first Borel-Cantelli lemma the events A_n do not happen infinitely often. Hence, we conclude that we will only observe a finite number of friends.

Points: Give 0.2 if the first Borel-Cantelli lemma is stated correctly. Subtract 0.2 per mistake and 0.5 per serious mistake.

Problem 3 Conditional expectation (3 points)

Let X be uniformly distributed on $[0,1]$, i.e. X is absolutely continuous with density $f_X(x) = 1$, $x \in [0,1]$. Moreover, let N be geometrically

(Continued on page 5.)

distributed on the set of natural numbers $\{1, 2, \dots\}$ with parameter $p \in (0,1)$, i.e. N has probability mass function $\mathbb{P}[N = k] = (1-p)^{k-1}p$, $k = 1, 2, \dots$. Assume that N and X are independent and define the random variable $Y = X^{1/N}$.

(a) Show that Y is absolutely continuous with density function given by

$$f_Y(y) = \frac{p}{(1-(1-p)y)^2}, \quad y \in [0,1]$$

and that

$$\mathbb{E}[Y] = \frac{1}{1-p} + \frac{p}{(1-p)^2}\log p.$$

Solution: Since X is uniform then $X \in (0,1)$ and since $N \geq 1$, then $X^{1/N} \in (0,1)$ as well. Let $y \in (0,1)$, conditioning on N, by the law of total probability, we have

$$\begin{aligned} F_Y(y) &= \sum_{k=1}^{\infty} \mathbb{P}[X^{1/N} \leq y | N = k]\mathbb{P}[N = k] \\ &= \sum_{k=1}^{\infty} \mathbb{P}[X^{1/k} \leq y](1-p)^{k-1}p \\ &= p\sum_{k=1}^{\infty} \mathbb{P}[X \leq y^k](1-p)^{k-1} \\ &= \frac{p}{1-p}\sum_{k=1}^{\infty}((1-p)y)^k \\ &= p\frac{y}{1-(1-p)y}. \end{aligned}$$

We can see that the distribution function is differentiable in $y \in (0,1)$ hence the density function is given by

$$f_Y(y) = \frac{d}{dy}\mathbb{P}[Y \leq y] = \frac{p}{(1-(1-p)y)^2}, \quad y \in (0,1).$$

Now, the expectation can be computed by using the density function

$$\mathbb{E}[Y] = \int_0^1 \frac{py}{(1-(1-p)y)^2}dy.$$

We use integration-by-parts, by choosing $u = py$ and $dv = (1-(1-$

(Continued on page 6.)

$p)y)^{-2}dy$ then $du = pdy$ and $v = \frac{1}{1-p}\frac{1}{1-(1-p)y}$. Therefore,

$$\begin{aligned}\mathbb{E}[Y] &= \int_0^1 \frac{py}{(1-(1-p)y)^2}dy \\ &= \frac{py}{1-p}\frac{1}{1-(1-p)y}\Big|_0^1 - \frac{p}{1-p}\int_0^1 \frac{dy}{1-(1-p)y} \\ &= \frac{1}{1-p} + \frac{p}{(1-p)^2}\log(1-(1-p)y)\Big|_0^1 \\ &= \frac{1}{1-p} + \frac{p}{(1-p)^2}\log p.\end{aligned}$$

Points: Give 0.5 for density and 0.5 for expectation. Give 0.2 for noticing that one has to condition on N. Subtract 0.2 per mistake and 0.5 per serious mistake.

(b) Compute $\mathbb{E}[Y|\sigma(N)]$ and use it to show that

$$\sum_{k=1}^{\infty} \frac{k}{k+1}\frac{1}{2^k} = 2\left(1-\log 2\right).$$

Solution: Since N is discrete, let us just fix an event $N = k$ and then we leave it as a function of N. We have

$$\mathbb{E}[Y|N=k] = \mathbb{E}[X^{1/N}|N=k] = \mathbb{E}[X^{1/k}] = \int_0^1 x^{1/k}dx = \frac{k}{k+1}.$$

Hence,

$$\mathbb{E}[Y|\sigma(N)] = \frac{N}{N+1}.$$

Note that it is not easy to compute $\mathbb{E}\left[\frac{N}{N+1}\right]$ since, by the law of the unconscious statistician, it is given by

$$\mathbb{E}\left[\frac{N}{N+1}\right] = \sum_{k=1}^{\infty}\frac{k}{k+1}(1-p)^{k-1}p.$$

But, we know that

$$\mathbb{E}\left[\frac{N+1}{N}\right] = \mathbb{E}\left[\mathbb{E}[Y|\sigma(N)]\right] = \mathbb{E}[Y],$$

by the tower property of the conditional expectation and we have computed $\mathbb{E}[Y]$. Hence,

$$\mathbb{E}\left[\frac{N}{N+1}\right] = \sum_{k=1}^{\infty}\frac{k}{k+1}(1-p)^{k-1}p = \frac{1}{1-p} + \frac{p}{(1-p)^2}\log p.$$

So,

$$\sum_{k=1}^{\infty}\frac{k}{k+1}(1-p)^{k-1} = \frac{1}{p(1-p)} + \frac{1}{(1-p)^2}\log p.$$

(Continued on page 7.)

Finally,

$$\sum_{k=1}^{\infty} \frac{k}{k+1}(1-p)^k = \frac{1}{p} + \frac{1}{1-p}\log p.$$

Plugging $p = \frac{1}{2}$ leads to the desired result.

Points: Give 0.5 for computing the conditional expectation and 0.25 for mentioning or using the tower property. Give 0.2 for connecting that to the given formula. Subtract 0.2 per mistake and 0.5 per serious mistake.

(c) Find the distribution of (N, Y).

Solution: We need to find $\mathbb{P}[N = k, Y \leq y]$ for $k = 1, 2, \ldots$ and $y \in (0, 1)$. It is obvious that N and Y are not independent.

$$\begin{aligned}
\mathbb{P}[N = k, Y \leq y] &= \mathbb{P}[N = k, X^{1/N} \leq y] \\
&= \mathbb{P}[N = k, X \leq y^N] \\
&= \mathbb{P}[X \leq y^N | N = k]\mathbb{P}[N = k] \\
&= \mathbb{P}[X \leq y^k]\mathbb{P}[N = k] \\
&= y^k(1-p)^{k-1}p.
\end{aligned}$$

Points: Give 0.5 for computing the conditional distribution. Subtract 0.2 per mistake and 0.5 per serious mistake.

Problem 4 Martingales (2 points)

Let $\{X_n\}_{n=0}^{\infty}$ a time-homogeneous Markov chain on a finite state space S, i.e. the Markov property holds a.s. $\mathbb{E}[f(X_{n+1})|\mathcal{F}_n] = \mathbb{E}[f(X_{n+1})|\sigma(X_n)]$ for all $n \geq 0$ and the transition probabilities $p_{ij} \triangleq \mathbb{P}[X_{n+1} = j | X_n = i]$, $i, j \in S$ do not depend on n. Let $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$, $n \geq 0$.

For any function $f : S \to \mathbb{R}$ such that $\mathbb{E}[|f(X_n)|] < \infty$, define the function $Pf : S \to \mathbb{R}$

$$Pf(i) = \mathbb{E}[f(X_{n+1})|X_n = i], \quad i \in S.$$

(a) Prove that $M_n \triangleq f(X_n) - \sum_{k=0}^{n-1}(Pf(X_k) - f(X_k))$, $n \geq 1$ is a martingale with respect to the given filtration.

Solution: We need to check for the three conditions: adaptedness, integrability and the martingale property.
Adaptedness: M_n is a function of $X_0, \ldots, X_n$ hence M_n is $\mathcal{F}_n$-measurable.
Integrability: $\mathbb{E}[|M_n|] \leq C\mathbb{E}[|f(X_0)|] < \infty$ where we used the tower property and the fact that $\mathbb{E}[|f(X_n)|] = \mathbb{E}[|f(X_0)|]$ since the chain is

(Continued on page 8.)

homogeneous.
Martingale property: First observe that

$$\begin{aligned} M_{n+1} &= f(X_{n+1}) - Pf(X_n) + f(X_n) - \sum_{k=0}^{n-1} (Pf(X_k) - f(X_k)) \\ &= f(X_{n+1}) - Pf(X_n) + M_n. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{E}[M_{n+1}|\mathcal{F}_n] &= \mathbb{E}\left[f(X_{n+1}) - Pf(X_n) + M_n|\mathcal{F}_n\right] \\ &= \mathbb{E}\left[f(X_{n+1}) - Pf(X_n)|\mathcal{F}_n\right] + M_n \\ &= \mathbb{E}\left[f(X_{n+1})|\mathcal{F}_n\right] - Pf(X_n) + M_n \end{aligned}$$

Now, we need to show that $\mathbb{E}[f(X_{n+1})|\mathcal{F}_n] - Pf(X_n) = 0$. Indeed, by the Markov property we have

$$\mathbb{E}[f(X_{n+1})|\mathcal{F}_n] = \mathbb{E}[f(X_{n+1})|X_n] = Pf(X_n)$$

and the result follows.

Points: Give 0.2 for adaptedness, 0.2 for integrability and 0.6 for the martingale property. Give 0.2 for defining only the concept of martingale. Subtract 0.2 per mistake and 0.5 per serious mistake.

(b) Show that for every $\alpha > 0$, $\mathbb{P}[M_n^* \geq \alpha] \leq \frac{(2n+1)c}{\alpha}$ where $M_n^* = \sup_{k=1,\dots,n} |M_k|$ and $c = \mathbb{E}[|f(X_0)|]$.

Solution: We simply use Doob's maximal inequality:

$$\mathbb{P}[M_n^* \geq \alpha] \leq \frac{\mathbb{E}[|M_n|]}{\alpha}.$$

Now, it remains to estimate $\mathbb{E}[|M_n|]$. Using the triangle inequality, Jensen's inequality (for conditional expectations) and the tower property we have

$$\mathbb{E}[|M_n|] \leq \mathbb{E}[|f(X_n)|] + \sum_{k=0}^{n-1} \mathbb{E}[|f(X_k)|] + \sum_{k=0}^{n-1} \mathbb{E}[|f(X_k)|]$$

and since $\mathbb{E}[|f(X_k)|] \leq \sup_k \mathbb{E}[|f(X_k)|] = c$ for all k we have

$$\mathbb{E}[|M_n|] \leq c + nc + nc = (2n+1)c$$

and the result follows.

Points: Give 0.5 for stating Doob's maximal inequality. Give 0.5 for estimating the expectation. Give 0.2 for mentioning Jensen and 0.2 for tower property. Subtract 0.2 per mistake and 0.5 per serious mistake.

GOOD LUCK!