Week 4: Kolmogorov 0 – 1 law, integration

Submission of solutions. Feedback can be given on Exercise 1 and any other exercise from the Training exercises. If you want to hand in, do it so by Monday 16/10/2023 17:00 (online) following the instructions on the course website

https://metaphor.ethz.ch/x/2023/hs/401-3601-00L/

Please pay attention to the quality, the precision and the presentation of your mathematical writing.

* * *

All random variables are defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

1 Exercise covered during the exercise class

The following exercise will be covered during the exercise class.

Exercise 1. Let μ and ν be measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

1) Assume that $\int_{\mathbb{R}} f(x)\mu(\mathrm{d}x) = \int_{\mathbb{R}} f(x)\nu(\mathrm{d}x)$ when f is any nonnegative measurable function. Show that $\mu = \nu$.
(2) Assume that $\mu(\mathbb{R}) = 1$. Assume that $\int_{\mathbb{R}} f(x)\mu(\mathrm{d}x) = \int_{\mathbb{R}} f(x)\nu(\mathrm{d}x)$ when f is any function of the form $f(x) = \mathbb{1}_{x\in(a,b)}$ with $a < b$ where $a, b \in \mathbb{R}$. Show that $\mu = \nu$.
(3) Assume that $\mu(\mathbb{R}) = 1$. Assume that $\int_{\mathbb{R}} f(x)\mu(\mathrm{d}x) = \int_{\mathbb{R}} f(x)\nu(\mathrm{d}x)$ when f is any continuous function with compact support. Show that $\mu = \nu$.

Solution:

(1) By taking $f = \mathbb{1}_A$ for $A \in \mathcal{B}(\mathbb{R})$, we get $\mu(A) = \nu(A)$ for every $A \in \mathcal{B}(\mathbb{R})$.
(2) Let us first check that $\mu(A) = \nu(\mathbb{R}) < \infty$. To this end, write

$$\mu((-n,n)) = \int_{\mathbb{R}} \mathbb{1}_{(-n,n)}\mu(\mathrm{d}x) = \int_{\mathbb{R}} \mathbb{1}_{(-n,n)}\nu(\mathrm{d}x) = \nu((-n,n)).$$

By taking $n \to \infty$ and continuity from below for measures ($\mathbb{R} = \bigcup_{n\geq 1}(-n,n)$), we get $\mu(A) = \nu(\mathbb{R}) < \infty$.

We conclude that $\mu = \nu$ by an application of the results from Lecture 3 (specifically, $\{(a,b) \in \mathcal{B}(\mathbb{R}) : a < b, \text{for} a, b \in \mathbb{R}\}$, the collection of all open intervals, is a π generating system of $\mathcal{B}(\mathbb{R})$.)
(3) Fix $a < b$. The idea is to approximate $\mathbb{1}_{(a,b)}$ by a continuous function. To this end, set $f_n(x) = 0$ if $x \notin (a,b)$, and for n sufficiently large set

$$f_n(x) = \min(1, n\min(b-x, x-a)).$$

Then the sequence (f_n) is increasing, and by the monotone convergence theorem

$$\int_{\mathbb{R}} f_n(x)\mu(\mathrm{d}x) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} \mathbb{1}_{(a,b)}\mu(\mathrm{d}x), \qquad \int_{\mathbb{R}} f_n(x)\nu(\mathrm{d}x) \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} \mathbb{1}_{(a,b)}\nu(\mathrm{d}x).$$

We conclude by question 2).

□

2 Training exercises

Exercise 2.

(1) Let $X \geq 0$ be a non-negative real-valued random variable. Show that $\mathbb{E}[\min(X,n)] \to \mathbb{E}[X]$ as $n \to \infty$.

(2) Let $X \geq 0$ be a non-negative real-valued random variable.

(a) Assume that $\mathbb{E}[X] < \infty$. Show that $n\mathbb{E}\left[\ln\left(1+\frac{X}{n}\right)\right] \to \mathbb{E}[X]$ as $n \to \infty$.

(b) Assume that $\mathbb{E}[X] = \infty$. Show that $n\mathbb{E}\left[\ln\left(1+\frac{X}{n}\right)\right] \to \infty$ as $n \to \infty$.

(3) Let X be a real-valued integrable random variable. Show that $\mathbb{E}\left[X\mathbb{1}_{|X|\geq n}\right] \to 0$ as $n \to \infty$.

Solution:

(1) This is a consequence of the monotone convergence theorem, since the sequence of nonnegative random variables $\min(X,n)$ is nondecreasing and converges pointwise to x.

(2) (a) Set $X_n = n\ln(1+X/n)$. Then using the inequality $\ln(1+x) \leq x$ valid for every $x \geq 0$, we have $|X_n| \leq X$ (integrable random variable which does not depend on n) and $X_n \to X$ pointwise. We conclude by dominated convergence.

(b) By Fatou's lemma,

$$\liminf_{n\to\infty} n\mathbb{E}\left[\ln\left(1+\frac{X}{n}\right)\right] \geq \mathbb{E}\left[\liminf_{n\to\infty} n\ln\left(1+\frac{X}{n}\right)\right] = \mathbb{E}[X] = \infty.$$

(3) Set $X_n = X\mathbb{1}_{|X|\geq n}$. Then $X_n \to 0$ pointwise, and $|X_n| \leq |X|$, which is integrable and independent of n. We conclude by dominated convergence.

□

* * *

Exercise 3. Let $X \geq 0$ be a non-negative real-valued random variable.

(1) Let $g : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing continuously differentiable function such that $g(0) = 0$. Show that

$$\mathbb{E}[g(X)] = \int_0^{+\infty} g'(t)\mathbb{P}(X \geq t)\,\mathrm{d}t.$$

Hint. Write $g(X)$ as an integral and use Fubini-Tonnelli's theorem.

(2) (Application) Let X be a nonnegative random variable. Show that $\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X \geq t)\mathrm{d}t$.

Solution:

(1) The function $g : \mathbb{R}_+ \to \mathbb{R}_+$ is continuously differentiable and $g(0) = 0$ so that for every $\omega \in \Omega$ we have

$$g(X(\omega)) = \int_0^{X(\omega)} g'(s)\,\mathrm{d}s.$$

Therefore

$$\mathbb{E}[g(X))] = \int_\Omega \int_{\mathbb{R}_+} g'(s)\mathbb{1}_{\{s \le X(\omega)\}}\,\mathrm{d}s\,\mathbb{P}(\mathrm{d}\omega).$$

In addition, the function $F : (\omega, s) \in \Omega \times \mathbb{R}_+ \mapsto g'(s)\mathbb{1}_{\{s \le X(\omega)\}}$ is nonnegative. Hence, by Fubini-Tonneli's theorem for nonnegative functions, we have

$$\int_\Omega g \circ X \, d\mathbb{P} = \int_{\mathbb{R}_+} g'(s) \int_\Omega \mathbb{1}_{\{s \le X(\omega)\}}\,\mathbb{P}(\mathrm{d}\omega)\,\mathrm{d}s = \int_{\mathbb{R}_+} g'(s)\mathbb{P}(X \ge s)\,\mathrm{d}s$$

since $\mathbb{E}\left[\mathbb{1}_{\{s \le X\}}\right] = \mathbb{P}(s \le X)$.

(2) Just take $g(x) = x$ in the previous question.

□

* * *

Exercise 4.

(1) Let $X : (\Omega, \mathcal{A}) \to \mathbb{R} \cup \{\pm\infty\}$ a random variable. Let $\mathcal{B} \subset \mathcal{A}$ be a σ-field such that for every $B \in \mathcal{B}$ we have $\mathbb{P}(B) = 0$ or $\mathbb{P}(B) = 1$. Assume that $X : (\Omega, \mathcal{B}) \to \mathbb{R} \cup \{\pm\infty\}$ is measurable. Show that there exists a (deterministic) constant $c \in \mathbb{R} \cup \{\pm\infty\}$ such that $\mathbb{P}(X = c) = 1$.

(2) (Application) Let (X_n) be a sequence of independent real-valued random variables and let (a_n) be a deterministic sequence with $a_n \to 0$ as $n \to \infty$. Show that there exists (deterministic) constants $C_\pm \in [-\infty, \infty]$ such that

$$\liminf_{n\to\infty} a_n \cdot (X_1 + \cdots + X_n) = C_- \quad \text{and} \quad \limsup_{n\to\infty} a_n \cdot (X_1 + \cdots + X_n) = C_+$$

almost surely.

Solution:

(1) Set

$$c = \inf\{x \in \mathbb{R} \colon \mathbb{P}(X \le x) > 0\}$$

with the usual convention that the infimum of an empty set is ∞. By assumption of measurability of X with respect to $\mathcal{B}$, for every $x \in \mathbb{R}$, $\{\omega \in \Omega : \{X(\omega) \le x\}\} \in \mathcal{B}$, so $\mathbb{P}(X \le x) = 0$ or $\mathbb{P}(X \le x) = 1$.

Thus, for every $a < c$ we have $\mathbb{P}(X \le a) = 0$ and for every $b > c$ we have $\mathbb{P}(X \le b) = 1$. Thus for every $n \ge 1$:

$$\mathbb{P}(c - 1/n \le X \le c + 1/n) = 1,$$

with the convention $\infty + x = \infty$ and $-\infty + x = -\infty$. Thus, by taking a countable decreasing intersection:

$$\mathbb{P}(X = c) = \mathbb{P}\left(\bigcap_{n\ge 1}\{c - 1/n \le X \le c + 1/n\}\right) = \lim_{n\to\infty} \mathbb{P}(c - 1/n \le X \le c + 1/n) = 1.$$

(2) Let $\mathcal{F}_m = \sigma(X_i : i \ge m)$ and $\mathcal{B} = \cap_{m\ge 1}\mathcal{F}_m$ be the tail σ-algebra. Since the (X_n) are assumed to be independent, we have by Kolmogorov's zero-one law that $\mathbb{P}(A) \in \{0, 1\}$ for all $A \in \mathcal{B}$. We then apply question (1) with the random variables

$$\liminf_{n\to\infty} a_n \cdot (X_1 + \cdots + X_n) \qquad \text{and} \qquad \limsup_{n\to\infty} a_n \cdot (X_1 + \cdots + X_n)$$

which are $\mathcal{B}$ measurable.

Indeed, let us check that for every $m \ge 1$ the random variable $\liminf_{n\to\infty} a_n \cdot (X_1 + \cdots + X_n)$ is $\mathcal{F}_m$-measurable (the same reasoning works with $\limsup_{n\to\infty} a_n \cdot (X_1 + \cdots + X_n)$). Since $a_n \to 0$ we have for all values of n such that $n \ge m$:

$$\begin{aligned}\liminf_{n\to\infty} a_n \cdot (X_1 + \cdots + X_n) &= \liminf_{n\to\infty} (a_n \cdot (X_1 + X_{m+1} + \cdots + X_{m-1}) + a_n \cdot (X_m + X_{m+1} + \cdots + X_n)) \\ &= 0 + \liminf_{n\to\infty} a_n \cdot (X_m + X_{m+1} + \cdots + X_n) = \liminf_{n\to\infty} a_n \cdot (X_m + X_{m+1} + \cdots + X_n).\end{aligned}$$

Here we have the fact that if x_n and y_n are two sequences of real numbers such that $x_n \to 0$ then $\liminf_{n\to\infty}(x_n + y_n) = \liminf_{n\to\infty} y_n$. For every $n \ge m$ the random variable $a_n \cdot (X_m + X_{m+1} + \cdots + X_n)$ is $\sigma(X_m, \ldots, X_n)$ measurable and thus $\mathcal{F}_m$-measurable. As a consequence, $\liminf_{n\to\infty} a_n \cdot (X_m + X_{m+1} + \cdots + X_n)$ is $\mathcal{F}_m$-measurable as a limit of $\mathcal{F}_m$-measurable functions.

The result then follows from the previous question by setting $X = \liminf_{n\to\infty} a_n \cdot (X_1 + \cdots + X_n)$.

□

3 More involved exercises (optional, will not be covered in the exercise class)

Exercise 5. Let X be a random variable taking values in $\mathbb{N} = \{0, 1, 2, \ldots\}$ and let (X_n) be a sequence of i.i.d. random variables with the same law as X.

(1) Show that $\mathbb{E}[X] = \sum_{n\ge 1} \mathbb{P}(X \ge n)$.

Now assume that $\mathbb{E}[X] = \infty$

(2) Show that $\limsup_{n\to\infty} X_n/n \ge k$ almost surely for all $k \in \mathbb{N}$.

Hint. First show that $\sum_{n\ge 1} \mathbb{P}(X \ge nk) = \infty$.

(3) Deduce that $\limsup_{n\to\infty} X_n/n = \infty$ almost surely.

Now consider any real-valued random variable Y satisfying $\mathbb{E}[|Y|] = \infty$ and let (Y_n) be i.i.d. random variables, each of which has the same law as Y.

(4) Using the previous questions, show that $\limsup_{n\to\infty} |Y_n|/n = \infty$ almost surely.

(5) Deduce that $\limsup_{n\to\infty} |Y_1 + \cdots + Y_n|/n = \infty$ almost surely.

Solution:

(1) We have

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{n\geq 1} \mathbb{1}_{X\geq n}\right] = \sum_{n\geq 1} \mathbb{P}(X \geq n)\,,$$

where the last equality holds by a result seen in the lecture (it is a consequence of the Monotone Convergence Theorem applied with the partial series of $\sum \mathbb{1}_{X\geq n}$.

(2) We have

$$\infty = \mathbb{E}[X] = \sum_{n\geq 1} \mathbb{P}(X \geq n) = \sum_{n\geq 0} \sum_{i=nk+1}^{n(k+1)} \mathbb{P}(X \geq i) \leq k + k \sum_{n\geq 1} \mathbb{P}(X \geq nk).$$

Since (X_n) is an i.i.d. sequence, we have that $\sum_{n\geq 1} \mathbb{P}(X_n/n \geq k) = \infty$, abd the events $\{X_n/n \geq k\}$ are independent. So by the second Borel-Cantelli Lemma, we have that almost surely the event $\{X_n/n \geq k\}$ occurs infinitely many times, and therefore

$$\limsup_{n\to\infty} X_n/n \geq k \quad \text{almost surely}\,.$$

(3) This follows from the equality of events $\{\limsup_{n\to\infty} X_n/n = \infty\} = \cap_{k\geq 1}\{\limsup_{n\to\infty} X_n/n \geq k\}$ (this is a countable intersection of sets of measure 1 which again has measure 1).

(4) Let us define $X_n := \lfloor |Y_n| \rfloor$ for every $n \geq 1$. We observe that (X_n) is an i.i.d. sequence of non-negative integer valued random variables satisfying $X_n \leq |Y_n| \leq X_n + 1$. Hence $\mathbb{E}[X_n] \geq \mathbb{E}[|Y_n|] - 1 = \infty$ and by question (3)

$$\limsup_{n\to\infty} |Y_n|/n \geq \limsup_{n\to\infty} X_n/n = \infty.$$

(5) Let us observe that

$$|Y_n| = |(Y_1 + \cdots + Y_n) - (Y_1 + \cdots + Y_{n-1})| \leq |Y_1 + \cdots + Y_n| + |Y_1 + \cdots + Y_{n-1}|.$$

Then, by the previous part

$$\infty = \limsup_{n\to\infty} |Y_n|/n \leq 2 \limsup_{n\to\infty} |Y_1 + \cdots + Y_n|/n$$

almost surely.

□

***

Exercise 6. (The Doob–Dynkin lemma does not hold for general σ-algebras) Set $E = F = \{1,2\}$ and equip F with the σ-algebra $\{\varnothing, F\}$. Let $f : E \to F$ be defined by $f(1) = f(2) = 1$. Find a measurable function $g = (E, \sigma(f)) \to (\mathbb{R}, \{\varnothing, \mathbb{R}\})$ which cannot be written in the form $g = h \circ f$ with $h : (F, \mathcal{F}) \to (\mathbb{R}, \{\varnothing, \mathbb{R}\})$ measurable.

Remark. Here $\mathbb{R}$ is equipped with the trivial σ-algebra $\{\varnothing, \mathbb{R}\}$ and not with the Borel σ-algebra.

Solution:

Note that $\sigma(f) = \{\varnothing, E\}$, so that any function $g = (E, \sigma(f)) \to (\mathbb{R}, \{\varnothing, \mathbb{R}\})$ is actually measurable. Set $g(1) = 1$ and $g(2) = 2$. Since $f(1) = f(2) = 1$, there does not exist $h : (F, \mathcal{F}) \to (\mathbb{R}, \{\varnothing, \mathbb{R}\})$ measurable such that $g = h \circ f$.

□

***

Exercise 7.

(1) Let f be a differentiable function on $[0,1]$, with bounded derivative f'. Show that

$$\int_0^1 f'(x)\mathrm{d}x = f(1) - f(0).$$

(2) Find a continuous function, almost everywhere differentiable on $[0,1]$ such that $f(0) = 0, f(1) = 1$ and $\int_0^1 f'(x)\mathrm{d}x = 0$.

Solution:

(1) Define on $[0,1]$ the sequence $(g_n)_{n\geq 1}$ by $g_n(x) = n(f(x+1/n) - f(x))$ if $x \leq 1 - 1/n$ and 0 otherwise. For fixed $x \in [0,1[$, $g_n(x)$ converges to $f'(x)$. Set $M = \sup_{x\in[0,1]} |f'(x)|$, which is finite by hypothesis. By the mean value theorem, $|g_n(x)| \leq M$ for every $x \leq 1 - 1/n$, and this inequality is clearly true for $x \in [1 - 1/n, 1]$. Hence $|g_n(x)| \leq M$ for every $x \in [0,1]$. By the dominated convergence theorem,

$$\int_0^1 f'(x)\mathrm{d}x = \lim_{n\to\infty} \int_0^1 g_n(x)\mathrm{d}x.$$

But, by writing $F(x) = \int_0^x f(t)dt$,

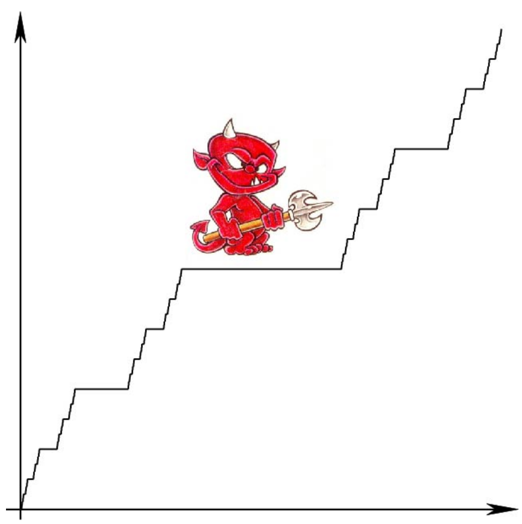
$$\begin{aligned}\int_0^1 g_n(x)\mathrm{d}x &= n\int_0^{1-1/n} f(x+1/n)\mathrm{d}x - n\int_0^{1-1/n} f(x)\mathrm{d}x \\ &= n\int_{1/n}^1 f(x)\mathrm{d}x - n\int_0^{1-1/n} f(x)\mathrm{d}x \\ &= n(F(1)-F(1-1/n)) - n(F(1/n)-F(0)),\end{aligned}$$

which converges to $f(1)-f(0)$ as $n\to\infty$. Indeed, the continuity of f at 0 and 1 implies that $F'(0)=f(0)$ and $F'(1)=f(1)$. The desired result follows.

Remark:

- Set $G(t)=\int_0^t f'(u)du$. By writing $(G(t+\epsilon)-G(t))/\epsilon = \int_0^1 f'(t+\epsilon u)du$, one cannot use the dominated convergence theorem to say that this quantity converges to $f'(t)$ as $\epsilon\to 0$. Indeed, we have no hypothesis on the continuity of f'.

(2) An example: the devil's staircase:



This is a continuous increasing function f such that $f(0)=0$, $f(1)=1$, whose derivative is 0 on the complement of the triadic Cantor set introduced in the exercise sheet of week 2. Therefore $f'=0$ almost everywhere! □

* * *

Exercise 8. Let $\varphi : ([0,1],\mathcal{B}([0,1])) \to (\mathbb{R},\mathcal{B}(\mathbb{R}))$ be an integrable function for the Lebesgue measure. Define $G:\mathbb{R}\to\mathbb{R}_+$ by

$$G(t) = \int_{[0,1]} |\varphi(x)-t|\mathrm{d}x.$$

(1) Show that G is continuous.

(2) Show that G is differentiable at $t\in\mathbb{R}$ if and only if $\lambda(\{\varphi=t\})=0$, where λ denotes the Lebesgue measure.

Solution:

(1) Let $t_n \to t$. Then $|\varphi(x) - t_n| \to |\varphi(x) - t|$ and $|\varphi(x) - t_n| \leq |\varphi(x)| + 1$, which is an integrable function on $[0,1]$ with respect to the Lebegsue measure, independent of n. We conclude by the dominated convergence theorem.

(2) For every $t \in \mathbb{R}$ and $h \neq 0$, we have

$$\frac{G(t+h) - G(t)}{h} = \int_0^1 \frac{|\varphi(x) - t - h| - |\varphi(x) - t|}{h} \, dx.$$

Let $(h_n)_{n\geq 0}$ be a sequence of positive real numbers, decreasing towards 0. For every $x \in [0,1]$, we have

$$\frac{|\varphi(x) - t - h_n| - |\varphi(x) - t|}{h_n} \xrightarrow[n\to\infty]{} \mathbb{1}_{[\varphi(x),+\infty[}(t) - \mathbb{1}_{]-\infty,\varphi(x)[}(t),$$

and $|\varphi(x)-t-h_n|-|\varphi(x)-t| \leq h_n$. Hence, by dominated convergence, the fonction G is differentiable to the right of t and

$$G'_{\text{right}}(t) = \lambda(\{\varphi \leq t\}) - \lambda(\{\varphi > t\}).$$

Similarly, G is differentiable to the left of t and

$$G'_{\text{left}}(t) = \lambda(\{\varphi < t\}) - \lambda(\{\varphi \geq t\}).$$

Therefore $G'_{\text{right}}(t) - G'_{\text{left}}(t) = \lambda(\{\varphi = t\})$, which gives the result.

□

4 Fun exercise (optional, will not be covered in the exercise class)

Exercise 9.

(1) In the logicians' prison, the following game is proposed to 100 prisoners. The 100 prisoners are lined up in a single line, so that each logician can see the logicians in front of her but not those behind her. A black or white hat is placed on the head of each of them, and then the logicians are asked one by one to guess the color of their hat. They start by asking the one at the back (the one who can see 99 hats), and then they work their way up to the one at the front (the one who can't see any). Each logician hears all the previous answers. If a logician correctly guesses the color of his hat, she is released. If she does not, she is sentenced to death. The logicians have the right to confer before the game and establish a strategy. What strategy can minimize the number of deaths?

(2) In Hilbert's prison, there is an infinite (countable) number of logicians. It is decided to make them play the same hat game. Again, each logician can see all the hats in front of her (now an infinite number), and must determine the color of her own hat. What strategy can save the maximum number of

logicians?

Solution:

(1) Obviously, the first to speak has no information, and whatever she says has a 50/50 chance of finding the color of her hat. So the best we can do is the strategy described below.

The first to speak counts the number of black hats among the 99 she sees. If this number is even, she says "white", and if it's odd, she says "black". The second speaker then counts the number of black hats in front of him, and can compare the parity of this number with what the last speaker said. If it's the same parity, she knows she has a white hat, and if not, she knows she has a black hat. Then the third player deduces the color of her hat from the two previous answers, and so on.

This allows to save 99 prisoners for sure.

(2) We begin by associating with each hat arrangement a number between 0 and 1 in binary basis:

$$x = 0, a_1 a_2 \dots a_n a_{n+1} \dots,$$

where $a_k = 0$ if the k-th hat is white, and $a_k = 1$ otherwise. We introduce the equivalence relation $\sim$ on $\mathbb{R}$ such that

$$x \sim y \quad \Longleftrightarrow \quad x - y \text{ has a finite number of digits in binary basis.}$$

In other words, $x \sim y$ if their writing in binary basis ends with the same infinity of digits. It's easy to check that is an equivalence relation. For each equivalence class $C \in \mathbb{R}/\sim$, logicians agree on a representative $x_C \in C$ (which is possible thanks to the axiom of choice).

During the game, since each logician sees all the hats in front of her, everyone knows which equivalence class C their configuration x belongs to . Furthermore, if $n \geq 1$ is the rank from which $x_m = C_m$ for every $m \geq n$, the $n-2$ logicians also know the integer n.

The first $n-1$ logicians can deduce the color of their hat using the previous strategy. The rest of the logicians recite the binary digits of x_C the k-th person will say "white" if the k-th number in the binary representation of x_C is a 0 and "black" otherwise.

This allows for sure to save all prisoners but one!

$\square$