THE KOLMOGOROV STUDENTS' COMPETITIONS ON PROBABILITY THEORY

A.S. Cherny

Moscow State University,
Faculty of Mechanics and Mathematics,
Department of Probability Theory,
119992 Moscow, Russia.
E-mail: `cherny@mech.math.msu.su`
Webpage: `http://mech.math.msu.su/~cherny`

Abstract. The paper contains the problems and the solutions of the four Kolmogorov Students' Competitions on Probability Theory. It also describes the connection of some problems with modern topics of scientific research.

Key words and phrases. Arbitrage pricing theory, coherent risk measures, convergence of σ-fields, financial mathematics, fundamental theorems of asset pricing, game-theoretic approach to statistics, Good Deals pricing, Lévy processes, percolation theory, randomization, stochastic orders, stochastic resonance.

1 Introduction

The Department of Probability Theory, Faculty of Mechanics and Mathematics, Moscow State University organized the Kolmogorov Students' Competitions on Probability Theory. Up to the present time, four such competitions have been held: in 2001, 2003, 2004, and 2005. The information about these competitions (including the problems and the names of the winners) can be found in the journal "Theory of Probability and Its Applications" [2]–[5] as well as on the Website of the Department of Probability Theory: `http://mech.math.msu.su/probab`.

Since 2003, the competitions have been held at the end of April as April 25 is the birthday of Andrey Nikolaevich Kolmogorov. In particular, the second competition was dedicated to the centennial of A.N. Kolmogorov.

The idea of conducting these competitions belongs to A.N. Shiryaev who was the chairman of the Organizing Committees of all four competitions. The members and Ph.D. students of the Department of Probability Theory: A.S. Cherny, S.V. Dilman, I.N. Medvedev, A.S. Mishchenko, A.V. Selivanov, A.P. Shashkin, and M.A. Urusov have made valuable contributions to arranging the competitions.

The number of students who took part in the First, Second, Third, and Fourth competitions is 23, 42, 93, and 110, respectively. Since the second one, the competitions have been arranged separately for the II-year students and for the students of the III–V years. The fourth competition was attended by students from Moscow, Kiev, Saint Petersburg, Samara, Tomsk, and Vologda.

The solutions of the problems are usually announced at the Big Seminar of the Department of Probability Theory. All the participants get prizes and the winners get special awards.

Most of the problems are of the "competition type", i.e. they admit a simple short solution, but it is hard to find it. However, at the fourth competition we have included problems that test the knowledge of some basic probabilistic concepts (see Problem 4.3). Some of the problems are in fact statements of the general theory that are well known to the specialists (for example, Problems 1.6, 1.8, 2.9, and 4.5). Some other problems are closely connected with topics of the modern research (for example, Problems 1.7, 2.10, 2.11, and 4.6).

In this paper, we give the solutions to all the problems. If available, several solutions are given. Moreover, we describe the origin of some problems as well as their connections with topics of the modern research in probability theory.

Of course, the reader is invited to solve as many problems as possible before passing on to the solutions.

We conclude the introduction by a brief survey of the literature related to problems in probability. Nice collections of problems are given in [21], [36], [46], [82]. Let us also mention the books [75] and [84] containing counterexamples in probability, statistics, and random processes. A very interesting collection of "probabilistic paradoxes" is given in [84].

We would also like to draw the reader's attention to the report [1], which contains a fascinating description of promising directions of future research in probability. Let us also mention the book [38], providing a description of promising research directions in mathematics in the 21st century.

Acknowledgement. I am thankful to I. Pavlyukevich for consultations about stochastic resonance and for having provided simulated paths illustrating this phenomenon (see Figure 7).

2 Problems

2.1 First Kolmogorov Students' Competition

Problem 1.1. $\left(\frac{18}{23}\right)$[1] Let $X = (X^1, X^2)$ be a 2-dimensional random vector with a continuous distribution, i.e. $\mathsf{P}(X = x) = 0$ for any $x \in \mathbb{R}^2$. Is it true that its distribution function $F(x^1, x^2) = \mathsf{P}(X^1 \le x^1, X^2 \le x^2)$ is continuous?

Problem 1.2. $\left(\frac{9}{23}\right)$ Let $X_1, X_2, \dots$ be a sequence of independent random variables that converges in probability as $n \to \infty$ to a random variable X. Prove that X is degenerate, i.e. there exists $x \in \mathbb{R}$ such that $X \overset{\text{a.s.}}{=} x$.

Problem 1.3. $\left(\frac{7}{23}\right)$ Give an example of 4 dependent random events A_1, A_2, A_3, A_4 such that any 3 of them are independent.

Problem 1.4. **a)** $\left(\frac{9}{23}\right)$ Let X and Y be square-integrable random variables such that $\mathsf{E}(X \mid Y) = Y$ and $\mathsf{E}(Y \mid X) = X$. Prove that $X \overset{\text{a.s.}}{=} Y$.

[1] The number in brackets is the *solvability coefficient*, i.e. the fraction whose numerator is the number of the students who solved the problem and whose denominator is the number of the students to whom the problem was proposed.

b) $\left(\frac{0}{23}\right)$ Let X and Y be integrable random variables such that $\mathsf{E}(X \mid Y) = Y$ and $\mathsf{E}(Y \mid X) = X$. Prove that $X \overset{\text{a.s.}}{=} Y$.

Problem 1.5. $\left(\frac{7}{23}\right)$ Let $(\mathsf{P}_n)_{n=1}^{\infty}$ be a sequence of probability measures on the real line such that

$$\forall \lambda \in \mathbb{Q}, \quad \int_{\mathbb{R}} e^{i\lambda x} \mathsf{P}_n(dx) \xrightarrow[n\to\infty]{} 1.$$

Is it true that $\mathsf{P}_n \xrightarrow{w} \delta_0$, where δ_0 is the delta-mass concentrated at zero?

Problem 1.6. $\left(\frac{3}{23}\right)$ Let P and Q be positive finite measures on the real line with no mass at zero (i.e. $\mathsf{P}(\{0\}) = 0$ and $\mathsf{Q}(\{0\}) = 0$) such that

$$\forall \lambda \in \mathbb{R}, \quad \int_{\mathbb{R}} (e^{i\lambda x} - 1) \mathsf{P}(dx) = \int_{\mathbb{R}} (e^{i\lambda x} - 1) \mathsf{Q}(dx).$$

Is it true that $\mathsf{P} = \mathsf{Q}$?

Problem 1.7. $\left(\frac{1}{23}\right)$ Let X be a bounded random variable on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let $\mathcal{F}_n \subseteq \mathcal{G}_n \subseteq \mathcal{H}_n$ be three sequences of sub-σ-fields of $\mathcal{F}$. It is known that there exists a random variable Y such that

$$\mathsf{E}(X \mid \mathcal{F}_n) \xrightarrow[n\to\infty]{\mathsf{P}} Y, \quad \mathsf{E}(X \mid \mathcal{H}_n) \xrightarrow[n\to\infty]{\mathsf{P}} Y.$$

Prove that $\mathsf{E}(X \mid \mathcal{G}_n) \xrightarrow{\mathsf{P}} Y$.

Problem 1.8. $\left(\frac{6}{23}\right)$ A random variable X is called infinitely divisible if for any $n \in \mathbb{N}$, there exist independent identically distributed random variables $X_1^n, \dots, X_n^n$ such that $X_1^n + \dots + X_n^n$ has the same distribution as X. Prove that a random variable with the distribution density $p(x) = |x| I(|x| \le 1)$ is not infinitely divisible.

2.2 Second Kolmogorov Students' Competition

Problem 2.1. $\left(\frac{38}{42}\right)$ On a table there are a black hat and a white hat containing the lottery tickets. The white hat is "better" in the sense that when one pulls a ticket out of this hat, the probability to get a lucky ticket is higher than when one pulls a ticket out of the black hat. On another table there are also a black hat and a white hat with lottery tickets, and the white hat is "better" in the sense described above. Let us imagine that the tickets from two white hats are put in one big white hat; the tickets from two black hats are put in one big black hat. Is is true that the big white hat is "better" than the big black hat in the sense described above?

Problem 2.2. $\left(\frac{32}{42}\right)$ Let $A, B, C_1, \dots, C_n$ be events on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Suppose that $\bigcup_{i=1}^n C_i = \Omega$ and $\mathsf{P}(C_i) > 0$, $\mathsf{P}(A \mid C_i) \ge \mathsf{P}(B \mid C_i)$ for any $i = 1, \dots, n$. Is it true that $\mathsf{P}(A) \ge \mathsf{P}(B)$?

Problem 2.3. $\left(\frac{10}{42}\right)$ Eight boys and seven girls bought tickets to the cinema in a 15-seat row. Assuming that all the 15! possible seatings are equally probable, compute the expected number of the pairs of female and male neighbors (for example, in the seating "b,b,b,b,b,b,b,g,b,g,g,g,g,g,g" there are three such pairs).

Problem 2.4. a) $\left(\frac{10}{29}\right)$[2] Let K be the unit circumference, P be a probability measure on K such that $\mathsf{P} \circ \varphi_\alpha^{-1} = \mathsf{P}$ for any $\alpha \in \mathbb{R}$. Here φ_α denotes the rotation of K by the

[2] Different denominators of the solvability coefficients for different problems reflect the fact that some problems were proposed to all the participants, while some others were proposed only to the students of the II year or to the students of the III–V years.

angle α and $\mathsf{P} \circ \varphi_\alpha^{-1}$ denotes the image of P under the map φ_α. Prove that P coincides with the normalized Lebesgue measure.

b) $\left(\frac{6}{29}\right)$ Prove the same assuming that there exists $\alpha \in \mathbb{R}$ such that $\alpha/\pi \notin \mathbb{Q}$ and $\mathsf{P} \circ \varphi_\alpha^{-1} = \mathsf{P}$.

Problem 2.5. a) $\left(\frac{23}{29}\right)$ Let X be a random variable such that $\mathsf{P}(X \neq 0) > 0$. Suppose that, for some real numbers a and b, the random variables aX and bX have the same distribution. Is it true that $a = b$?

b) $\left(\frac{14}{29}\right)$ Is it true that $a = b$ under the additional assumption that $a, b \geq 0$?

Problem 2.6. $\left(\frac{5}{42}\right)$ Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ be a σ-field that includes all the P-null sets from $\mathcal{F}$. Let X be a random variable such that, for any random variable Y that is independent of $\mathcal{G}$, X and Y are independent. Is it true that X is $\mathcal{G}$-measurable?

Problem 2.7. $\left(\frac{10}{42}\right)$ What is the maximal possible value of the variance of a random variable X taking on values in the interval $[0, 1]$?

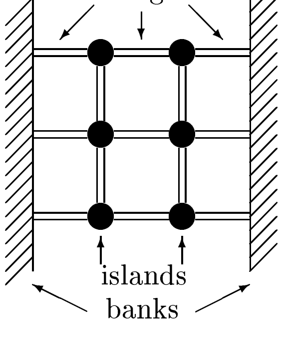
Problem 2.8. a) $\left(\frac{9}{13}\right)$ A sequence (X_n) of random variables converges in probability to X. Is it true that $S_n/n \xrightarrow{\mathsf{P}} X$, where $S_n = X_1 + \cdots + X_n$?

b) $\left(\frac{4}{13}\right)$ Is it true that $S_n/n \xrightarrow{\mathsf{P}} X$ under the additional assumption that $|X_n| \leq 1$ for any n?

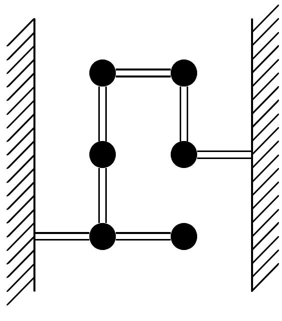
Problem 2.9. $\left(\frac{2}{13}\right)$ Let X_n be a random vector with a uniform distribution on a unit sphere in $\mathbb{R}^n$. (The uniform distribution is characterized by the property that it is invariant under the orthogonal mappings.) Let Y_n denote the first coordinate of X_n. Prove that the sequence $\sqrt{n}Y_n$ converges in distribution to a Gaussian random variable with mean 0 and variance 1.

Problem 2.10. $\left(\frac{2}{13}\right)$ Let $X_1, X_2, \ldots$ be independent Bernoulli random variables with $\mathsf{P}(X_n = -1) = 1/3$, $\mathsf{P}(X_n = +1) = 2/3$. Does there exist a probability measure Q equivalent to P such that $\mathsf{E}_\mathsf{Q} X_n = 0$ for any $n \in \mathbb{N}$? (Recall that Q is equivalent to P if $\mathsf{Q}(A) = 0 \Leftrightarrow \mathsf{P}(A) = 0$.)

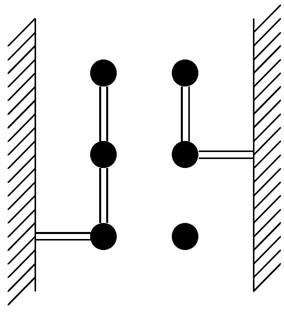
Problem 2.11. $\left(\frac{5}{42}\right)$ On a river there are 6 islands connected by a system of bridges (see Figure 1.a). During the summer flood a part of the bridges has been destroyed. Each bridge is destroyed with probability 1/2, independently of the other bridges. What is the probability that after the described flood it is possible to cross the river using the remaining bridges? (In the case shown in Figure 1.b it is possible to cross the river; in the case shown in Figure 1.c it is not possible to cross the river.)



`Figure 1.a`



`Figure 1.b`



`Figure 1.c`

2.3 Third Kolmogorov Students' Competition

Problem 3.1. $\left(\frac{88}{93}\right)$ Let A and B be events such that $0 < \mathsf{P}(A) < 1$ and $\mathsf{P}(B \mid A) = \mathsf{P}(B \mid A^c)$ (A^c denotes the complement to A). Is it true that A and B are independent?

Problem 3.2. a) $\left(\frac{61}{93}\right)$ Let X, Y, Z be random variables on some probability space. Suppose that Y stochastically dominates X, i.e. $\mathsf{P}(X \le x) \ge \mathsf{P}(Y \le x)$ for any $x \in \mathbb{R}$. Is it true that $Y + Z$ stochastically dominates $X + Z$?

b) $\left(\frac{29}{93}\right)$ Is the above statement true under the additional assumption that X, Y are independent and Y, Z are also independent?

Problem 3.3. $\left(\frac{59}{93}\right)$ One hundred passengers bought tickets in a 100-seat carriage. One seat was reserved for each passenger. The first 99 passengers took the seats at random, so that all 100! variants of their seating have the same probability. However, the last passenger decided to take his/her reserved seat and asked the passenger who had taken his/her seat (if it was occupied) to change the seat. The passenger who was disturbed asked the passenger who had taken his/her seat (if it was occupied) to change the seat, and so on. Compute the expected number of the passengers who were disturbed (the hundredth passenger is not included in this number).

Problem 3.4. $\left(\frac{30}{93}\right)$ We have two dice with their sides marked by numbers $1, \dots, 6$. Is it possible to attach the probabilities of occurrence to the sides of each die (these probabilities might be different for these two dice) in such a way that the sum of the occurred numbers after a simultaneous throw of both dice has the uniform distribution on the set $\{2, \dots, 12\}$?

Problem 3.5. $\left(\frac{0.5}{53}\right)$ Let X and Y be independent random variables such that $\mathsf{E}|X + Y| < \infty$. Is it true that $\mathsf{E}|X| < \infty$?

Problem 3.6. $\left(\frac{0}{40}\right)$ Let $X_1, X_2, \dots$ be independent identically distributed random variables. Set $S_n = X_1 + \dots + X_n$. Suppose that $\limsup\limits_{n\to\infty} \frac{X_n}{n} < \infty$ a.s. Is it true that $\limsup\limits_{n\to\infty} \frac{S_n}{n} < \infty$ a.s.?

Problem 3.7. $\left(\frac{33}{93}\right)$ Let A be a Borel set on a circumference such that $\mu(A) = 2/3$, where μ is the uniform distribution (i.e. the normalized Lebesgue measure) on the circumference. The points of A are marked red, while the points of its complement are marked blue. Prove that it is possible to inscribe a square in the circumference in such a way that at least its 3 vertices are red.

Problem 3.8. $\left(\frac{1}{93}\right)$ Let $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ be two sequences of random variables on some probability space. Suppose that X_n and Y_n are independent for every $n \in \mathbb{N}$ and $X_n + Y_n \xrightarrow{\mathsf{P}} 0$. Prove that there exist real numbers $a_1, a_2, \dots$ such that $X_n - a_n \xrightarrow{\mathsf{P}} 0$.

Problem 3.9. $\left(\frac{3}{40}\right)$ Let $X_1, X_2, \dots$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Suppose that $X_n \xrightarrow{\text{a.s.}} 0$ and $|X_n| \le 1$ for any $n \in \mathbb{N}$. Let $\mathcal{G}_1, \mathcal{G}_2, \dots$ be sub-σ-fields of $\mathcal{F}$. Is it true that $\mathsf{E}(X_n \mid \mathcal{G}_n) \xrightarrow{\text{a.s.}} 0$?

Problem 3.10. $\left(\frac{0}{93}\right)$ Let μ be a probability measure on the Borel σ-field $\mathcal{B}$ of the unit circumference. Let X and Y be independent random points on the circumference with the distribution μ (i.e. $\mathsf{P}(X \in A, Y \in B) = \mu(A)\mu(B)$ for all $A, B \in \mathcal{B}$). Denote by α the angle between X and Y (so that $\alpha \in [0, \pi]$). Prove that $\mathsf{P}(\alpha \le 2\pi/3) \ge 1/2$.

2.4 Fourth Kolmogorov Students' Competition

Problem 4.1. $\left(\frac{49}{56}\right)$ The riflemen A and B are shooting at a target in turn. (A shoots first.) The rifleman A hits the target with probability P_A, while the rifleman B hits the target with probability P_B (the results of different shots are independent). The winner is the one who hits the target first. Find the probability that A wins.

Problem 4.2. $\left(\frac{52}{110}\right)$ Let $X_1, X_2, \ldots$ be independent identically distributed random variables. Set $Y_n = \frac{X_n + X_{n+1}}{1+|X_n+X_{n+1}|}$, $S_n = Y_1 + \cdots + Y_n$. Prove that there exists a constant c such that $\frac{S_n}{n} \xrightarrow{\text{a.s.}} c$.

Problem 4.3. $\left(\frac{15.5}{54}\right)$ Let X be a random variable that it uniformly distributed on $[0, \pi]$. Find $\mathsf{E}(X \mid \sin X)$.

Problem 4.4. $\left(\frac{56}{110}\right)$ Let X, Y, Z be independent uniformly distributed points on a circumference (i.e. $\mathsf{P}(X \in A, Y \in B, Z \in C) = \mu(A)\mu(B)\mu(C)$, where μ is the normalized Lebesgue measure). Find the probability that the triangle XYZ is acute.

Problem 4.5. **a)** $\left(\frac{84.5}{110}\right)$ Let X, Y, ξ be independent random variables with $\mathsf{P}(\xi = \pm 1) = 1/2$. Prove that $|X + \xi Y|$ and $|Y + \xi X|$ have the same distribution.

b) $\left(\frac{13.5}{110}\right)$ Let X, Y, Z, ξ, η be independent random variables with $\mathsf{P}(\xi = \pm 1) = \mathsf{P}(\eta = \pm 1) = 1/2$. Prove that $|X + \xi|Y + \eta Z||$ and $||X + \xi Y| + \eta Z|$ have the same distribution.

Problem 4.6. **a)** $\left(\frac{9}{110}\right)$ Let X be a bounded random variable on some probability space and λ be a number from $(0, 1]$. Denote $u_\lambda(X) = \inf \mathsf{E}(ZX)$, where inf is taken over the set $\mathcal{D}_\lambda$ of random variables Z such that $0 \le Z \le \lambda^{-1}$ and $\mathsf{E}Z = 1$. Prove that there exists a random variable $Z_* \in \mathcal{D}_\lambda$ such that $\mathsf{E}(Z_* X) = u_\lambda(X)$.

b) $\left(\frac{10.5}{110}\right)$ Let X, Y be independent nondegenerate bounded random variables and λ be a number from $(0, 1)$ (recall that X is degenerate if there exists a constant c such that $X \overset{\text{a.s.}}{=} c$). Is it true that $u_\lambda(X + Y) > u_\lambda(X) + u_\lambda(Y)$?

c) $\left(\frac{5}{110}\right)$ Let X, Y be independent nondegenerate random variables taking on a finite number of values and let μ be a probability measure on $[0, 1]$ such that $\mu((a, b)) > 0$ for any $0 \le a < b \le 1$. Prove that

$$\int_{[0,1]} u_\lambda(X + Y)\mu(d\lambda) > \int_{[0,1]} u_\lambda(X)\mu(d\lambda) + \int_{[0,1]} u_\lambda(Y)\mu(d\lambda). \tag{1}$$

d) $\left(\frac{0}{110}\right)$ Let X, Y be independent nondegenerate bounded random variables and let μ be a probability measure on $[0, 1]$ such that $\mu((a, b)) > 0$ for any $0 \le a < b \le 1$. Prove that (1) is true.

Problem 4.7. $\left(\frac{3}{54}\right)$ Let B be a Brownian motion and $t_1, t_2, \ldots$ be a sequence of positive numbers with $t_n \to \infty$. Is it true that $\limsup\limits_{n\to\infty} \frac{B_{t_n}}{\sqrt{2t_n \ln\ln t_n}} \overset{\text{a.s.}}{=} 1$?

Problem 4.8. $\left(\frac{4.5}{110}\right)$ Let X, Y, Z be independent identically distributed random vectors in $\mathbb{R}^n$ taking on a finite number of values. Let L denote the (random) linear subspace of $\mathbb{R}^n$ generated by X and Y and let $d(L)$ denote its dimension. We assume that $\mathsf{P}(d(L) = 2) > 0$ and $\mathsf{P}(d(L) = 1) > 0$. Is it true that $\mathsf{P}(Z \in L \mid d(L) = 2) \ge \mathsf{P}(Z \in L \mid d(L) = 1)$?

3 Solutions

3.1 First Kolmogorov Students' Competition

Problem 1.1. The answer is negative. Consider a random vector X such that $X^1 = 0$ and X^2 is uniformly distributed on $[0, 1]$.

Problem 1.2. *First solution.* After passing on to a subsequence, we can assume that $X_n \xrightarrow{\text{a.s.}} X$. Now, the result follows from the Kolmogorov 0–1 law.

Second solution. This solution is more complicated, but it employs a trick that is sometimes useful. (For example, this trick is employed in the proof of the Yamada-Watanabe theorem related to stochastic differential equations; see [72; Ch. IX, Th. 1.7].)

We have $(X_n, X_{n+1}) \xrightarrow{\text{law}} (X, X)$. On the other hand, as X_n and X_{n+1} are independent, $(X_n, X_{n+1}) \xrightarrow{\text{law}} (X, X')$, where X' is an independent copy of X. Hence, $(X, X) \overset{\text{law}}{=} (X, X')$. It is easy to see that this is possible only if X is degenerate.

Problem 1.3. Consider

$$\Omega = \big\{(x_1, x_2, x_3, x_4) : x_i \in \{0, 1\}, x_1 + \dots + x_4 = 0(2)\big\}$$

(recall that the notation $a = b(c)$ means that c divides $b - a$). Let P be the uniform measure (i.e. it attributes mass 1/8 to each elementary outcome). Then the sets

$$A_i = \{(x_1, x_2, x_3, x_4) \in \Omega : x_i = 0\}, \quad i = 1, \dots, 4$$

satisfy the desired conditions.

Problem 1.4. a) We have

$$\mathsf{E}(X - Y)^2 = \mathsf{E}(\mathsf{E}(X^2 - 2XY + Y^2 \mid X)) = \mathsf{E}(X^2 - 2X^2 + \mathsf{E}(Y^2 \mid X)) = \mathsf{E}Y^2 - \mathsf{E}X^2.$$

Similarly, $\mathsf{E}(X - Y)^2 = \mathsf{E}X^2 - \mathsf{E}Y^2$. As a result, $\mathsf{E}(X - Y)^2 = 0$.

b) There exist various solutions of this problem. We give only two of them.

First solution. (Proposed by O. Dragoshansky.) Let us first assume additionally that X and Y are bounded below. For any $n \in \mathbb{N}$, we can write

$$\mathsf{E}(X \wedge n \mid Y) \le \mathsf{E}(X \mid Y) \wedge n = Y \wedge n$$

($x \wedge y$ denotes $\min\{x, y\}$). Consequently,

$$\mathsf{E}(X \wedge n \mid Y \wedge n) \le Y \wedge n.$$

Similarly,

$$\mathsf{E}(Y \wedge n \mid X \wedge n) \le X \wedge n.$$

By taking expectations of two inequalities above, we get that these inequalities are actually equalities. It follows from a) that $X \wedge n \overset{\text{a.s.}}{=} Y \wedge n$. Letting $n \to \infty$, we get $X \overset{\text{a.s.}}{=} Y$.

In a similar way the statement is proved for X and Y bounded above. Employing now the truncation procedure mentioned above to integrable X and Y, we get the desired statement.

Second solution. (Proposed by G. Peskir.) Consider $f(x) = \arctan x$. Then

$$\mathsf{E}f(X)(X - Y) = \mathsf{E}(\mathsf{E}(f(X)X - f(X)Y \mid X)) = \mathsf{E}f(X)X - \mathsf{E}f(X)X = 0.$$

Similarly, $\mathsf{E} f(Y)(X-Y)=0$. Consequently, $\mathsf{E}(f(X)-f(Y))(X-Y)=0$, which yields the desired statement.

`Remark`. The problem admits also the following modification. Assume that X and Y are positive random variables such that $\mathsf{E}(X \mid Y)$ and $\mathsf{E}(Y \mid X)$ are finite a.s. and $\mathsf{E}(X \mid Y)=Y$, $\mathsf{E}(Y \mid X)=X$. Then $X \overset{\text{a.s.}}{=} Y$. This is seen from the first solution.

Problem 1.5. The answer is negative. Consider $\mathsf{P}_n=\delta_{2\pi n!}$.

Problem 1.6. The answer is positive. Without loss of generality, we can assume that $\mathsf{P}(\mathbb{R}) \geq \mathsf{Q}(\mathbb{R})$. Then, for the measure $\widetilde{\mathsf{Q}}=\mathsf{Q}+(\mathsf{P}(\mathbb{R})-\mathsf{Q}(\mathbb{R}))\delta_0$, we have

$$\forall \lambda \in \mathbb{R}, \quad \int_{\mathbb{R}}\big(e^{i\lambda x}-1\big)\mathsf{P}(dx)=\int_{\mathbb{R}}\big(e^{i\lambda x}-1\big)\widetilde{\mathsf{Q}}(dx).$$

As $\mathsf{P}(\mathbb{R})=\widetilde{\mathsf{Q}}(\mathbb{R})$, we get

$$\forall \lambda \in \mathbb{R}, \quad \int_{\mathbb{R}} e^{i\lambda x}\mathsf{P}(dx)=\int_{\mathbb{R}} e^{i\lambda x}\widetilde{\mathsf{Q}}(dx).$$

Hence, $\mathsf{P}=\widetilde{\mathsf{Q}}$, which yields the desired statement.

`Comments`. This problem is closely connected with the *Lévy-Khintchine representation* of the infinitely divisible distributions (see, for example, [76; § 8]). It states that the characteristic function of any infinitely divisible random variable X (see Problem 1.8 for the definition) has the form

$$\mathsf{E} e^{i\lambda X}=\exp\Big\{i\lambda b-\frac{c}{2}\lambda^2+\int_{\mathbb{R}}\big(e^{i\lambda x}-1-i\lambda x I(|x|\leq 1)\big)\nu(dx)\Big\}, \quad \lambda \in \mathbb{R},$$

where $b \in \mathbb{R}$, $c \in \mathbb{R}_+$, and ν is a positive measure on $\mathbb{R}$ such that $\int_{\mathbb{R}} 1 \wedge x^2 \nu(dx)<\infty$. Moreover, such a triplet (b, c, ν) is unique. It is called the *triplet of characteristics* of X.

The infinitely divisible distributions are in one-to-one correspondence with *Lévy processes*. Recall that $(X_t)_{t \geq 0}$ is a Lévy process if

(i) $X_0=0$;
(ii) X has stationary increments (i.e. the distribution of X_t-X_s depends only on $t-s$);
(iii) X has independent increments (i.e. for any $0 \leq t_0 \leq \cdots \leq t_n$, the random variables $X_{t_1}-X_{t_0}, \ldots, X_{t_n}-X_{t_{n-1}}$ are independent);
(iv) X has càdlàg (i.e. right-continuous with left-hand limits) paths.

Clearly, if X is a Lévy process, then each X_t is infinitely divisible. Conversely, for any infinitely divisible distribution Q, there exists a Lévy process X such that $\operatorname{Law} X_1=\mathsf{Q}$ (see [76; § 7]). Thus, there is one-to-one correspondence between the distributions of Lévy processes and the infinitely divisible distributions. The Lévy-Khintchine representation for Lévy processes has the form

$$\mathsf{E} e^{i\lambda X_t}=\exp\Big\{t\Big[i\lambda b-\frac{c}{2}\lambda^2+\int_{\mathbb{R}}\big(e^{i\lambda x}-1-i\lambda x I(|x|\leq 1)\big)\nu(dx)\Big]\Big\}, \quad t \in \mathbb{R}_+,\ \lambda \in \mathbb{R}.$$

Here $(X_t)_{t \geq 0}$ is a Lévy process, $b \in \mathbb{R}$, $c \in \mathbb{R}_+$, and ν is a positive measure on $\mathbb{R}$ such that $\int_{\mathbb{R}} 1 \wedge x^2 \nu(dx)<\infty$. The collection (b, c, ν) is unique and is called the *triplet of characteristics* of the process X. Here b is called the *drift coefficient*, c is called the *diffusion coefficient*, and ν is called the *Lévy measure* of X.

Standard examples of Lévy processes are Brownian motion and Poisson process. Another important example is a compound Poisson process. A *compound Poisson process with the Lévy measure* ν, where ν is a positive finite measure on $\mathbb{R}$, is defined as $X_t = \sum_{n=1}^{N_t} \xi_n$, where N is a Poisson process with intensity $\nu(\mathbb{R})$ and $\xi_1, \xi_2, \dots$ are independent random variables (they are also independent of N) with $\operatorname{Law} \xi_n = \frac{\nu}{\nu(\mathbb{R})}$ (thus, X jumps at the same times as N, and the n-th jump is ξ_n; see Figure 2). The characteristic function of X has the form

$$\begin{aligned}\mathsf{E}e^{i\lambda X_t} &= \sum_{n=0}^{\infty} e^{-\nu(\mathbb{R})t} \frac{(\nu(\mathbb{R})t)^n}{n!} \mathsf{E}e^{i\lambda(\xi_1+\dots+\xi_n)} = \sum_{n=0}^{\infty} e^{-\nu(\mathbb{R})t} \frac{(\nu(\mathbb{R})t)^n}{n!} \Big(\int_{\mathbb{R}} \frac{e^{i\lambda x}}{\nu(\mathbb{R})}\, \nu(dx)\Big)^n \\ &= \exp\Big\{t \int_{\mathbb{R}} (e^{i\lambda x} - 1)\nu(dx)\Big\}\end{aligned}$$

(see [76; § 4] for more information on compound Poisson processes). The left-hand side of this equality determines the finite-dimensional distributions of X (note that X has stationary independent increments). Hence, the equality

$$\forall \lambda \in \mathbb{R}, \quad \int_{\mathbb{R}} (e^{i\lambda x} - 1)\nu(dx) = \int_{\mathbb{R}} (e^{i\lambda x} - 1)\widetilde{\nu}(dx)$$

leads to $\nu = \widetilde{\nu}$. This provides another solution to Problem 1.6.

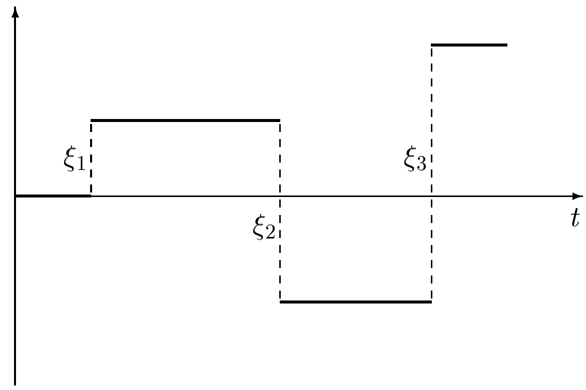


Figure 2. A path of a compound Poisson process

Any Lévy process can be represented as a sum of three independent processes $X^1 + X^2 + X^3$, where $X_t^1 = bt$ is a drift process, $X_t^2 = \sigma B_t$ is a scaled Brownian motion, and X^3 is a "pure jump" process, which can be obtained as a limit of compound Poisson processes. This is known as the *Lévy-Itô decomposition* (see [76; § 19]).

Figure 3 shows simulated paths of some Lévy processes. The first (upper) graph represents Brownian motion. The second one represents the *Cauchy process*, i.e. the Lévy process, for which $X_t - X_s$ is a Cauchy random variable with the density $p_{t-s}(x) = \frac{t-s}{x^2+(t-s)^2}$. It is known that the characteristics of this process have the form $b = 0$, $c = 0$, and $\nu(dx) = \frac{1}{\pi x^2}\, dx$. It is a pure jump process, i.e. it moves only by jumps. The third graph represents the *one-sided Cauchy process*, i.e. the Lévy process with the characteristics $b = 0$, $c = 0$, and $\nu(dx) = \frac{I(x>0)}{x^2}\, dx$. Its qualitative behavior is very interesting: it moves upwards by jumps, and it moves downwards in a continuous manner (i.e. all its jumps are positive). Finally, the fourth graph represents the *Gamma process*, i.e. the Lévy process, for which $X_t - X_s$ has Gamma distribution with the density $p_{t-s}(x) = \frac{x^{t-s-1}e^{-x}}{\Gamma(t-s)} I(x > 0)$.

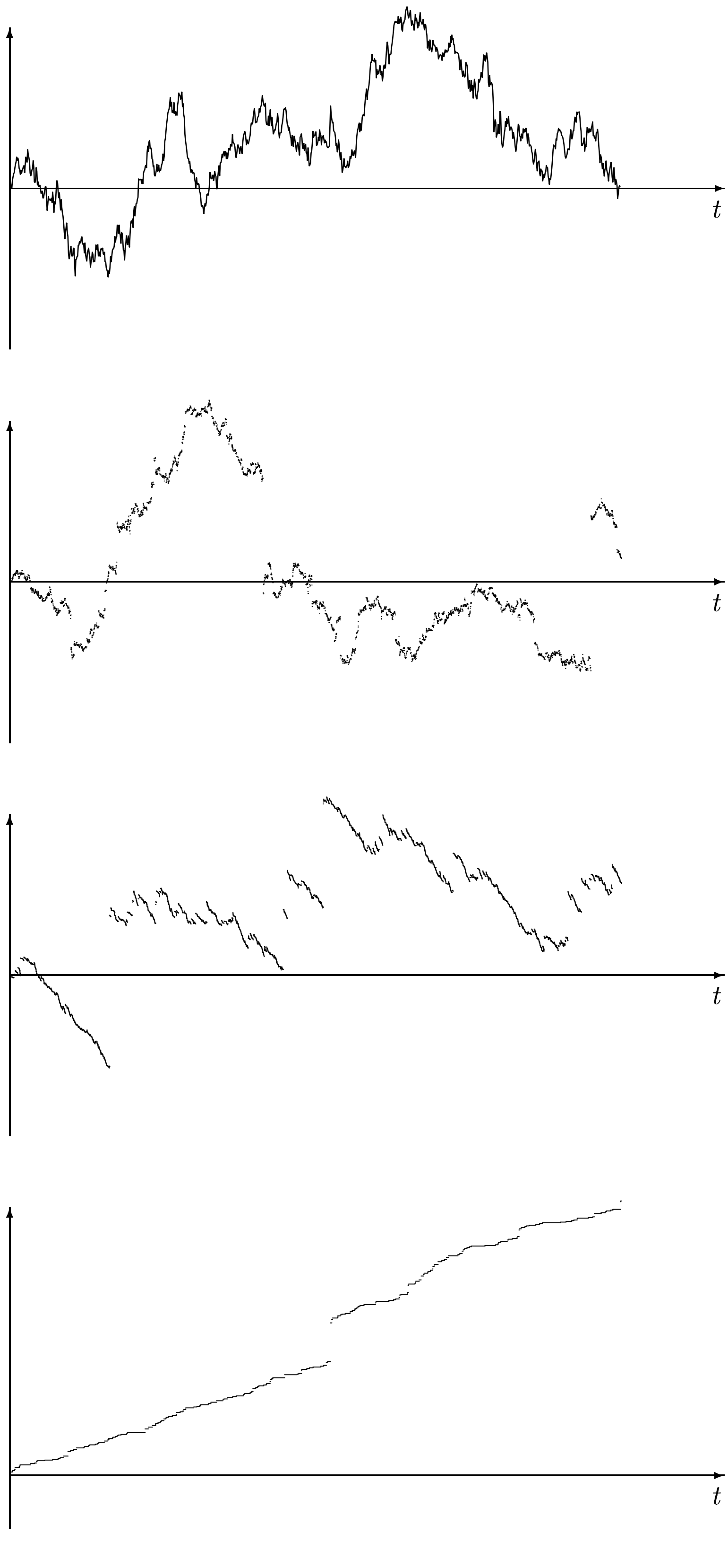


Figure 3. Simulated paths of Lévy processes

This process has increasing paths. Lévy processes with increasing paths are termed *subordinators*.

The theory of Lévy processes is a significant part of theory of random processes (see [14], [76], [83]). In the last decade, Lévy processes have been attracting much attention due to their applications in financial mathematics (see [29], [81; Ch. III]). Let us describe some financial models, in which Lévy processes are used.

In the famous *Black–Scholes–Merton model*, the price evolution of a financial asset is modelled as

$$S_t = S_0 e^{\mu t + \sigma B_t}, \quad t \geq 0,$$

where B is a Brownian motion, S_0 is the initial price of the asset, $\mu \in \mathbb{R}$, and $\sigma \in \mathbb{R}_+$. In 1973, F. Black, M. Scholes [16], and R. Merton [66] employed this model to obtain a formula for the fair price of a European call option. This formula is known as the *Black–Scholes formula* or the *Black–Scholes–Merton formula* (see, for example [54; Ch. 11], [81; Ch. VIII, § 1]). In 1990, M. Scholes and R. Merton were awarded the Nobel prize for their work.

The Black–Scholes–Merton model is widely used in practice. However, there is a number of discrepancies between this model and the real data. Two most important ones are as follows:

- In the Black–Scholes–Merton model, the logarithmic price process $\ln S_t = \ln S_0 + \mu t + \sigma B_t$ has Gaussian increments, while the empirical logarithmic price increments have much heavier tails than the Gaussian distribution.
- In the Black–Scholes–Merton model, the increments of $\ln S$ over nonoverlapping time intervals are independent, while the real prices have the effect of *clustering*, which consists in the following: if the increment of prices over some interval is large in the absolute value, then one should expect that the price increments over the neighbor intervals are also large in the absolute value. In other words, there are some periods of low activity and some periods of high activity.

There are many financial models that cope with the problems described above. A popular class of models is the class of *exponential Lévy models*, in which

$$S_t = S_0 e^{L_t}, \quad t \geq 0,$$

where L is a Lévy process (note that the Black–Scholes–Merton model is of this form). These models are free of the first drawback described above, but still the increments over nonoverlapping intervals are independent. Recently P. Carr, H. Geman, D. Madan, and M. Yor [20] proposed the *time-changed exponential Lévy model*, in which

$$S_t = S_0 e^{L_{\tau_t}}, \quad t \geq 0.$$

Here L is a Lévy process and τ is an increasing process that is independent of L (if τ has smooth paths, then τ' means the "inner" or the "operational" time). These models are free of both drawbacks described above. Although they are not so simple as the Black–Scholes–Merton model, they are rather elegant and analytically tractable (see [44], [64], [79]).

Problem 1.7. *First solution.* This solution applies to the case, where X is square integrable. Denote $\xi_n = \mathsf{E}(X \mid \mathcal{F}_n)$, $\eta_n = \mathsf{E}(X \mid \mathcal{G}_n)$, $\zeta_n = \mathsf{E}(X \mid \mathcal{H}_n)$. It follows from the Jensen inequality $\xi_n^2 \leq \mathsf{E}(X^2 \mid \mathcal{F}_n)$ that the family $(\xi_n^2)_{n \in \mathbb{N}}$ is uniformly integrable. Hence, $\xi_n \xrightarrow{L^2} Y$. Similarly, $\zeta_n \xrightarrow{L^2} Y$. As η_n is the projection of ζ_n on $L^2(\Omega, \mathcal{G}_n, \mathsf{P})$, the vectors $\zeta_n - \eta_n$ and $\eta_n - \xi_n$ are orthogonal in L^2. As $\zeta_n - \xi_n \xrightarrow{L^2} 0$, we get $\eta_n - \xi_n \xrightarrow{L^2} 0$, which leads to the desired statement.

Second solution. (Proposed by P. Yarykin.) This solution applies to the case, where X is integrable. Let ξ_n, η_n, and ζ_n be the same as above. As $(\xi_n)_{n\in\mathbb{N}}$ is uniformly integrable, $\xi_n \xrightarrow{L^1} Y$. Similarly, $\zeta_n \xrightarrow{L^1} Y$. It follows from the equality $\eta_n - \xi_n = \mathsf{E}(\zeta_n - \xi_n \mid \mathcal{G}_n)$ that $\mathsf{E}|\eta_n - \xi_n| \le \mathsf{E}|\zeta_n - \xi_n|$. Hence, $\eta_n - \xi_n \xrightarrow{L^1} 0$, which leads to the desired statement.

`Comments`. The statement of Problem 1.7 may be called the *stochastic lemma about two policemen*. (Recall that the lemma about two policemen states that if (a_n), (b_n), and (c_n) are three sequences of real numbers such that $a_n \le b_n \le c_n$ for any n and $a_n \to \alpha$, $c_n \to \alpha$, then also $b_n \to \alpha$.) The origin of this problem is as follows.

D. Hoover [53] introduced the notion of weak convergence of σ-fields. According to this definition, a sequence $(\mathcal{G}_n)$ of sub-σ-fields of $\mathcal{F}$ (we have a fixed probability space $(\Omega, \mathcal{F}, \mathsf{P})$) converges *weakly* to a sub-σ-field $\mathcal{G}$ of $\mathcal{F}$ if $\mathsf{E}(X \mid \mathcal{G}_n) \xrightarrow{\mathsf{P}} \mathsf{E}(X \mid \mathcal{G})$ for any bounded random variable X. One can also introduce the notion of strong convergence of σ-fields by analogy with convergence of sets. Namely, a sequence $(\mathcal{G}_n)$ converges *strongly* to $\mathcal{G}$ if $\liminf_n \mathcal{G}_n = \limsup_n \mathcal{G}_n = \mathcal{G}$, where the *lower* and *upper limits* of σ-fields are defined as follows:

$$\liminf_{n\to\infty} \mathcal{G}_n = \bigvee_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \mathcal{G}_m, \qquad \limsup_{n\to\infty} \mathcal{G}_n = \bigcap_{n=1}^{\infty} \bigvee_{m=n}^{\infty} \mathcal{G}_m$$

($\bigvee_m \mathcal{G}_m$ denotes the smallest σ-field that contains all $\mathcal{G}_m$). The question arises: what is the relationship between the two notions of convergence?

One can show that the weak convergence does not imply the strong convergence (a problem for the reader!). The reverse is true. In order to prove this, assume that $(\mathcal{G}_n)$ converges strongly to $\mathcal{G}$ and consider $\mathcal{F}_n = \bigcap_{m=n}^{\infty} \mathcal{G}_m$, $\mathcal{H}_n = \bigvee_{m=n}^{\infty} \mathcal{G}_m$. Fix a bounded random variable X. It follows from the martingale convergence theorem for forward and backward martingales (see [72; Ch. II, Cor. 2.4]) that $\mathsf{E}(X \mid \mathcal{F}_n) \xrightarrow{\text{a.s.}} \mathsf{E}(X \mid \mathcal{G})$ and $\mathsf{E}(X \mid \mathcal{H}_n) \xrightarrow{\text{a.s.}} \mathsf{E}(X \mid \mathcal{G})$. In order to complete the proof, one should apply the result of Problem 1.7.

Problem 1.8. Let us prove a more general statement: any bounded nondegenerate random variable X is not infinitely divisible. Assume the contrary. Without loss of generality, $|X| \le 1$. Then $|X_k^n| \le 1/n$ for any $n \in \mathbb{N}$, $k \le n$. Consequently, $\mathsf{D}X_k^n \le n^{-2}$, and therefore, $\mathsf{D}X \le n^{-1}$ (D denotes the variance). But this is possible only if X is degenerate.

`Remark`. Another way to solve Problem 1.8, which employs the particular form of the density of X, is as follows. One can see by a direct calculation that the characteristic function of X has zeros. On the other hand, the characteristic function of an infinitely divisible distribution has no zeros, which follows from the Lévy-Khintchine representation.

3.2 Second Kolmogorov Students' Competition

Problem 2.1. The answer is negative. Consider the example: the first black hat contains 100 tickets, 99 being lucky, the first white hat contains only 1 ticket, 1 being lucky, the second black hat contains 1 ticket, 0 being lucky, and the second white hat contains 100 tickets, 1 being lucky.

`Comments`. The effect described above has an interpretation as a "paradox" in mathematical statistics (see [86; Ch. II]).

Problem 2.2. The answer is negative. Consider the example: $\Omega = \{\omega_1, \omega_2, \omega_3\}$, P is the uniform measure, $A = \{\omega_2\}$, $B = \{\omega_1, \omega_3\}$, $C_1 = \{\omega_1, \omega_2\}$, $C_2 = \{\omega_2, \omega_3\}$.

Problem 2.3. Let X_n, $n = 1, \dots, 14$ be the indicator of the event that in the n-th pair one member is male and the other one is female. Then

$$\mathsf{E}X_n = \mathsf{P}(X_n = 1) = \frac{2 \cdot 8 \cdot 7 \cdot 13!}{15!} = \frac{8}{15}, \quad n = 1, \dots, 14.$$

Hence, the required expectation is $\mathsf{E}(X_1 + \dots + X_{14}) = \frac{112}{15}$.

`Remark`. The "indicator method" used in the solution of this problem is often applied in the variety of similar problems. The most popular one is the *problem about the absent-minded secretary.* It is as follows. A secretary has N letters to send to different addressees and N envelopes with their addresses. He/she inserts the letters in the envelopes at random, so that $N!$ allocations are equally probable. Find the expected number of the addressees who will receive his/her own letter.

Let us also remark that another problem related to the absent-minded secretary is to find $\lim_{N\to\infty} P_N$, where P_N is the probability that at least one addressee will receive his/her own letter. (The answer is $\lim_{N\to\infty} P_N = 1 - e^{-1}$.)

Problem 2.4. a) Let us identify the circumference with the interval $(0, 2\pi]$ and let F denote the distribution function of P. For any $n \in \mathbb{N}$, we have

$$F\Big(\frac{2\pi}{n}\Big) - F(0) = \dots = F(2\pi) - F\Big(\frac{2\pi(n-1)}{n}\Big).$$

Consequently, $F(2\pi q) = q$ for any $q \in \mathbb{Q} \cap [0, 1]$. Hence, $F(x) = \frac{x}{2\pi}$, $x \in [0, 2\pi]$.

b) For any $\beta \in \mathbb{R}$, there exists a sequence $n(k)$ of natural numbers such that $\varphi_\alpha^{n(k)} \to \varphi_\beta$ pointwise. Furthermore, it is clear that P attributes no mass to any one-point set. Combining these two properties, we deduce that $\mathsf{P} \circ \varphi_\beta^{-1}$ coincides with P on any interval. Employing the above arguments, we get the desired statement.

Problem 2.5. a) The answer is negative. Consider a symmetric random variable X (i.e. $X \overset{\text{law}}{=} -X$) and $a = 1$, $b = -1$.

b) The answer is positive. Indeed, assume that $a \neq b$. Without loss of generality, $b \neq 0$. Let F be the distribution function of X. Then $F(x) = F\big(\big(\frac{b}{a}\big)^n x\big)$ for any $x \in \mathbb{R}$, $n \in \mathbb{N}$. This leads to: $F(0+) = 1$, $F(0-) = 0$, i.e. $X \overset{\text{a.s.}}{=} 0$, which is a contradiction.

Problem 2.6. The answer is negative. Consider the example: $\Omega = \{\omega_1, \omega_2, \omega_3\}$, P is the uniform measure, $\mathcal{G} = \big\{\emptyset, \{\omega_1\}, \{\omega_2, \omega_3\}, \Omega\big\}$, $X(\omega) = I(\omega = \omega_3)$. (Note that any random variable Y that is independent of $\mathcal{G}$ should be degenerate.)

Problem 2.7. Let us first solve the problem with $[0, 1]$ replaced by $[-1, 1]$. For any random variable X taking on values in $[-1, 1]$, we have $\mathsf{D}X \le \mathsf{E}X^2 \le 1$. On the other hand, if $\mathsf{P}(X = \pm 1) = 1/2$, then $\mathsf{D}X = 1$. Thus, the answer in the original problem is $1/4$.

Problem 2.8. a) The answer is negative. Consider the example: $\Omega = [0, 1]$, P is the Lebesgue measure, and the sequence $X_1, X_2, \dots$ is given by $Y_1^1, Y_1^2, Y_2^2, Y_1^3, \dots$, where

$$Y_k^n = 2^n I\bigg(\bigg(\frac{k-1}{2^n}, \frac{k}{2^n}\bigg]\bigg), \quad n \in \mathbb{N},\ k = 1, \dots, 2^n.$$

Then $S_{2^n-1} = 2^n - 1$.

b) The answer is positive. It is sufficient to note that $X_n \xrightarrow{L^1} X$.

`Remark`. In a), one can construct a counterexample with independent $X_1, X_2, \dots$ (see [84; § 14.18]).

Problem 2.9. Let $\xi_1, \xi_2, \dots$ be independent Gaussian random variables with mean 0 and variance 1. Then $X_n \overset{\text{law}}{=} \frac{Z_n}{\|Z_n\|}$, where $Z_n = (\xi_1, \dots, \xi_n)$. Consequently,

$$\sqrt{n} Y_n \overset{\text{law}}{=} \frac{\xi_1}{\sqrt{(\xi_1^2 + \dots + \xi_n^2)/n}}.$$

It follows from the law of large numbers that the right-hand sides of this equality converge to ξ_1 a.s. as $n \to \infty$.

`Comments`. This statement is known as *Poincaré's lemma*. It was proposed to us by M. Yor. (See [21; § 4.2] for a discussion related to this lemma.)

Problem 2.10. The answer is negative. Indeed, suppose that such a measure Q exists. By the law of large numbers, $S_n/n \xrightarrow{\mathsf{P}\text{-a.s.}} 1/3$, where $S_n = X_1 + \dots + X_n$. Hence, $S_n/n \xrightarrow{\mathsf{Q}\text{-a.s.}} 1/3$. As $|S_n/n| \le 1$, we get $\mathsf{E_Q} S_n/n \to 1/3$, which is a contradiction.

`Comments`. This problem has its origin in financial mathematics. One of the basic problems of finance is the *pricing problem*, which consists in finding the fair prices of derivative financial contracts. There exist several approaches to solving this problem (their financial description can be found, for example, in a nice introduction to finance [15]; their mathematical description can be found, for example, in a nice introduction to financial mathematics [41]). One of the basic theories is the *arbitrage pricing theory*. Let us illustrate its main results on two models.

First, we consider a one-period model with a finite number of assets. The model consists of a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, a vector $S_0 \in \mathbb{R}^d$, and a random vector $S_1 : \Omega \to \mathbb{R}^d$. From the financial point of view, S_0^i is the price of the i-th asset at the present moment, while S_1^i is the price of the i-th asset at a fixed future time 1. (As the current prices are known, S_0 is a nonrandom vector; as the future prices are unknown, S_1 is a random vector.) Consider the set of random variables

$$A = \left\{ \sum_{i=1}^{d} h^i (S_1^i - S_0^i) : h^i \in \mathbb{R} \right\}. \tag{2}$$

From the financial point of view, A is the set of incomes that can be obtained by trading at dates 0 and 1 in the model under consideration. Indeed, if h^i units of the i-th asset are bought at time 0 and sold back at time 1, then the income of this operation is $h^i(S_1^i - S_0^i)$ (the negative value of h^i corresponds to the so-called *short selling* of the asset, i.e. borrowing it at time 0, selling it immediately, then buying it at time 1 and returning it).

The model satisfies the *No Arbitrage* (*NA*) condition if there exists no $X \in A$ with the properties: $X \ge 0$ a.s. and $\mathsf{P}(X > 0) > 0$ (the existence of such X would mean a possibility to gain something with no risk). The *fundamental theorem of asset pricing* (its proof can be found in the textbooks [41; Th. 1.6], [81; Ch. V, § 2], or [82; Ch. VII, § 11]) states that

$$\text{NA} \iff \mathcal{M} \ne \emptyset,$$

where

$$\mathcal{M} = \{\mathsf{Q} \sim \mathsf{P} : \mathsf{E}_{\mathsf{Q}}|S_1| < \infty \text{ and } \mathsf{E}_{\mathsf{Q}} S_1 = S_0\}$$

is the set of *equivalent martingale measures* (these are the probability measures equivalent to P, under which the sequence (S_0, S_1) is a martingale).

Now, let F be a random variable meaning the payoff of some derivative contract, so that $F(\omega)$ is the sum the holder of the contract gets at time 1 in the elementary outcome ω. A *NA price* of F is a real number x such that there exist no $X \in A$, $h \in \mathbb{R}$ with the properties: $X + h(F - x) \geq 0$ a.s. and $\mathsf{P}(X + h(F - x) > 0) > 0$. Applying the fundamental theorem of asset pricing to the extended $d+1$-dimensional model with $\widetilde{S}_0 = (S_0^1, \dots, S_0^d, x)$, $\widetilde{S}_1 = (S_1^1, \dots, S_1^d, F)$, we conclude that the set $I_{NA}(F)$ of NA prices has the form

$$I_{NA}(F) = \{\mathsf{E}_{\mathsf{Q}} F : \mathsf{Q} \in \mathcal{M},\ \mathsf{E}_{\mathsf{Q}}|F| < \infty\}. \tag{3}$$

The analog of this theorem is also true in the dynamic setting. Consider a model that consists of a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n=0,\dots,N}, \mathsf{P})$ and an $(\mathcal{F}_n)$-adapted d-dimensional process $(S_n)_{n=0,\dots,N}$. From the financial point of view, S_n^i is the price of the i-th asset at time n. The set of incomes that can be obtained by dynamic trading is naturally defined as

$$A = \left\{ \sum_{n=1}^{N} H_n^i (S_n^i - S_{n-1}^i) : H_n^i \text{ is } \mathcal{F}_{n-1}\text{-measurable} \right\}.$$

The NA condition and the NA price interval are defined in this model in the same way as above. The fundamental theorem of asset pricing states that

$$\text{NA} \iff \mathcal{M} \neq \emptyset,$$

where

$$\mathcal{M} = \{\mathsf{Q} \sim \mathsf{P} : S \text{ is an } (\mathcal{F}_n, \mathsf{Q})\text{-martingale}\}. \tag{4}$$

This theorem was first proved by J. Harrison and S. Pliska [48] in the case of a finite Ω and by R. Dalang, A. Morton, W. Willinger [31] in the case of a general Ω. (For these reasons, it is often called the *Harrison-Pliska theorem* or the *Dalang-Morton-Willinger theorem.*) Simpler proofs given later can be found in the textbooks [41; Th. 5.17], [81; Ch. V, § 2]. Furthermore,

$$I_{NA}(F) = \{\mathsf{E}_{\mathsf{Q}} F : \mathsf{Q} \in \mathcal{M},\ \mathsf{E}_{\mathsf{Q}}|F| < \infty\} \tag{5}$$

(see [41; Th. 5.30], [81; Ch. VI, § 1c]).

However, in the model with an infinite number of assets the analog of the theorem stated above is no longer true. Indeed, consider a model with a countable number of assets whose prices are given by $S_0^i = 1$, $S_1^i = 1 + X_i$, where $X_1, X_2, \dots$ are the random variables described in Problem 2.10. It is natural to define the set of attainable incomes in this model as

$$A = \left\{ \sum_{i=1}^{N} h^i (S_1^i - S_0^i) : N \in \mathbb{N},\ h^i \in \mathbb{R} \right\}.$$

It is easy to check that this model satisfies the NA condition. However (as the solution of the problem shows), there exists no measure $\mathsf{Q} \sim \mathsf{P}$ such that $\mathsf{E}_{\mathsf{Q}} S_1^i = S_0^i$ for any i.

This example shows that in complicated models (for instance, those with an infinite number of assets) the standard NA condition is too weak to guarantee the existence of an equivalent martingale measure. The same effect arises for continuous-time models with a finite number of assets (see [34; Ex. 7.7]). It has been recognized in financial mathematics that the standard NA condition should be strengthened in order to obtain fundamental theorems of asset pricing for complicated models. This means that in the definition of arbitrage we take some closure of A instead of A. As there is no canonical way to take the closure of A, there exist various strengthenings of the NA condition: *No Free Lunch*, *No Free Lunch with Vanishing Risk*, *No Generalized Arbitrage*, etc. (see, for example, [25], [34], [47], [62]). Each of the papers mentioned above contains, in particular, a fundamental theorem of asset pricing, i.e. an equivalence between the corresponding strengthening of the NA property and the existence of an equivalent martingale measure (also the term *equivalent risk-neutral measure* is often used). Let us remark that in complex models there is no canonical definition of a risk-neutral measure, so that different authors employ different definitions.

Problem 2.11. (The problem was proposed by F. Hubalek.) Suppose that there are two persons: a walker who tries to cross the river and a boatman who tries to go down the river. Assume that the boatman cannot go beneath a bridge, so he/she can pass only through the destroyed bridges. The possible ways of the boatman lie along the dashed lines in Figure 4. It is clear that

$$p := \mathsf{P}(\text{the walker can cross the river}) = \mathsf{P}(\text{the boatman can go down the river}) =: q.$$

It is easy to see that the boatman can go down the river if and only if the walker cannot cross the river. Thus, $p + q = 1$. As a result, $p = 1/2$.

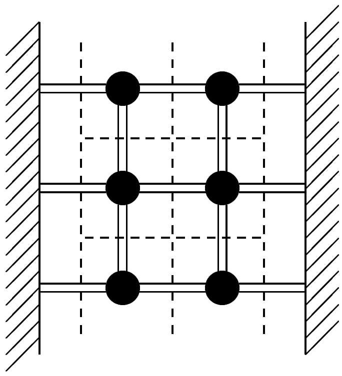


Figure 4

Comments. Problem 2.11 is closely connected with *percolation theory*. Let us briefly describe its basic object. Consider the square lattice $\mathbb{Z}^d$. Suppose that each edge of the lattice is open with probability p ($p \in (0,1)$) and is closed with probability $1-p$ (different edges are given independent designations). The central question is as follows: does there exist with a strictly positive probability an infinite cluster of open edges that contains the origin? In other words, if water is supplied at the origin and flows along open edges only, can it reach infinitely many vertices with a strictly positive probability? It is known that there exists a critical probability $p_c \in (0,1)$ such that the answer is yes if $p < p_c$ and is no if $p > p_c$. Examples of further questions are: What is the value of p_c? Is there percolation at the critical value (i.e. if $p = p_c$, does there exist with a strictly positive probability an infinite cluster of open edges that contains the origin)? Although these problems are easy to formulate, they are very hard to solve. For example, it is known that, for sufficiently high dimensions ($d \geq d_0$), there is no percolation at the critical value. But for low dimensions (already for $d = 3$), this problem remains unsolved.

However, for $d = 2$, the problems described above are completely resolved. (This was done by H. Kesten in his famous paper [61].) Namely, it is known that $p_c = 1/2$ and there is no percolation at the critical value. The most elegant way to prove the latter statement is based on considering the dual lattice to $\mathbb{Z}^d$. This is exactly the idea underlying the solution of Problem 2.11.

Let us mention that in the percolation theory there is a large variety of other interesting problems as well as other objects of study (for instance, one can consider other lattices, dynamic percolation, etc.). There are many open questions in this theory. The principal reference on percolation is [45]. The percolation theory is closely connected with other interesting topics of the modern probability, like superprocesses, random fractals, etc. Some interactions are described in [37].

3.3 Third Kolmogorov Students' Competition

Problem 3.1. The answer is positive. The proof is straightforward.

Problem 3.2. a) The answer is negative. Consider independent random variables X, Y with $\mathsf{P}(X = \pm 1) = \mathsf{P}(Y = \pm 1) = 1/2$ and take $Z = -X$.

b) The answer is positive. Indeed, an application of Fubini's theorem yields the line

$$\begin{aligned}\mathsf{P}(X + Z \leq x) &= \int_{\mathbb{R}} \mathsf{Q}_X((-\infty, x - z))\mathsf{Q}_Z(dz) \\ &\geq \int_{\mathbb{R}} \mathsf{Q}_Y((-\infty, x - z))\mathsf{Q}_Z(dz) = \mathsf{P}(Y + Z \leq x),\end{aligned}$$

where Q_X, Q_Y, and Q_Z are the distributions of X, Y, and Z, respectively.

`Comments`. The stochastic order considered in Problem 3.2 is often called the *monotone stochastic order* and is sometimes denoted as $\preccurlyeq_{\text{mon}}$. Note that $X \preccurlyeq_{\text{mon}} Y$ if and only if $\mathsf{E}f(X) \leq \mathsf{E}f(Y)$ for any increasing function $f : \mathbb{R} \to \mathbb{R}$ (provided the expectations exist). It is well known that $X \preccurlyeq_{\text{mon}} Y$ if and only if there exist random variables $\widetilde{X}$, $\widetilde{Y}$ (defined, possibly, on another probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathsf{P}})$) such that $\widetilde{X} \overset{\text{law}}{=} X$, $\widetilde{Y} \overset{\text{law}}{=} Y$, and $\widetilde{X} \leq \widetilde{Y}$ a.s. The proof can be found in [80; § 1.A], but it is a nice problem for the reader to prove this independently.

There also exist many other important stochastic orders. For example, Y dominates X in the *concave order* (notation: $X \preccurlyeq_{\text{con}} Y$) if $\mathsf{E}f(X) \leq \mathsf{E}f(Y)$ for any concave function $f : \mathbb{R} \to \mathbb{R}$ (provided the expectations exist). It is known that, for integrable X and Y, $X \preccurlyeq_{\text{con}} Y$ if and only if there exist random variables $\widetilde{X}$, $\widetilde{Y}$ (defined, possibly, on another probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathsf{P}})$) such that $\widetilde{X} \overset{\text{law}}{=} X$, $\widetilde{Y} \overset{\text{law}}{=} Y$, and $\mathsf{E}(X \mid Y) = Y$ (see [41; § 2.6]). Let us remark that the concave stochastic order is sometimes used in financial mathematics to express the view that one portfolio is less risky than another.

Presumably, the stochastic order that is most important for financial mathematics is the *monotone concave order*. A random variable Y dominates a random variable X (notation: $X \preccurlyeq_{\text{mc}} Y$) if $\mathsf{E}f(X) \leq \mathsf{E}f(Y)$ for any increasing concave function $f : \mathbb{R} \to \mathbb{R}$ (provided the expectations exist). The financial meaning is as follows: the expected utility of Y exceeds the expected utility of X for any choice of the utility function. It is known that, for integrable X and Y, $X \preccurlyeq_{\text{mc}} Y$ if and only if there exist random variables $\widetilde{X}$, $\widetilde{Y}$

(defined, possibly, on another probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathsf{P}})$) such that $\widetilde{X} \overset{\text{law}}{=} X$, $\widetilde{Y} \overset{\text{law}}{=} Y$, and $\mathsf{E}(X \mid Y) \le Y$ (see [41; § 2.6]). This stochastic order is also related to coherent risk measures. Namely, let u_λ denote the function introduced in Problem 4.6 (it corresponds to a very important coherent risk measure called Tail V@R; see Comments following the solution of Problem 4.6). Then, for integrable X and Y,

$$X \preccurlyeq_{\text{mc}} Y \iff \forall \lambda \in [0,1],\ u_\lambda(X) \le u_\lambda(Y)$$

(see [41; Rem. 4.44]). Note also that

$$X \preccurlyeq_{\text{mon}} Y \implies X \preccurlyeq_{\text{mc}} Y,$$
$$X \preccurlyeq_{\text{con}} Y \implies X \preccurlyeq_{\text{mc}} Y.$$

Along with three stochastic orders described above, there exist many other stochastic orders. One of the principal references on the subject is [80].

Let us now give a financial application of the results described above. One of the basic problems of the modern financial mathematics is to find nicer fair price intervals of derivative contracts than those provided by arbitrage pricing theory (the NA price intervals are known to be unacceptably large in many models; see, in particular, the discussion in [1; Sect. 5]). This requires new ideas. One of the techniques used is as follows. At the present time, the market of derivative securities is so large that one can consider actively traded derivatives as basic assets. Below we describe one of the models of this form.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n=0,\dots,N}, \mathsf{P})$ be a filtered probability space and $(S_n)_{n=0,\dots,N}$ be an $\mathbb{R}_+$-valued $(\mathcal{F}_n)$-adapted process describing the price of evolution of some asset. Assume that, for each $n \in \{1, \dots, N\}$ and $K \in \mathbb{R}_+$, there exists a *European call option* on this asset with maturity n and strike price K, i.e. a contract whose holder obtains the amount $(S_n - K)^+$ at time n (x^+ denotes $\max\{x, 0\}$). Let $\varphi_n(K)$ be the price at time 0 of such a contract. For this model, it is reasonable to consider a strengthening of the NA condition (see the discussion following Problem 2.10). Here we follow the approach of [25], where the *No Generalized Arbitrage* (*NGA*) condition was proposed. The fundamental theorem of asset pricing in the form proposed in [25] states that

$$\text{NGA} \iff \mathcal{M} \ne \emptyset,$$

where

$$\mathcal{M} = \{\mathsf{Q} \sim \mathsf{P} : S \text{ is an } (\mathcal{F}_n, \mathsf{Q})\text{-martingale such that } \operatorname{Law}_{\mathsf{Q}} S_n = \varphi_n'' \text{ for any } n\} \qquad (6)$$

is the set of equivalent martingale measures with given marginal distributions. Here $\operatorname{Law}_{\mathsf{Q}} S_n$ denotes the distribution of S_n under Q and φ_n'' is the second derivative of φ_n taken in the sense of distributions, i.e. it is a measure on $\mathbb{R}_+$ defined by $\varphi_n''((a,b]) = \varphi_n'(b) - \varphi_n'(a)$ (under the NGA condition, φ_n'' is a probability measure; see [25] for details). Furthermore, the set of (appropriately defined) *NGA prices* of a derivative contract F has the form

$$I_{NGA}(F) = \{\mathsf{E}_{\mathsf{Q}} F : \mathsf{Q} \in \mathcal{M},\ \mathsf{E}_{\mathsf{Q}}|F| < \infty\}.$$

As the set $\mathcal{M}$ given by (6) is smaller than the set $\mathcal{M}$ given by (4), this interval is smaller as compared to the corresponding NA price interval provided by (5).

These considerations justify the following problem. Let $\mu_0, \dots, \mu_N$ be a sequence of probability measures on $\mathbb{R}$ with $\int_{\mathbb{R}} |x| \mu_n(dx) < \infty$. Under which conditions does there exist a martingale $M_0, \dots, M_N$ on some filtered probability space such that Law $M_n = \mu_n$ for each n? Applying the above described results related to the concave stochastic order, one can deduce that the necessary and sufficient condition for the existence of such a martingale is $\mu_N \preccurlyeq_{\text{con}} \cdots \preccurlyeq_{\text{con}} \mu_0$. The same result is also true in the continuous-time setting (see [60]). Some explicit constructions of martingale measures with given marginal distributions can be found in [65].

Problem 3.3. (The problem was proposed by I. Kurkova.)

First solution. Let E_n be the expected number of the disturbed passengers if the carriage contains n seats. Then $E_0 = 0$ and

$$E_n = \mathsf{P}(\text{the } n\text{-th passenger is not on his seat})(1 + E_{n-1}) = \frac{n-1}{n}(1 + E_{n-1}).$$

This yields $E_n = \frac{n-1}{2}$, so that the answer is $\frac{99}{2}$.

Second solution. The seating described in the problem defines a permutation on $\{1, \dots, 100\}$ ($i \to j$ if the j-th passenger occupies the seat of the i-th one). Let ξ be the length of the cycle of the permutation that contains the 100-th passenger. Clearly,

$$\mathsf{P}(\xi = k) = C_{99}^{k-1} \frac{(k-1)!(99-k)!}{100!} = \frac{1}{100}, \quad k = 1, \dots, 100.$$

Thus, the required expected number E is

$$E = \sum_{k=1}^{100} \frac{k-1}{100} = \frac{99}{2}.$$

Problem 3.4. The answer is negative. Indeed, suppose that such probabilities (p_{ij}) exist (p_{ij} denotes the probability of occurrence of the j-th face for the i-th dice). Let Σ denote the sum of occurred numbers. Then $\mathsf{P}(\Sigma = 2) = p_{11}p_{21} = 1/11$, $\mathsf{P}(\Sigma = 12) = p_{16}p_{26} = 1/11$. Hence,

$$\mathsf{P}(\Sigma = 7) \geq p_{16}p_{21} + p_{11}p_{26} \geq 2\sqrt{p_{16}p_{21}p_{11}p_{26}} = \frac{2}{11}.$$

Problem 3.5. The answer is positive. Let Q_X and Q_Y denote the distributions of X and Y, respectively. Then, by Fubini's theorem,

$$\mathsf{E}|X + Y| = \int_{\mathbb{R}} \int_{\mathbb{R}} |x + y| \mathsf{Q}_X(dx) \mathsf{Q}_Y(dy) = \int_{\mathbb{R}} \mathsf{E}|X + y| \mathsf{Q}_Y(dy).$$

As $\mathsf{E}|X + Y| < \infty$, we get $\mathsf{E}|X + y| < \infty$ for Q_Y-a.e. y. This yields $\mathsf{E}|X| < \infty$.

Problem 3.6. (The problem was proposed by S. Dilman.) By the Kolmogorov 0-1 law, there exists $a \geq 0$ such that $\limsup_n \frac{X_n}{n} \overset{\text{a.s.}}{=} a$. By the Borel-Cantelli lemma,

$$\sum_{n=1}^{\infty} \mathsf{P}(X_n > (a+1)n) < \infty.$$

Clearly,

$$\mathsf{E}\frac{X_1^+}{a+1} \leq 1 + \sum_{n=1}^{\infty} \mathsf{P}\Big(\frac{X_1}{a+1} > n\Big) = 1 + \sum_{n=1}^{\infty} \mathsf{P}(X_1 > (a+1)n) < \infty.$$

By the strong law of large numbers, for $\widetilde{S}_n = X_1^+ + \cdots + X_n^+$, we have $\lim_n \frac{\widetilde{S}_n}{n} = \mathsf{E}X_1^+$. As $\widetilde{S}_n \geq S_n$, we get the desired statement.

Problem 3.7. Let ξ be a random point that has a uniform distribution on the circumference and consider random points $X_1 = \xi$, $X_2 = \xi + \frac{\pi}{2}$, $X_3 = \xi + \pi$, $X_4 = \xi + \frac{3\pi}{2}$. Let $Y_i = I(X_i$ belongs to the red region$)$. Then

$$\mathsf{E}(Y_1 + \cdots + Y_4) = \sum_{i=1}^{4} \mathsf{P}(X_i \text{ belongs to the red region}) = \frac{8}{3}.$$

As $Y_1 + \cdots + Y_4$ takes on integer values, there exists ω, for which $Y_1(\omega) + \cdots + Y_4(\omega) = 3$. This is the desired statement.

`Comments`. The method of solution of Problem 3.7 is based on introducing a probabilistic object (we introduced a random variable ξ). In this problem, the probability is in fact given (one can consider ξ to be the identical map from the circumference to itself and take $\mathsf{P} = \mu$). But there are also some problems whose solution requires an introduction of a new probability, which is not present in the original formulation.

Let us give an example of such a problem. (This is one of the most exciting problems on basic probability). Suppose that two players A and B play the following game. The player A chooses a probability distribution μ on the real line with no atoms and draws two independent random numbers X and Y, both of which have distribution μ. Then A shows X to the player B. The player B by looking at this number should decide whether he/she takes it or rejects it. If B takes X, then A gets the number Y; if B rejects X, then B gets Y, and A gets X. Thus, each of the players gets a number. Then they compare the numbers, and the one who has a bigger number is declared the winner (as μ has no atoms, X and Y are different a.s.). The question is: does B have a strategy such that whatever μ is chosen by A, the probability that B wins is strictly greater than 1/2?

Surprisingly enough, the answer is positive. The solution requires as a first stage a thorough analysis of what should be called the class of possible strategies. If a strategy is understood as a function $F : \mathbb{R} \to \{0, 1\}$ (so that B calculates $F(X)$ and takes X if and only if $F(X) = 1$), then a strategy with the desired properties does not exist. Indeed, if μ is concentrated on the set $\{F = 1\}$, then B would always take X, so that the probability of his/her winning is exactly 1/2. But B can also employ randomized strategies. A *randomized strategy* is a function $F : \mathbb{R} \to [0, 1]$; B should calculate $F(X)$ and then take X with probability $F(X)$ and reject X with probability $1 - F(X)$. If F is strictly increasing on the whole real line, then the corresponding strategy provides a positive answer to the problem posed above. The proof that B wins with a probability strictly greater than 1/2 can be based on the comparison of some integrals. But there also exists a simple proof based on the following observation. Suppose that B draws a random variable Z whose distribution function is F and then applies the following strategy: if $X > Z$, he/she takes X; otherwise, he/she rejects X. Clearly, B would take X exactly with probability $F(X)$, so that this is a realization of the strategy described above. Using this representation of the strategy, one can easily show that B wins with probability strictly greater than 1/2 (a problem for the reader!).

The method of introducing a new probability is used rather often. Let us mention one of its most surprising applications. A general opinion is that random noise is an unpleasant feature of any device and it should be filtered out. However, in some physical systems random noise has been found to play a positive role: the presence of noise can amplify the signal. The corresponding effect is known as *stochastic resonance*. This phenomenon has well been known to physicists for 25 years (see the surveys [7], [43]) and has recently attracted the attention of probabilists (see, [13], [18], [42], [50], [51], [52], [55], [56], [70]). This effect and related subjects are described in the monograph [6]. Let us briefly describe this interesting effect.

Consider a particle moving in a double-well potential U (see Figure 5) in the presence of friction, i.e. a particle whose motion is governed by the differential equation

$$\frac{d^2}{dt}X(t) = -\gamma\frac{d}{dt}X(t) - U'(X(t))$$

Here γ is the friction coefficient and U is the potential (an example is provided by $U(x) = x^4 - x^2$). We assume that γ is large, so that $\frac{d^2}{dt}X(t)$ can be eliminated, and (after replacing U by U/γ) the equation gets the form

$$\frac{d}{dt}X(t) = -U'(X(t)).$$

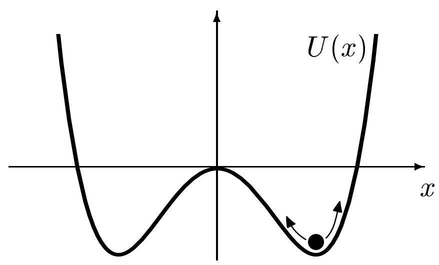


Figure 5

Now, suppose that the particle's motion is disturbed by a periodic signal, i.e. the equation gets the form

$$\frac{d}{dt}X(t) = -U'(X(t)) + f(t), \quad t \geq 0,$$

where f is a periodic function (for example, $f(t) = a\sin(\omega t + \varphi)$). This can be interpreted as a periodic perturbation of the potential, i.e. we can write $\frac{d}{dt}X(t) = -\frac{\partial}{\partial x}U(x,t)$, where $U(x,t) = U(x) + xf(t)$ (see Figure 6).

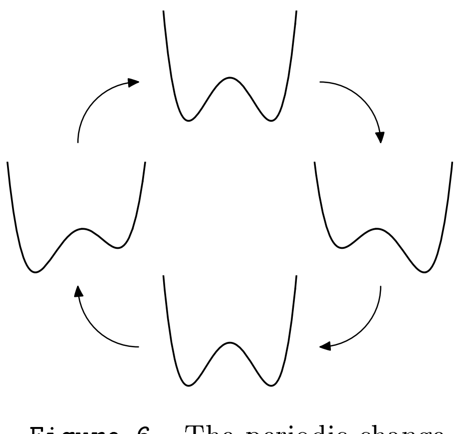


Figure 6. The periodic change of the potential $U(\cdot, t)$

Let us consider the following problem: how to filter the signal by observing the particle's motion? It is clear that if f is small as compared to U, then the particle would stay in one of two local minima, so that one cannot filter the signal by observing the particle. But let us now imagine that there is random noise disturbing the particle, i.e. its motion is governed by the stochastic differential equation

$$dX_t = [-U'(x) + f(t)]dt + \sigma dB_t, \quad t \geq 0,$$

where B is a Brownian motion and $\sigma > 0$ is the noise level (for basic facts on stochastic differential equations, one may consult [28; Ch. 1], [69], [72; Ch. XI]). Then the system has one of 3 types of behavior. If σ is very small, then the particle jumps very rarely between the wells. This is known as the *trivial behavior* (see the upper graph in Figure 7). If σ is large enough, then the particle jumps very often between the wells. This is known as the *chaotic behavior* (see the bottom graph in Figure 7). But there also exists some intermediate σ, for which the jumps of the particle between the wells become coherent with the signal. This is known as the *stochastic resonance* (see the middle graph in Figure 7). Thus, if we have chosen an optimal value of the noise level, then the particle's motion follows the signal, so we can filter the signal by observing the particle.

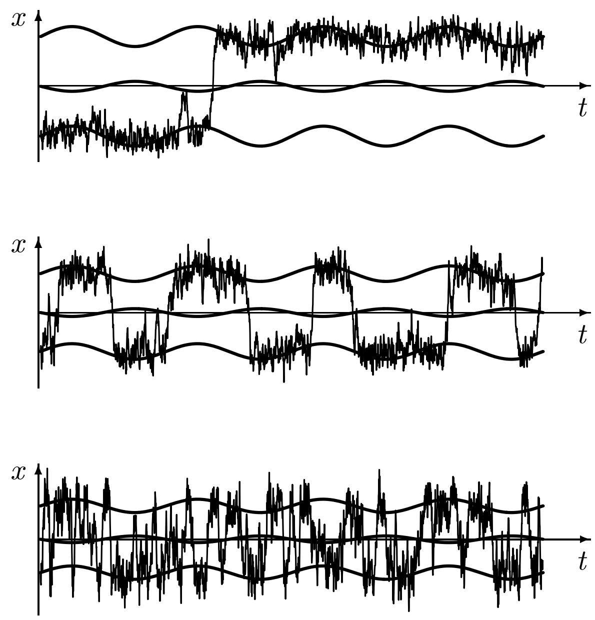


Figure 7. Simulated paths of X_t illustrating stochastic resonance. The upper, middle, and lower graphs show the trivial behavior, the stochastic resonance, and the chaotic behavior, respectively. The upper and lower sinusoids in each graph show the positions of two local minima of the potential $U(\cdot, t)$ (cf. Figure 6). The middle sinusoid on each graph shows the position of the local maximum of the potential $U(\cdot, t)$.

Let us now describe a problem, which is purely nonprobabilistic in formulation, but which can be solved only with the help of randomization. This is a *two-person zero-sum game*.

There are two players. The first one has an amount S_1 of strategies and the second one has an amount S_2 of strategies (the sets S_1 and S_2 might be finite or infinite). The game is determined by the *payoff function* $A : S_1 \times S_2 \to \mathbb{R}$. The number $A(i, j)$ means the amount the first player should pay to the second player if the first one chooses the strategy i and the second one chooses the strategy j. The problem is to find the optimal strategies of each player.

A basic notion of the corresponding theory is the notion of a saddle point. A pair $(i_0, j_0) \in S_1 \times S_2$ is called a *saddle point* if $A(i_0, j) \leq A(i_0, j_0)$ for any $j \in S_2$ and $A(i, j_0) \geq A(i_0, j_0)$ for any $i \in S_1$. It is easy to understand that if such a point exists, then it is optimal for the first player to employ the strategy i_0 and it is optimal for the second one to employ the strategy j_0. Let us also remark that a saddle point corresponds to the *minimax* and the *maximin* strategies, i.e.

$$i_0 = \operatorname*{argmin}_{i \in S_1} \sup_{j \in S_2} A(i, j),$$
$$j_0 = \operatorname*{argmax}_{j \in S_2} \inf_{i \in S_1} A(i, j).$$

However, a saddle point need not exist (it is sufficient to consider the example $S_1 = S_2 = \{1, 2\}$, $A(1,1) = A(2,2) = 1$, $A(1,2) = A(2,1) = 0$). But note that one can extend the class of strategies by considering randomized ones. A *randomized strategy* of the first player is a probability measure P_1 on S_1 (i.e. the first player chooses his strategy at random according to the measure P_1). Similarly, a randomized strategy of the second player is a probability measure P_2 on S_2. The payoff function is extended to these strategies in a straightforward way:

$$A(\mathsf{P}_1, \mathsf{P}_2) = \int_{S_1} \int_{S_2} A(x, y) \mathsf{P}_1(dx) \mathsf{P}_2(dy).$$

The basic theorem proved by J. von Neumann [67] in 1928 states that if both S_1 and S_2 are finite, then there exists a saddle point in the class of randomized strategies. This means that the game can be solved in the class of randomized strategies.

The paper [67] was the starting point of *game theory*. An essential part of the theory was created by J. von Neumann. Nowadays, game theory is one of the main instruments of analysis in various applied disciplines. It is very important for economics (see the monograph [68]). A way to apply game theory to option pricing has recently been proposed (see [89]). There also exists a game-theoretic approach to statistics, which enables one to embed most statistical problems in the framework of a two-person zero-sum game described above (see [17; Ch. 6]). Let us briefly describe how this is done.

Consider a statistical space $(\Omega, \mathcal{F}, (\mathsf{P}_\theta)_{\theta \in \Theta})$ and a random element $X : (\Omega, \mathcal{F}) \to (E, \mathcal{A})$. Suppose that we are trying to estimate $\varphi(\theta)$ by observing X, where $\varphi : \Theta \to \widetilde{\Theta}$. A *statistical game* is a game of two players: the *statistician* and the *nature*. The statistician's amount of strategies is some collection $\mathfrak{F}$ of functions $f : E \to \widetilde{\Theta}$. The nature's amount of strategies is Θ. The payoff function (in statistical games it is called the *risk function*) is defined as

$$A(f, \theta) = \mathsf{E}_{\mathsf{P}_\theta} w(f(X), \varphi(\theta)), \quad f \in \mathfrak{F}, \, \theta \in \Theta,$$

where $w : \widetilde{\Theta} \times \widetilde{\Theta} \to \mathbb{R}$ is the *loss function* (for example, if $\widetilde{\Theta} = \mathbb{R}$, then a very popular choice is $w(x, y) = (x - y)^2$). Now, the problem of finding an optimal statistical estimate of $\varphi(\theta)$ is the problem of finding an optimal statistician's strategy in the game described above. There are three typical ways to solve a statistical game.

First, one can assume that instead of having the amount Θ of possible strategies, the nature has only one, but randomized, strategy. Namely, it is assumed that there exists a probability measure μ on Θ such that the nature chooses θ at random according to this distribution. This is known as the *Bayesian approach*. Then the risk function becomes a function of one argument:

$$B(f) = \int_\Theta A(F,\theta)\mu(d\theta) = \int_\Theta \mathsf{E}_{\mathsf{P}_\theta} w(f(X),\varphi(\theta))\mu(d\theta), \quad f \in \mathfrak{F}.$$

So, in order to find an optimal strategy one should minimize $B(f)$ over $\mathfrak{F}$.

Second, if $\mathfrak{F}$ is small enough, then it might happen that there exists a uniformly optimal strategy, i.e. there exists $f_0 \in \mathfrak{F}$ such that $A(f_0,\theta) \le A(f,\theta)$ for any $f \in \mathfrak{F}$, $\theta \in \Theta$. Such a situation occurs indeed in some natural cases (recall the Cramer-Rao inequality).

If we have neither the first, nor the second situation, we should employ the technique of game theory, i.e. one should look for

$$f_0 = \operatorname*{argmin}_{f\in\mathfrak{F}} \sup_{\theta\in\Theta} A(f,\theta)$$

(possibly, $\mathfrak{F}$ should be extended to include randomized strategies). Sufficient conditions for the existence of a saddle point in statistical games as well as more information on the subject can be found in [17; Ch. 6].

Problem 3.8. Let b_n be a median of Y_n (i.e. b_n is a number such that $\mathsf{P}(Y_n \ge b_n) \ge \frac{1}{2}$ and $\mathsf{P}(Y_n \le b_n) \ge \frac{1}{2}$). Set $a_n = -b_n$. It follows from the inequalities

$$\mathsf{P}(X_n + Y_n \ge \delta) \ge \mathsf{P}(X_n \ge a_n + \delta)\mathsf{P}(Y_n \ge b_n) \ge \frac{1}{2}\mathsf{P}(X_n \ge a_n + \delta),$$
$$\mathsf{P}(X_n + Y_n \le -\delta) \ge \mathsf{P}(X_n \le a_n - \delta)\mathsf{P}(Y_n \le b_n) \ge \frac{1}{2}\mathsf{P}(X_n \le a_n - \delta)$$

that $X_n - a_n \xrightarrow{\mathsf{P}} 0$.

Problem 3.9. The answer is negative. Consider the example: $\Omega = [0,1]$, P is the Lebesgue measure, and the sequences $X_1, X_2, \dots$ and $\mathcal{G}_1, \mathcal{G}_2, \dots$ are given by $Y_1^1, Y_1^2, Y_2^2, Y_1^3, \dots$ and $\mathcal{H}_1^1, \mathcal{H}_1^2, \mathcal{H}_2^2, \mathcal{H}_1^3, \dots$, where $Y_k^n(\omega) = I(\omega \in [0, \frac{1}{n})$ and $\mathcal{H}_k^n = \{\emptyset, A_k^n, \Omega \setminus A_k^n, \Omega\}$ with

$$A_k^n = \left(0, \frac{1}{n}\right] \cup \left(\frac{k-1}{n}, \frac{k}{n}\right], \quad n \in \mathbb{N},\ k = 1, \dots, n.$$

Then $\mathsf{E}(Y_k^n \mid \mathcal{H}_k^n) = \frac{1}{2} I_{A_k^n}$ $n \in \mathbb{N},\ k = 2, \dots, n$.

Problem 3.10. Let $\mathfrak{G}$ be the set of all closed arcs of the angle $2\pi/3$. By the compactness argument, there exists $\Gamma_0 \in \operatorname{argmax}_{\Gamma\in\mathfrak{G}} \mu(\Gamma)$. Denote $\mu(\Gamma_0)$ by p and set $q = 1 - p$. Then, by Fubini's theorem,

$$\mathsf{P}\Big(\alpha \le \frac{2\pi}{3}\Big) \ge \mathsf{P}(X \in \Gamma_0,\, Y \in \Gamma_0) + \mathsf{P}\Big(X \in \Gamma_0^c,\, \alpha \le \frac{2\pi}{3}\Big) = p^2 + \int_{\Gamma_0^c} (1 - \mu(\widetilde{\Gamma}_x))\mu(dx)$$
$$\ge p^2 + \int_{\Gamma_0^c} q\mu(dx) = p^2 + q^2 \ge \frac{1}{2},$$

where Γ_0^c is the complement to Γ_0 and $\widetilde{\Gamma}_x$ is the arc that consists of points y such that the angle between x and y exceeds $2\pi/3$ (clearly, $\mu(\widetilde{\Gamma}_x) \le p$).

`Comments`. This is presumably the hardest problem of the four competitions. It was proposed by M. Jacobsen. The solution given above was proposed independently by Yu. Bakhtin and by F. Delbaen.

3.4 Fourth Kolmogorov Students' Competition

Problem 4.1. Clearly,

$$\mathsf{P}(\text{A wins}) = P_A + (1-P_A)(1-P_B)P_A + \dots = \frac{P_A}{P_A + P_B - P_A P_B}.$$

Problem 4.2. It is sufficient to note that $Y_1, Y_3, Y_5, \dots$ are independent and also $Y_2, Y_4, Y_6, \dots$ are independent.

Problem 4.3. There are many possible solutions. We give only the most straightforward one.

Any set $A \in \sigma(\sin X)$ has the form $A = \{\sin X \in B\}$, where B is a Borel subset of $[0,1]$. Clearly, $\{\sin X \in B\} = \{X \in C\}$, where C is a Borel subset of $[0,\pi]$ that is symmetric around $\pi/2$. Thus,

$$\mathsf{E}_\mathsf{P} I_A X = \frac{1}{\pi}\int_C x\,dx = \frac{1}{\pi}\int_C \frac{\pi}{2}\,dx = \mathsf{E}_\mathsf{P} I_A \frac{\pi}{2}.$$

As a result, $\mathsf{E}(X \mid \sin X) = \pi/2$.

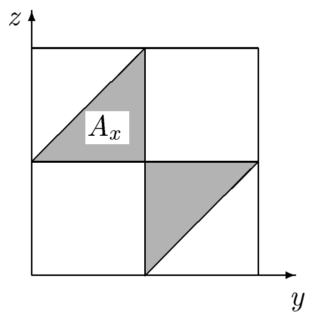
`Comments`. This problem has a different style compared to other problems. It tests not the ability to solve tricky problems, but rather the acquaintance with basic probability objects. The unexpectedly low solvability coefficient reflects the fact that students are not used to deal with conditional expectations.

Problem 4.4. (The problem was proposed by A. Mishchenko.)

First solution. By Fubini's theorem,

$$\begin{aligned}\mathsf{P}(XYZ \text{ is acute}) &= \int_C\int_C\int_C I((x,y,z)\in A)\mu(dx)\mu(dy)\mu(dz)\\ &= \int_C \mathsf{P}((Y,Z)\in A_x)\mu(dx),\end{aligned}$$

where C denotes the circumference, A is the set of points $(x,y,z) \in C^3$ corresponding to acute triangles, and A_x is the set of points $(y,z)\in C^2$ such that xyz is acute. In order to find $\mathsf{P}((Y,Z)\in A_x)$, we identify the circumference with $[0,1)$ in such a way that x corresponds to 0 and draw A_x (see Figure 8). Clearly, $\mathsf{P}((Y,Z)\in A_x) = 1/4$, so that the answer is $1/4$.



`Figure 8`

Second solution. By Fubini's theorem,

$$\begin{aligned}\mathsf{P}(XYZ \text{ is acute}) &= \int_C \int_C \int_C I((x,y,z) \in A)\mu(dx)\mu(dy)\mu(dz) \\ &= \int_C \int_C \mathsf{P}(Z \in A_{xy})\mu(dx)\mu(dy),\end{aligned}$$

where A_{xy} is the set of points $z \in C$ such that the triangle xyz is acute. It is seen from Figure 9 that $\mathsf{P}(Z \in A_{xy})$ equals the angle between x and y divided by 2π. Simple integration yields the answer.

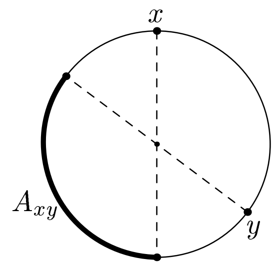


Figure 9

Third solution. Let $\widetilde{X}$, $\widetilde{Y}$, and $\widetilde{Z}$ denote the points symmetric to X, Y, and Z, respectively. If XYZ is acute, then each of triangles $\widetilde{X}YZ$, $X\widetilde{Y}Z$, $XY\widetilde{Z}$ is obtuse-angled (see Figure 10). Thus, we have established a map from points $T \in \mathfrak{A}$ to triples of points in $\mathfrak{T} \setminus \mathfrak{A}$, where $\mathfrak{T}$ denotes the set of all triangles whose vertices belong to the circumference and $\mathfrak{A}$ is the subset of acute triangles. It should be checked that different acute triangles yield different obtuse-angled triangles, any obtuse-angled triangle belongs to some triple of this form, and the map thus constructed preserves the measure μ^3. As a result, $\mu^3(\mathfrak{A}) = 1/4$.

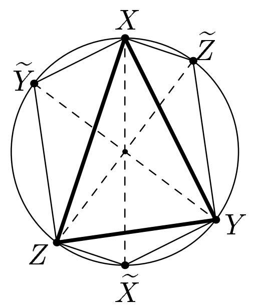


Figure 10

Problem 4.5. a) We have

$$|X + \xi Y| = |\xi||X + \xi Y| = |\xi X + Y|,$$

so that these random variables coincide a.s.

b) There exist various solutions. We give only the most standard one.

For any Borel subset A of the real line, we have, by Fubini's theorem,

$$\mathsf{P}(|X + \xi|Y + \eta Z|| \in A) = \int_{\mathbb{R}^3} \mathsf{P}(|x + \xi|y + \eta z|| \in A)\mathsf{Q}_X(dx)\mathsf{Q}_Y(dy)\mathsf{Q}_Z(dz),$$

where Q_X, Q_Y, and Q_Z denote the distributions of X, Y, and Z, respectively. Now, it is easy to see that the four-element set $\{|x \pm |y \pm z||\}$ coincides with the four-element set $\{||x \pm y| \pm z|\}$, so that $\mathsf{P}(|x + \xi|y + \eta z|| \in A) = \mathsf{P}(||x + \xi y| + \eta z| \in A)$, and the proof is easily completed.

`Comments`. The origin of the problem is as follows. V.M. Zolotarev considered in [90; § 1.4] the following operation: with a pair of independent random variables X, Y we associate a random variable $|X + \xi Y|$, where ξ is independent of (X, Y) and $\mathsf{P}(\xi = \pm 1) = 1/2$ (in fact, this is an operation on distributions). The statement that this operation is commutative and associative is exactly Problem 4.5. This problem was proposed by A.V. Lebedev.

Problem 4.6. a) *Constructive solution.* Let $q_\lambda(X)$ denote the λ-quantile of X, i.e. $q_\lambda(X) = \inf\{x : \mathsf{P}(X \le x) > x\}$. Consider Z_* of the form

$$Z_* = \lambda^{-1} I(X < q_\lambda(X)) + \xi I(X = q_\lambda(X)), \tag{7}$$

where ξ is a random variable taking on values in $[0, \lambda^{-1}]$ such that $\mathsf{E} Z_* = 1$. Let us prove that $\mathsf{E}(Z_* X) = u_\lambda(X)$. Without loss of generality, $q_\lambda(X) = 0$. Then, for any $Z \in \mathcal{D}_\lambda$,

$$\mathsf{E} Z X - \mathsf{E} Z_* X = (Z - \lambda^{-1}) X I(X < 0) + Z X I(X > 0) \ge 0.$$

`Remark. (i)` It is seen from the explicit form of Z_* that

$$u_\lambda(X) = \lambda^{-1} \int_{(-\infty, q_\lambda(X))} x \mathsf{Q}(dx) + c q_\lambda(X), \tag{8}$$

where $\mathsf{Q} = \operatorname{Law} X$ and $c = 1 - \lambda^{-1} \mathsf{Q}((-\infty, q_\lambda(X)))$.

`(ii)` The analysis of the proof given above shows that any element $Z_* \in \mathcal{D}_\lambda$ such that $\mathsf{E} Z_* X = u_\lambda(X)$ should be of the form (7).

Nonconstructive solution. By the Dunford-Pettis criterion, $\mathcal{D}_\lambda$ is relatively compact in the weak topology $\sigma(L^1, L^\infty)$ (for basic facts related to topological vector spaces, see [74]). It is easy to see that $\mathcal{D}_\lambda$ is L^1-closed. As $\mathcal{D}_\lambda$ is convex, any point of $L^1 \setminus \mathcal{D}_\lambda$ can be separated from $\mathcal{D}_\lambda$ by a linear continuous functional. As any such functional is given by an element of L^∞, $\mathcal{D}_\lambda$ is $\sigma(L^1, L^\infty)$-closed, and hence, it is $\sigma(L^1, L^\infty)$-compact. As the function $\mathcal{D}_\lambda \ni Z \mapsto \mathsf{E} Z X$ is $\sigma(L^1, L^\infty)$-continuous, it attains its minimum.

b) The answer is negative. Consider $\lambda = 1/4$ and let X, Y be independent random variables with $\mathsf{P}(X = \pm 1) = \mathsf{P}(Y = \pm 1) = 1/2$. Then it is easy to see (one may use (8)) that $u_\lambda(X) = u_\lambda(Y) = -1$ and $u_\lambda(X + Y) = -2$.

c) Clearly, we can assume from the outset that X, Y are positive and $P_X := \mathsf{P}(X = 0) > 0$, $P_Y = \mathsf{P}(Y = 0) > 0$. Without loss of generality, $P_X \le P_Y$. Take $\lambda_0 \in (P_X P_Y, P_X)$. As $\lambda_0 \le P_X$ and $\lambda_0 \le P_Y$, it is easy to see from (8) that $u_{\lambda_0}(X) = u_{\lambda_0}(Y) = 0$. As $X + Y$ is positive and $\mathsf{P}(X + Y = 0) = P_X P_Y < \lambda_0$, $u_{\lambda_0}(X + Y) > 0 = u_{\lambda_0}(X) + u_{\lambda_0}(Y)$. It is obvious that $u_\lambda(X + Y) \ge u_\lambda(X) + u_\lambda(Y)$ for any λ and the functions $u_\lambda(X)$, $u_\lambda(Y)$, and $u_\lambda(X + Y)$ are continuous in λ. This leads to the desired result.

d) As X and Y are nondegenerate, there exist two different essential values $x_1 < x_2$ of X and two different essential values $y_1 < y_2$ of Y (recall that x is called an essential value of X if $\mathsf{P}(X \in (x - \varepsilon, x + \varepsilon)) > 0$ for any $\varepsilon > 0$). Furthermore, we can choose x_1, x_2, y_1, y_2 in such a way that $x_2 - x_1 \ne y_2 - y_1$ unless both X and Y take on only two values, but the latter case has already been resolved. Without loss of generality, $x_2 - x_2 < y_2 - y_1$. Choose c such that $x_2 + y_1 < c < x_1 + y_2$ and take $\lambda_0 = \mathsf{P}(X + Y \le c)$. It is seen from (8) that $u_{\lambda_0}(X + Y) = \mathsf{E}Z_0(X + Y)$, where $Z_0 = \lambda^{-1}I(X + Y \le c)$. Furthermore, the analysis of the constructive proof of a) shows that any random variable $Z_* \in \mathcal{D}_\lambda$, for which $\mathsf{E}Z_*X = u_{\lambda_0}(X)$, should a.s. coincide with a random variable of the form $\lambda^{-1}I(X < q_{\lambda_0}) + \xi I(X + q_{\lambda_0})$. As $Z_0 = \lambda^{-1}$ on $\{X + Y \in A\}$, where A is a sufficiently small neighborhood of (x_2, y_1), and $Z_0 = 0$ on $\{X + Y \in B\}$, where B is a sufficiently small neighborhood of (x_1, y_2), we get that $\mathsf{E}Z_0X > u_{\lambda_0}(X)$. As $\mathsf{E}Z_0Y \ge u_{\lambda_0}(Y)$, we get $u_{\lambda_0}(X + Y) > u_{\lambda_0}(X) + u_{\lambda_0}(Y)$. Now, the proof is completed in the same way as above.

`Comments.` `(i)` Problem 4.6.d admits the following generalization: Let X, Y be bounded random variables and μ be a probability measure on $[0,1]$ such that $\mu((a,b)) > 0$ for any $0 \le a < b \le 1$. Then

$$\int_{[0,1]} u_\lambda(X + Y)\mu(d\lambda) > \int_{[0,1]} u_\lambda(X)\mu(d\lambda) + \int_{[0,1]} u_\lambda(Y)\mu(d\lambda) \tag{9}$$

if and only if X and Y are not comonotone. Recall that X and Y are *comonotone* if (any of) the following equivalent conditions are satisfied:

(i) $(X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \ge 0$ for $\mathsf{P} \times \mathsf{P}$ a.e. ω_1, ω_2;

(ii) there exist $\widetilde{X} \overset{\text{a.s.}}{=} X$ and $\widetilde{Y} \overset{\text{a.s.}}{=} Y$ such that $(\widetilde{X}(\omega_1) - \widetilde{X}(\omega_2))(\widetilde{Y}(\omega_1) - \widetilde{Y}(\omega_2)) \ge 0$ for any ω_1, ω_2;

(iii) there exists a function $f : \mathbb{R} \to \mathbb{R}^2$ such that each of its coordinates is an increasing function and $\mathsf{P}((X,Y) \in f(\mathbb{R})) = 1$.

The "only if" part of the statement above is clear from the constructive proof of a). Let us prove the "if" part. Suppose that (9) is not true. As $u_\lambda(X + Y) \ge u_\lambda(X) + u_\lambda(Y)$ for any λ, we get

$$\int_{[0,1]} u_\lambda(X + Y)\mu(d\lambda) = \int_{[0,1]} u_\lambda(X)\mu(d\lambda) + \int_{[0,1]} u_\lambda(Y)\mu(d\lambda).$$

As the integrands here are continuous in λ, we get $u_\lambda(X + Y) = u_\lambda(X) + u_\lambda(Y)$ for any $\lambda \in (0,1]$. This is possible only if for any $\lambda \in (0,1]$ there exists $Z_* \in \mathcal{D}_\lambda$ such that $\mathsf{E}Z_*X = u_\lambda(X)$, $\mathsf{E}Z_*Y = u_\lambda(Y)$. Using Remark (ii) above, we deduce that

$$\begin{aligned} &\mathsf{P}\big((X,Y) \in (-\infty, q_\lambda(X)) \times (q_\lambda(Y), \infty)\big) = 0, \quad \lambda \in (0,1],\\ &\mathsf{P}\big((X,Y) \in (q_\lambda(X), \infty) \times (-\infty, q_\lambda(Y))\big) = 0, \quad \lambda \in (0,1]. \end{aligned}$$

From this, it is easy to deduce that $\mathsf{P}((X,Y) \in f((0,1])) = 1$, where $f(\lambda) = (q_\lambda(X), q_\lambda(Y))$. Thus, X and Y are comonotone.

As a corollary, we obtain the statements of Problems 4.6.c, 4.6.d. Another corollary (it will be used below) is as follows: if the vector (X,Y) has a joint density (with respect to the Lebesgue measure), then (9) is true.

`(ii)` Problem 4.6 has its origin in the theory of *coherent risk measures*, which is a very new, popular, and important topic of the modern financial mathematics. The notion of a coherent risk measure was introduced by P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath in [8] and [9]. These are in fact two versions of the same paper: [8] is a financial version, while [9] is a mathematical version. The motivation of these papers is as follows. Currently, the most common way to measure the risk of a financial position is based on *Value at Risk* (abbreviated as *V@R*). Recall that $\mathrm{V@R}_\lambda(X) = -q_\lambda(X)$, where $\lambda \in (0,1]$

is a parameter (in practice, λ is fixed as a small number like 0.05 or 0.01). Despite its popularity, V@R has very serious drawbacks. These drawbacks are actively discussed in the financial literature and some of them are described in [8] and [9] (see also the discussion in [1; Sect. 5]). The goal of the authors of these papers was to introduce a way to measure risk that is similar to V@R, but free of these drawbacks; to put if briefly, a way to measure risk that is smarter than V@R. Let us describe the basic object introduced in these papers.

Let Ω be a finite set. A *coherent utility function* is a map $u : L^0 \to \mathbb{R}$ (here L^0 denotes the space of all functions $\Omega \to \mathbb{R}$) satisfying the following properties:

(i) (Superadditivity) $u(X+Y) \geq u(X) + u(Y)$;
(ii) (Monotonicity) if $X \leq Y$, then $u(X) \leq u(Y)$;
(iii) (Positive homogeneity) $u(\lambda X) = \lambda u(X)$ for $\lambda \in \mathbb{R}_+$;
(iv) (Translation invariance) $u(X+m) = u(X) + m$ for $m \in \mathbb{R}$.

A *coherent risk measure* is a coherent utility function taken with the minus sign. (It is more convenient to deal with coherent utility functions because it enables one to get rid of numerous minus signs.)

The basic result of [8] and [9] is the representation theorem. It states that any coherent utility function can be represented as

$$u(X) = \inf_{\mathsf{Q} \in \mathcal{D}} \mathsf{E}_{\mathsf{Q}} X, \tag{10}$$

where $\mathcal{D}$ is a convex closed set of probability measures on Ω (clearly, the right-hand side of (10) is a coherent utility function for any choice of $\mathcal{D}$). Furthermore, such a set $\mathcal{D}$ is unique.

F. Delbaen [32] extended the study of coherent risk measures to general probability spaces. This is done as follows. Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space. A *coherent utility function* is a map $u : L^\infty \to \mathbb{R}$ (L^∞ is the space of bounded random variables on $(\Omega, \mathcal{F}, \mathsf{P})$) that satisfies conditions (i)–(iv) above. It is easy to show that any such map admits representation (10) with some set $\mathcal{D}$ of positive finitely additive measures on $\mathcal{F}$ with total mass 1 that are absolutely continuous with respect to P. However, dealing with finitely additive measures is of course unpleasant. F. Delbaen introduced a continuity axiom, which he called the *Fatou property*. It is as follows

(v) if $X_n \xrightarrow{\mathsf{P}} X$ and $|X_n| \leq 1$, then $\limsup_n u(X_n) \leq u(X)$.

(It is possible to show that this is equivalent to the following property: if a sequence (X_n) decreases to X and $|X_n| \leq 1$, then $u(X_n) \to u(X)$.) It was proved in [32] that any function $u : L^\infty \to \mathbb{R}$ satisfying conditions (i)–(v) can be represented in the form (10), where $\mathcal{D}$ is an L^1-closed convex set of probability measures on $\mathcal{F}$ that are absolutely continuous with respect to P (we identify measures that are absolutely continuous with respect to P with their densities). Furthermore, such a set $\mathcal{D}$ is unique.

An important example of a coherent risk measure is *Tail V@R* (some authors also use the terms *Average V@R*, *Conditional V@R*, and *expected shortfall*). Tail V@R of order $\lambda \in [0,1]$ is the coherent risk measure $\rho_\lambda = -u_\lambda$, where u_λ is defined in Problem 4.6. (Clearly, $u_0(X) = \operatorname{essinf}_\omega X(\omega)$, where $\operatorname{essinf}_\omega X(\omega) = \sup\{x \in \mathbb{R} : X \geq x \text{ a.s.}\}$.) It is seen from representation (8) that in the case, where $\operatorname{Law} X$ has no atoms, $u_\lambda(X)$ is the conditional expectation of X given the set $\{X < q_\lambda(X)\}$. Another important example is *Weighted V@R*. Weighted V@R corresponding to a probability measure μ on $[0,1]$ is the coherent risk measure $\rho_\mu = -u_\mu$, where

$$u_\mu(X) = \int_{[0,1]} u_\lambda(X)\mu(dx).$$

After the first papers on coherent risk measures, this theory has rapidly been evolving. Let us mention just a few papers that followed. H. Föllmer and A. Schied [39], [40], [41; Ch. 4] introduced the notion of a *convex risk measure*. It is defined similarly to a coherent risk measure with condition (iii) being dropped and condition (i) replaced by

(i') $u(\alpha X+(1-\alpha)Y)\geq \alpha u(X)+(1-\alpha)u(Y)$ for any $X,Y\in L^\infty$, $\alpha\in[0,1]$.

S. Kusuoka [63] obtained a description of *law invariant* coherent utility functions (a coherent utility function u is law invariant if $u(X)$ depends only on the distribution of X). Namely, he proved that if $(\Omega,\mathcal{F},\mathsf{P})$ has no atoms, then any law invariant coherent utility function can be represented as

$$u(X)=\inf_{\mu\in\mathfrak{M}}u_\mu(X), \tag{11}$$

where $\mathfrak{M}$ is a set of probability measures on $[0,1]$ and u_μ is Weighted V@R. (It is easy to see from representation (8) that u_λ is law invariant; hence, u_μ is law invariant, so that the right-hand side of (11) is a law invariant coherent utility function for any choice of $\mathfrak{M}$.)

Nice expositions on coherent risk measures are given in [33], [41; Ch. 4], [77], [85].

All the papers mentioned above are related to *static* risk measures. One of the most promising topics of the modern research is the theory of *dynamic* risk measures; see [10], [11], [22], [23], [24], [30], [35], [58], [73], [87]. The study of dynamic risk measures is closely connected with the theory of *backward stochastic differential equations*; see [71].

While the first papers on coherent risk measures were aimed at the study of risk measures themselves, recent literature is characterized by an increasing interest to applications of coherent risk measures to the problems of finance. An important direction is the so-called *No Good Deals* (*NGD*) pricing (see [19], [26], [57]). This technique is aimed at obtaining finer price intervals of derivative contracts than those provided by arbitrage considerations (see the Comments following Problem 3.2, where another technique was discussed). We illustrate the idea of the NGD pricing by a one-period model with a finite number of assets, which has been considered in the Comments following Problem 2.10.

Let $(\Omega,\mathcal{F},\mathsf{P})$ be a probability space, $S_0\in\mathbb{R}^d$ be the vector of initial prices of several assets, and S_1 be a random vector of their terminal prices. Let A be the set of attainable incomes given by (2). First of all, we define a coherent utility function u on the space of all random variables (this is necessary because the elements of A need not be bounded). This is done as follows. Let $\mathcal{D}$ be an L^1-closed convex set of probability measures that are absolutely continuous with respect to P. For any $X\in A$ (note that X does not necessarily belong to L^∞), we define $u(X)$ by (10), where the expectation $\mathsf{E}_\mathsf{Q}X$ is understood as $\mathsf{E}_\mathsf{Q}X^+-\mathsf{E}_\mathsf{Q}X^-$ with the convention $\infty-\infty=-\infty$. Here $X^+=\max\{X,0\}$ and $X^-=\max\{-X,0\}$. We will assume that, for any i, $\sup_{\mathsf{Q}\in\mathcal{D}}|\mathsf{E}_\mathsf{Q}S_1^i|<\infty$ (in particular, this implies that S_1 is integrable with respect to any $\mathsf{Q}\in\mathcal{D}$).

The model satisfies the *NGD* condition if there exists no $X\in A$ with $u(X)>0$. The financial interpretation is as follows. A random variable X with $u(X)>0$ represents a trading opportunity with a negative risk. Such an opportunity is attractive to all the market participants, so that if it occurred, everyone would try to exploit it, which would lead to its disappearing. The fundamental theorem of asset pricing (see [26; Subsect. 2.2]) states (under minor additional assumptions) that

$$\text{NGD}\iff\mathcal{D}\cap\mathcal{M}\neq\emptyset,$$

where

$$\mathcal{M}=\{\mathsf{Q}\ll\mathsf{P}:\mathsf{E}_\mathsf{Q}|S_1|<\infty\text{ and }\mathsf{E}_\mathsf{Q}S_1=S_0\}$$

is the set of martingale measures that are absolutely continuous with respect to P (recall that $\mathsf{Q} \ll \mathsf{P}$ if $\mathsf{P}(A) = 0 \Rightarrow \mathsf{Q}(A) = 0$).

Now, let F be a random variable meaning the payoff of some derivative contract. We assume that $\sup_{\mathsf{Q}\in\mathcal{D}} |\mathsf{E}_{\mathsf{Q}} F| < \infty$. An *NGD price* of F is a real number x such that there exist no $X \in A$, $h \in \mathbb{R}$ with $u(X + h(F - x)) > 0$. Applying the fundamental theorem of asset pricing to the extended $d+1$-dimensional model with $\widetilde{S}_0 = (S_0^1, \dots, S_0^d, x)$, $\widetilde{S}_1 = (S_1^1, \dots, S_1^d, F)$, we conclude that the set $I_{NGD}(F)$ of NGD prices has the form

$$I_{NGD}(F) = \{\mathsf{E}_{\mathsf{Q}} F : \mathsf{Q} \in \mathcal{D} \cap \mathcal{M},\ \mathsf{E}_{\mathsf{Q}}|F| < \infty\},$$

which is smaller than the corresponding NA price interval provided by (3).

Along with pricing, a basic problem of the modern financial mathematics is the problem of finding the optimal structure of a portfolio. Various settings of this problem based on coherent risk measures were considered in [26] and [78]. Another basic problem is related to equilibrium; equilibrium based on coherent risk measures was studied in [12], [26], [49], and [59].

Let us describe one of the optimization problems considered in [26]. (As will be seen below, this motivates Problems 4.6.c and 4.6.d.) Consider a one-period model with a finite number of assets described above. For $X \in A$, we define its *Risk-Adjusted Return on Capital* (*RAROC*) as

$$\mathrm{RAROC}(X) = \begin{cases} +\infty & \text{if } \mathsf{E}_{\mathsf{P}} X > 0 \text{ and } u(X) \geq 0, \\ \dfrac{\mathsf{E}_{\mathsf{P}} X}{-u(X)} & \text{otherwise} \end{cases}$$

with the convention $\frac{0}{0} = 0$, $\frac{\infty}{\infty} = 0$. Let us now consider the following optimization problem:

$$\mathrm{RAROC}(\langle h, S_1 - S_0 \rangle) \xrightarrow[h\in\mathbb{R}^d]{} \max, \tag{12}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^d$.

We give a simple geometric solution of this problem. The key step is to introduce the set $C = \mathrm{cl}\{\mathsf{E}_{\mathsf{Q}} S_1 : \mathsf{Q} \in \mathcal{D}\}$, where "cl" denotes the closure. It is easy to see that C is a convex compact in $\mathbb{R}^d$. Set $E = \mathsf{E}_{\mathsf{P}} S_1$. It is natural to assume that $\mathsf{P} \in \mathcal{D}$, so that $E \in C$. Clearly, for any $h \in \mathbb{R}^d$, we have

$$\mathsf{E}_{\mathsf{P}} \langle h, S_1 - S_0 \rangle = \langle h, E - S_0 \rangle, \tag{13}$$

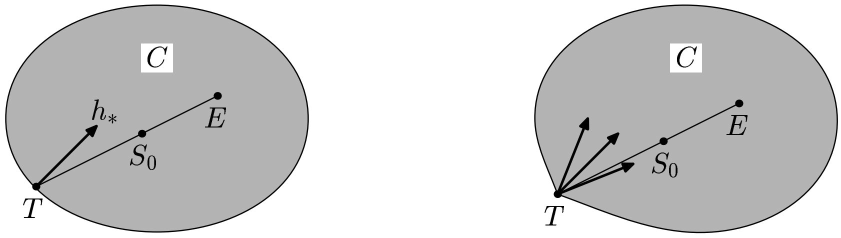
$$u(\langle h, S_1 - S_0 \rangle) = \min_{x\in C} \langle h, x - S_0 \rangle. \tag{14}$$

We will assume that $S_0 \in C^\circ \setminus \{E\}$, where C° denotes the relative interior of C. (Recall that the relative interior of C is its interior in the relative topology of the smallest affine subspace containing C.) The solution of (12) is given by the following procedure (see Figure 11). Let T denote the intersection of the ray (E, S_0) with the border of C. Then

$$\sup_{h\in\mathbb{R}^d} \mathrm{RAROC}(\langle h, S_1 - S_0 \rangle) = \frac{|E - S_0|}{|S_0 - T|}$$

and

$$\operatorname*{argmax}_{h\in\mathbb{R}^d} \mathrm{RAROC}(\langle h, S_1 - S_0 \rangle) = \{h \in \mathbb{R}^d : \forall x \in C^\circ,\ \langle h, x - T \rangle > 0\}.$$



`Figure 11.` Solution of the optimization problem. In the left graph, the solution h_* of (12) is unique up to multiplication by a positive constant. In the right graph, the border of C has a break at the point T and the optimal solution is not unique.

(If C has a nonempty interior, this is the set of inner normals to C at the point T.) These statements can be derived from (13) and (14) with the help of elementary geometric considerations (the proof is given in [26; Subsect. 3.2]).

The connection between property (9) (it might be called the *strict diversification* property) and the optimization problem discussed above is as follows. Let $u = u_\mu$, where μ is a measure on $[0,1]$ such that $\mu((a,b)) > 0$ for any $0 \le a < b \le 1$ (u_μ is Weighted V@R). Suppose that the distribution of S_1 has a density (with respect to the Lebesgue measure). Then a solution of problem (12) is unique up to multiplication by a positive constant. Indeed, suppose that there exist two solutions h_*^1 and h_*^2 that are not collinear. After multiplying h_*^1 by a positive constant, we can assume that

$$\mathsf{E}_\mathsf{P}\langle h_*^1, S_1 - S_0\rangle = \mathsf{E}_\mathsf{P}\langle h_*^2, S_1 - S_0\rangle.$$

As h_*^1 and h_*^2 solve (12),

$$u_\mu(\langle h_*^1, S_1 - S_0\rangle) = u_\mu(\langle h_*^2, S_1 - S_0\rangle).$$

Consider $h_* = \frac{h_*^1 + h_*^2}{2}$. Then

$$\mathsf{E}_\mathsf{P}\langle h_*, S_1 - S_0\rangle = \mathsf{E}_\mathsf{P}\langle h_*^i, S_1 - S_0\rangle,$$

while it follows from (9) that

$$u_\mu(\langle h_*, S_1 - S_0\rangle) > u_\mu(\langle h_*^i, S_1 - S_0\rangle).$$

Thus,

$$\mathrm{RAROC}(\langle h_*, S_1 - S_0\rangle) > \mathrm{RAROC}(\langle h_*^i, S_1 - S_0\rangle),$$

which is a contradiction.

We conclude this comment by an ideological remark. S. Kusuoka [63] proved that Tail V@R of order λ is the smallest law invariant coherent risk measure that dominates V@R of order λ (note that $\mathrm{V@R}_\lambda$ is not a coherent risk measure because it does not satisfy the superadditivity axiom (i)). This property shows that Tail V@R is one of the most important classes of coherent risk measures. However, we believe that Weighted V@R is a better class. One obvious advantage of Weighted V@R over Tail V@R is that the former risk measure employs the whole distribution of a random variable (provided that $\mu((a,b)) > 0$ for any $0 \le a < b \le 1$), while the latter one depends only on the tail of distribution to the left of its λ-quantile (see representation (8)). Another reason is that Weighted V@R possesses the strict diversification property (9) (provided that $\mu((a,b)) > 0$ for any $0 \le a < b \le 1$), which leads to the uniqueness of a solution of various optimization problems; on the other hand, as shown by Problem 4.6.b, Tail V@R does not have this property. To put it briefly, Weighted V@R is "smoother" than Tail V@R (see [27] for a more detailed discussion and for further results on Weighted V@R).

Problem 4.7. The answer is negative. Indeed, $\frac{B_m}{\sqrt{2m \ln \ln m}} \xrightarrow{\mathsf{P}} 0$, and therefore, $\frac{B_{m_n}}{\sqrt{2m_n \ln \ln m_n}} \xrightarrow{\text{a.s.}} 0$ for some subsequence m_n. Now, it is sufficient to take $t_n = m_n$.

Problem 4.8. The answer is negative. Fix $n \in \mathbb{N}$ and consider the example: $\mathsf{P}(X = e_1) = 1/2$, $\mathsf{P}(X = e_2) = \dots = \mathsf{P}(X = e_{n+1}) = 1/2n$, where $e_1, \dots, e_{n+1}$ is a basis in $\mathbb{R}^{n+1}$. Then

$$\begin{aligned}\mathsf{P}(Z \in L \mid d(L) = 1) &= \frac{\mathsf{P}(Z = e_1,\, X = Y = e_1) + \sum_{k=2}^{n+1} \mathsf{P}(Z = e_k,\, X = Y = e_k)}{\mathsf{P}(X = Y = e_1) + \sum_{k=2}^{n+1} \mathsf{P}(X = Y = e_k)} \\ &= \frac{\frac{1}{2}\frac{1}{4} + \frac{n}{(2n)^3}}{\frac{1}{4} + \frac{n}{(2n)^2}} \xrightarrow[n\to\infty]{} \frac{1}{2}.\end{aligned}$$

Similarly,

$$\mathsf{P}(Z \in L \mid d(L) = 2) \xrightarrow[n\to\infty]{} \frac{1}{3}.$$

Thus, for a sufficiently large n, we get a desired counterexample.

`Comments`. When I had proposed this problem to L.C.G. Rogers, he sent me the following E-mail 20 minutes later: "The answer is negative. The example is

$$\begin{aligned} P_1 &= 0.0084692, \\ P_2 &= 0.9489620, \\ P_3 &= 0.0170750, \\ P_4 &= 0.0058926, \\ P_5 &= 0.0196012." \end{aligned}$$

This mysterious solution is as follows.

Assume that $\mathsf{P}(X = e_1) = P_1, \dots, \mathsf{P}(X = e_n) = P_n$, where $e_1, \dots, e_n$ is a basis in $\mathbb{R}^n$. Then it is easy to compute both sides of the inequality and to write a computer program that picks at random many vectors $P_1, \dots, P_n$ with $P_i \ge 0$, $P_1 + \dots + P_n = 1$ and verifies whether the inequality in question is true. For $n = 5$, the program immediately yields a counterexample!

References

[1] Current and emerging research opportunities in probability. Workshop report, available at: `http://www.math.cornell.edu/~durrett/probrep/probrep.html`.

[2] Information on the First "Kolmogorov Students' Competition on Probability Theory". — Theory of Probability and Its Applications, **46** (2001), No. 4.

[3] Information on the Second "Kolmogorov Students' Competition on Probability Theory". — Theory of Probability and Its Applications, **48** (2003), No. 2.

[4] Information on the Third "Kolmogorov Students' Competition on Probability Theory". — Theory of Probability and Its Applications, **49** (2004), No. 3.

[5] Information on the Fourth "Kolmogorov Students' Competition on Probability Theory". — Theory of Probability and Its Applications, **50** (2005), No. 3.

[6] *V.S. Anishchenko, V. Astakhov, A.B. Neiman, T. Vadivasova, L. Schimansky-Geier.* Nonlinear Dynamics of Chaotic and Stochastic Systems. Springer, 2001.

[7] *V.S. Anishchenko, A.B. Neiman, F. Moss, L. Schimansky-Geier.* Stochastic resonance: noise-enhanced order. Phys. Usp., **42** (1999), No. 1, p. 7–36.

[8] *P. Artzner, F. Delbaen, J.-M. Eber, D. Heath.* Thinking coherently. Risk, **10** (1997), No. 11, p. 68–71.

[9] *P. Artzner, F. Delbaen, J.-M. Eber, D. Heath.* Coherent measures of risk. Mathematical Finance, **9** (1999), No. 3, p. 203–228.

[10] *P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, H. Ku.* Coherent multiperiod risk adjusted values and Bellman's principle. Preprint, available at: `http://www.math.ethz.ch/~delbaen`.

[11] *P. Barrieu, N. El Karoui.* Optimal derivatives design under dynamic risk measures. In: Yin, George (Eds.). Mathematics of finance. Proceedings of an AMS-IMS-SIAM joint summer research conference on mathematics of finance, 2003. AMS Contemporary Mathematics **351**, (2004) p. 13–25.

[12] *P. Barrieu, N. El Karoui.* Inf-convolution of risk measures and optimal risk tranfer. Finance and Stochastics, **9** (2005), p. 269–298.

[13] *N. Berglund, B. Gentz.* A sample-paths approach to noise-induced syncronization: stochastic resonance in a double-well potential. Annals of Applied Probability, **12** (2002), No. 4, p. 1419–1470.

[14] *J. Bertoin.* Lévy processes. Cambridge, 1998.

[15] *Z. Bodie, R. Merton.* Finance. Prentice Hall, 1999.

[16] *F. Black, M. Scholes.* The pricing of options and corporate liabilities. Journal of Political Economy, **81** (1973), No. 3, p. 637–659.

[17] *A.A. Borovkov.* Mathematical statistics. Gordon & Breach, 1998.

[18] *S.C. Carmona, M.I. Freidlin.* On logarithmic asymptotics of stochastic resonance frequencies. Stochastics and Dynamics, **3** (2003), No. 1, p. 55–71.

[19] *P. Carr, H. Geman, D. Madan.* Pricing and hedging in incomplete markets. Journal of Financial Economics, **62** (2001), p. 131–167.

[20] *P. Carr, H. Geman, D. Madan, M. Yor.* Stochastic volatility for Lévy processes. Mathematical Finance, **13** (2003), No. 3, p. 345–382.

[21] *L. Chaumont, M. Yor.* Exercises in probability. A guided tour from measure theory to random processes, via conditioning. Cambridge, 2003.

[22] *P. Cheridito, F. Delbaen, M. Kupper.* Coherent and convex monetary risk measures for bounded càdlàg processes. Stochastic Processes and their Applications, **112** (2004), No. 1, p. 1–22.

[23] *P. Cheridito, F. Delbaen, M. Kupper.* Coherent and convex monetary risk measures for unbounded càdlàg processes. Preprint, available at: `http://www.princeton.edu/~dito/papers`.

[24] *P. Cheridito, F. Delbaen, M. Kupper.* Dynamic monetary risk measures for bounded discrete-time processes. Article math.PR/0410453 on Mathematics ArXiv, `http://arxiv.org`.

[25] *A.S. Cherny.* General arbitrage pricing model. Preprint, available at: `http://mech.math.msu.su/~cherny`.

[26] *A.S. Cherny.* Pricing, optimality, and equilibrium based on coherent risk measures. Preprint, available at: `http://mech.math.msu.su/~cherny`.

[27] *A.S. Cherny.* Weighted V@R and its properties. Preprint, available at: `http://mech.math.msu.su/~cherny`.

[28] *A.S. Cherny, H.-J. Engelbert.* Singular stochastic differential equations. Lecture Notes in Mathematics, 1858 (2004).

[29] *R. Cont, P. Tankov.* Financial modelling with jump processes. Chapman & Hall, 2004.

[30] *J. Cvitanić, I. Karatzas.* On dynamic measures of risk. Finance and Stochastics, **3** (1999), p. 451–482.

[31] *R.C. Dalang, A. Morton, W. Willinger.* Equivalent martingale measures and no-arbitrage in stochastic securities market models. Stochastics and Stochastic Reports, **29** (1990), No. 2, p. 185–201.

[32] *F. Delbaen.* Coherent risk measures on general probability spaces. In: Advances in Finance and Stochastics. Essays in honor of Dieter Sondermann. K. Sandmann, P. Schönbucher (Eds.). Springer, 2002, p. 1–37.

[33] *F. Delbaen.* Coherent monetary utility functions. Preprint, available at `http://www.math.ethz.ch/~delbaen` under the name “Pisa lecture notes”.

[34] *F. Delbaen, W. Schachermayer.* A general version of the fundamental theorem of asset pricing. Mathematische Annalen, **300** (1994), p. 463–520.

[35] *K. Detlefsen, G. Scandolo.* Conditional and dynamic convex risk measures. Preprint, available at: `http://www.dmd.unifi.it/scandolo`.

[36] *A.Ya. Dorogovtsev, D.S. Silvestrov, A.V. Skorokhod, M.I. Yadrenko.* Probability theory — collection of problems. Translations of Mathematical Monographs. AMS, 1997.

[37] *R. Durrett.* Probability theory — an introduction to its applications. In: B. Engquist, W. Schmid. (Eds.) Mathematics unlimited: 2001 and beyond. Springer, 2001, p. 393–405.

[38] *B. Engquist, W. Schmid.* (Eds.) Mathematics unlimited: 2001 and beyond. Springer, 2001.

[39] *H. Föllmer, A. Schied.* Convex measures of risk and trading constraints. Finance and Stochastics, **6** (2002), p. 429–447.

[40] *H. Föllmer, A. Schied.* Robust preferences and convex risk measures. In: K. Sandmann, Ph. Schonbucher (Eds.). Advances in Finance and Stochastics, Essays in Honor of Dieter Sondermann. Springer, 2002, p. 39–56.

[41] *H. Föllmer, A. Schied.* Stochastic finance. An introduction in discrete time. 2nd Ed., Walter de Gruyter, 2004.

[42] *M. Freidlin.* Quasi-deterministic approximation, metastability, and stochastic resonance. Physica D, **137** (2000), p. 333-352.

[43] *L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni.* Stochastic resonance. Reviews of Modern Physics, **70** (1998), No. 1, p. 223–287.

[44] *H. Geman, D. Madan, M. Yor.* Time changes hidden in Brownian subordination. Preprint, available at: `http://www.rhsmith.umd.edu/faculty/dmadan`.

[45] *G. Grimmett.* Percolation. 2nd Ed., Springer, 1999.

[46] *G. Grimmett, D. Stirzaker.* One thousand exercises in probability. 2nd Ed., Oxford, 2001.

[47] *J.M. Harrison, D.M. Kreps.* Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory, **20** (1979), p. 381–408.

[48] *J.M. Harrison, S.R. Pliska.* Martingales and stochastic integrals in the theory of continuous trading. Stochastic Processes and Their Applications, **11** (1981), No. 3, p. 215–260.

[49] *D. Heath, H. Ku.* Pareto equilibria with coherent measures of risk. Mathematical Finance, **14** (2004), p. 163–172.

[50] *S. Herrmann, P. Imkeller.* Barrier crossings characterize stochastic resonance. Stochastics and Dynamics, **2** (2002), No. 3, p. 413–436.

[51] *S. Herrmann, P. Imkeller.* The exit problem for diffusions with time-periodic drift and stochastic resonance. Annals of Applied Probability, **15** (2005), No. 1A, p. 39–68.

[52] *S. Herrmann, P. Imkeller, I. Pavlyukevich.* Stochastic resonance: non-robust and robust tuning notions. In: K. Haman, B. Jakobiak, J. Zabczyk (Eds.). Probabilistic Problems in Atmospheric and Water Sciences. Proceedings of the workshop held in Bedlewo Mathematical Conference Center. Fizyka Atmosfery, 2003, p. 61–90.

[53] *D.N. Hoover.* Convergence in distribution and Skorokhod convergence for the general theory of processes. Probability Theory and Related Fields, **89** (1991), p. 239–259.

[54] *J.C. Hull.* Options, futures, and other derivatives. 5th Ed., Prentice Hall, 2002.

[55] *P. Imkeller, I. Pavlyukevich.* Stochastic resonance in two-state Markov chains. Archives of Mathematics, **77** (2001), No. 1, p. 107–115.

[56] *P. Imkeller, I. Pavlyukevich.* Model reduction and stochastic resonance. Stochastics and Dynamics, **2** (2002), No. 4, p. 463–506.

[57] *S. Jaschke, U. Küchler.* Coherent risk measures and good deal bounds. Finance and Stochastics, **5** (2001), p. 181–200.

[58] *A. Jobert, L.C.G. Rogers.* Pricing operators and dynamic convex risk measures. Preprint, available at: `http://www.statslab.cam.ac.uk/~chris`.

[59] *E. Jouini, W. Schachermayer, N. Touzi.* Optimal risk sharing for law invariant monetary utility functions. Preprint, available at: `http://www.fam.tuwien.ac.at/~wschach/pubs`.

[60] *H. Kellerer.* Markov-Komposition und eine Anwendung auf Martingale. Mathematische Annalen, **198** (1972), p. 99–122.

[61] *H. Kesten.* The critical probability of bond percolation on the square lattice equals $1/2$. Communications in Mathematical Physics, **74** (1980), p. 41–59.

[62] *D. Kreps.* Arbitrage and equilibrium in economies with infinitely many commodities. Journal of Mathematical Economics, **8** (1981), p. 15–35.

[63] *S. Kusuoka.* On law invariant coherent risk measures. Advances in Mathematical Economics, **3** (2001), p. 83–95.

[64] *D. Madan, E. Seneta.* The VG model for Share Market Returns. Journal of Business, **63** (1990), No. 4, p. 511-524.

[65] *D. Madan, M. Yor.* Making Markov martingales meet marginals: with explicit constructions. Bernoulli, **8** (2002), No. 4, p. 509–536.

[66] *R.C. Merton.* Theory of rational option pricing. Bell Journal of Economics and Management Science, **4** (1973), p. 141–183.

[67] *J. von Neumann.* Zur theorie der gesellschaftsspiele. Mathematische Annalen, **100** (1928), p. 295–320.

[68] *J. von Neumann, O. Morgenstern.* Theory of games and economic behavior. 2nd Ed., Princeton, 1947.

[69] *B. Øksendal.* Stochastic differential equations: an introduction with applications. 6th Ed., Springer, 2003.

[70] *I. Pavlyukevich.* Stochastic resonance. Ph.D. thesis, available at: `http://www.mathematik.hu-berlin.de/~pavljuke`.

[71] *S. Peng.* Nonlinear expectations, nonlinear evaluations, and risk measures. Lecture Notes in Mathematics, **1856** (2004), p. 143–218.

[72] *D. Revuz, M. Yor.* Continuous martingales and Brownian motion. 3rd Ed., Springer, 1999.

[73] *F. Riedel.* Dynamic coherent risk measures. Stochastic Processes and Their Applications, **112** (2004), p. 185-200.

[74] *A. Robertson, W. Robertson.* Topological vector spaces. Cambridge, 1980.

[75] *J.P. Romano, A.F. Siegel.* Counterexamples in probability and statistics. Chapman & Hall, 1986.

[76] *K.-I. Sato.* Lévy processes and infinitely divisible distributions. Cambridge, 1999.

[77] *A. Schied.* Risk measures and robust optimization problems. Lecture notes of a mini-course held at the 8th symposium on probability and stochastic processes. Preprint.

[78] *J. Sekine.* Dynamic minimization of worst conditional expectation of shortfall. Mathematical Finance, **14** (2004), p. 605–618.

[79] *A.V. Selivanov.* On martingale measures in exponential Lévy models. Theory of Probability and Its Applications, **49** (2004), No. 2.

[80] *M. Shaked, J. Shanthikumar.* Stochastic orders and their applications. Academic Press, 1994.

[81] *A.N. Shiryaev.* Essentials of stochastic finance. World Scientific, 1999.

[82] *A.N. Shiryaev.* Probability. 3rd Russian Ed., Moscow, MCCME, 2004.

[83] *A.V. Skorokhod.* Random processes with independent increments. Kluwer, 1991.

[84] *J.M. Stoyanov.* Counterexamples in probability. 2nd Ed., Wiley, 1997.

[85] *G. Szegö* (Ed.). Risk measures for the 21st century. Wiley, 2004.

[86] *G.J. Szekely.* Paradoxes in Probability Theory and Mathematical Statistics. Springer, 2001.

[87] *S. Weber.* Distribution-invariant risk measures, information, and dynamic consistency. Preprint, available at: `http://www.bwl3.uni-bonn.de/Seminar`.

[88] *D. Williams.* Probability with martingales. Cambridge, 1991.

[89] *A. Ziegler.* A game theory analysis of options. 2nd Ed., Springer, 2004.

[90] *V.M. Zolotarev.* Modern theory of summation of random variables. Utrecht, 1997.

Subject Index