Решения задач - СЕМНАДЦАТАЯ КОЛМОГОРОВСКАЯ ОЛИМПИАДА ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ

1. (**1-2**) Дано множество из n гирь с весами $\{1, ..., n\}$. Из них составляется случайный набор S (каждая гиря берется с вероятностью p, независимо от прочих). Рассмотрим два события: $A = \{$среди S есть три гири, две из которых вместе весят как третья$\}$, $B = \{$в S есть гиря с нечетным весом$\}$. Докажите, что $\mathsf{P}(AB) \geq \mathsf{P}(A)\mathsf{P}(B)$.

   Решение. Пусть A_1 - событие, состоящее в том, что есть три гири с четным весом, две из которых в сумме весят как третья, и $A_2 = A \setminus A_1$. Тогда $A_2 \subseteq B$, A_1 не зависит от B и потому

$$\mathsf{P}(AB) = \mathsf{P}(A_1B) + \mathsf{P}(A_2B) = \mathsf{P}(A_1)\mathsf{P}(B) + \mathsf{P}(B) \geq \mathsf{P}(A_1)\mathsf{P}(B) + \mathsf{P}(A_2)\mathsf{P}(B) = \mathsf{P}(A)\mathsf{P}(B).$$

2. Имеется N монет, все попарно различного веса. Наугад выберем две, взвесим и оставим более тяжелую; из оставшихся $N-2$ выберем еще одну, сравним с оставшейся при первом взвешивании, и т.д. Пусть после m-го взвешивания ($m < N-1$) оставлена монета A. Найти вероятность, что она окажется тяжелее в $m+1$-м взвешивании.

   Решение. Пусть первые выбранные $m+2$ монеты имеют номера $i_1, \dots, i_{m+2}$. Монета, которая останется после m взвешиваний, очевидно, самая тяжелая из монет $i_1, \dots, i_{m+1}$. Вероятность того, что она НЕ окажется тяжелее и в следующем взвешивании — это вероятность того, что монета i_{m+2} самая тяжелая из перечисленных; в силу симметрии эта вероятность равна $\frac{1}{m+2}$.

   Ответ: $\frac{m+1}{m+2}$.

3. Пусть $\{S_n, n \geq 0\}$ — простое случайное блуждание ($S_n = \sum_{i=1}^n \xi_i$, где $\{\xi_n, n \geq 1\}$ н.о.р., $\mathsf{P}(\xi_1) = p$, $\mathsf{P}(\xi_1 = -1) = 1-p$) и $0 < m \leq n < 2N$. Найти условную ковариацию S_m и S_n при условии $S_{2N} = 0$.

   Решение. Имеем

$$\mathsf{E}(\xi_1 | S_{2N} = 0) = \frac{1}{2N}\mathsf{E}(S_{2N} | S_{2N} = 0) = 0, \quad \mathsf{E}(\xi_1^2 | S_{2N} = 0) = 1,$$

$$\mathsf{E}(\xi_1\xi_2 | S_{2N} = 0) = \frac{\mathsf{E}\xi_1\xi_2 I\{\xi_1 + \xi_2 + S_{2N-2} = 0\}}{\mathsf{P}(S_{2N} = 0)} =$$

$$= \frac{\mathsf{P}(\xi_1 = \xi_2 = 1)\mathsf{P}(S_{2N-2} = -2) + \mathsf{P}(\xi_1 = \xi_2 = -1)\mathsf{P}(S_{2N-2} = 2) - \mathsf{P}(\xi_1 \neq \xi_2)\mathsf{P}\{S_{2N-2} = 0\}}{\mathsf{P}(S_{2N} = 0)} =$$

$$= \frac{1}{2}\frac{(2N-2)!/(N!(N-2)!)2^{-2N+2} - (2N-2)!/((N-1)!)^2 2^{-2N+2}}{(2N)!/(N!)^2 2^{-2N}} =$$

$$= \frac{1}{2}\frac{4(N(N-1) - N^2)}{2N(2N-1)} = -\frac{1}{2N-1}.$$

Следовательно, в силу одинаковой распределенности и независимости слагаемых

$$\mathsf{E}(S_m | S_{2N} = 0) = \sum_{i=1}^m \mathsf{E}(\xi_i | S_{2N} = 0) = 0,$$

$$\mathsf{E}(S_m | S_{2N} = 0) = \sum_{i=1}^m \sum_{j=1}^n \mathsf{E}(\xi_i\xi_j | S_{2N} = 0) = \sum_{i=1}^m \mathsf{E}(\xi_i^2 | S_{2N} = 0) + \sum_{i=1}^m \sum_{j \neq i, j \leq n} \mathsf{E}(\xi_i\xi_j | S_{2N} = 0) =$$

$$= m + (m(m-1) + m(n-m))\mathsf{E}(\xi_1\xi_2 | S_{2N} = 0) = m - \frac{m(n-1)}{2N-1}.$$

Ответ: $m - \frac{m(n-1)}{2N-1}$.

4. Пусть $X_1, X_2, \dots$ – последовательность одинаково распределенных (возможно, зависимых) случайных величин с конечным математическим ожиданием, и $M_n = \max\{X_1, \dots, X_n\}$, $n \in \mathbb{N}$. Доказать, что $M_n/n \to 0$ п.н. при $n \to \infty$.

   Решение. $|M_n| \le |X_1| + \max\{X_1^+, \dots, X_n^+\}$, где $x^+ = \max\{x, 0\}$ (первое слагаемое – учитывает случай, когда все величины отрицательны), очевидно $|X_1|/n \to 0$ п.н., поэтому далее можем считать, что наши случайные величины неотрицательны. Заметим, что если для элементарного исхода имеем $X_n/n \to 0$ при $n \to \infty$, то $M_n/n \to 0$. В самом деле, для любого $m \in \mathbb{N}$ есть такое K, что $X_n/n < m^{-1}$ при $n > K$; тогда

$$\limsup_{n\to\infty} \frac{M_n}{n} \le \limsup_{n\to\infty} \frac{\max(X_1, \dots, X_K)}{n} + \limsup_{n\to\infty} \frac{\max(X_{K+1}, \dots, X_n)}{n} \le$$

$$\le \limsup_{n\to\infty} \max\left(\frac{X_{K+1}}{K+1}, \dots, \frac{X_n}{n}\right) \le m^{-1}.$$

   Итак, $\mathsf{P}(M_n/n \not\to 0) \le \mathsf{P}(X_n/n \not\to 0) \le \mathsf{P}(\cup_{m=1}^{\infty}\{X_n/n \ge m^{-1}\text{ бч}\})$. Так как $\mathsf{E}X_1 < \infty$, то $\sum_n \mathsf{P}(X_n \ge m^{-1}n) < \infty$, так что по лемме Бореля-Кантелли каждое из событий под знаком обьединения имеет нулевую вероятность.

5. Тренер волейбольной команды утверждает, что его команда чаще побеждает, если ее предыдущая встреча тоже была выиграна. Рассмотрим долю тех побед, после которых была одержана победа (их число обозначим Y), среди всех побед кроме последней в турнире (их число обозначим X). Пусть турнир состоит из 5 встреч и команда выигрывает встречу с вероятностью 1/2, независимо от результата прочих игр (т.е. слова тренера неверны). Найти $\mathsf{E}\left(\frac{Y}{X} | X > 0\right)$.

   Решение. Запишем все последовательности (1 - выигрыш, 0 - проигрыш), для которых $X > 0$, то есть не начинающиеся с 4 нулей, их всего 30, и для каждого запишем значение дроби Y/X: (см следующую страницу, наверняка таблица уехала туда)

   Все исходы равновероятны, так что ответ – среднее этих дробей $49/120 \approx 0,41$; таким образом, выборочное значение данной доли побед, близкое к 1/2, указывает на то, что тренер скорее всего прав.

   Ответ: $\frac{49}{120}$

6. Пусть (X, Y) – гауссовский случайный вектор со средним 0 и матрицей коварияций $\begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$, где $|r| < 1$. Найти а) $\mathsf{E}\max\{X, Y\}$ б) $\mathsf{D}\max\{X, Y\}$.

   Решение. Заметим, что $\max\{x, y\} = \frac{|x-y|+x+y}{2}$ для любых $x, y \in \mathbb{R}$. Так как $X - Y \sim N(0, 2-2r)$, то

$$\mathsf{E}\max\{X, Y\} = \frac{\mathsf{E}|X-Y| + \mathsf{E}(X+Y)}{2} = \frac{1}{2}\sqrt{\frac{2}{\pi}(2-2r)}.$$

   Далее, случайные величины $X-Y, X+Y$ независимы так как $cov(X-Y, X+Y) = 1+r-r-1 = 0$. Следовательно по той же формуле

$$\mathsf{D}\max\{X, Y\} = \frac{1}{4}(\mathsf{D}(X+Y) + \mathsf{E}(X-Y)^2 - (\mathsf{E}|X-Y|)^2) = \frac{1}{4}(2 + 2r + 2 - 2r - \frac{2}{\pi}(2-2r)).$$

   Заметим, что при $r \to \pm 1$ ответы переходят в матожидание и дисперсию X и $|X|$, как и должно быть.

   Ответ: а) $\sqrt{\frac{1-r}{\pi}}$ б) $1 - \frac{1-r}{\pi}$.

7. Пусть $X_1, \dots, X_{50}$ — н.о.р. случайные величины с таким распределением: $\mathsf{P}(X_1 = -1) = 1/9, \mathsf{P}(X_1 = -2/5) = 4/9, \mathsf{P}(X_1 = 1/5) = 4/9$. Найти $\mathsf{P}(X_1 + \ldots + X_{50} = -17)$.

   Решение. Удобно положить $Y_i = 5(X_i + 1)/3$, тогда $Y_1, \dots, Y_{50}$ – нор величины с биномиальным распределением с параметрами $n = 2, p = 2/3$. Но тогда их сумма распределена как

00010	0/1
00011	1/1
00100	0/1
00101	0/1
00110	1/2
00111	2/2
01000	0/1
01001	0/1
01010	0/2
01011	1/2
01100	1/2
01101	1/2
01110	2/3
01111	3/3
10000	0/1
10001	0/1
10010	0/2
10011	1/2
10100	0/2
10101	0/2
10110	1/3
10111	2/3
11000	1/2
11001	1/2
11010	1/3
11011	2/3
11100	2/3
11101	2/3
11110	3/4
11111	4/4

число успехов в 100 испытаниях Бернулли с вероятностью успеха 2/3, и $\mathsf{P}(Y_1 + \ldots + Y_{50} = 5(-17+50)/3) = C_{100}^{55} 2^{55}/3^{100}$.

Ответ: $C_{100}^{55} 2^{55}/3^{100}$.

8. (**3-6**) Пусть случайная величина $X \geq 0$ имеет плотность p, которая непрерывна при $x \geq 0$ и дифференцируема при $x > 0$, причем $p(0) = 0$, $p(x) > 0$ при $x > 0$, $p(x) \to 0$ при $x \to \infty$ и $d(\ln p)/dx$ ограничена снизу при $x > 0$. Доказать, что $X = X_1 + X_2$ по распределению, где X_1 и X_2 независимы и X_1 имеет показательное распределение.

   Решение. Пусть φ — характеристическая функция X. Докажем, что для некоторого $a > 0$ функция $\varphi(t)(1 - iat)$ — характеристическая функция некоторого распределения. Тогда мы возьмем его в качестве распределения X_2, и из того, что $1/(1-iat) = \mathsf{E}e^{it\xi}$, где $\xi \sim \exp(a^{-1})$, получим требуемое утверждение.

   Действительно, рассмотрим функцию $q(x) = p(x)(1 + ap'(x)/p(x))$ $(q(0) = 0)$. Покажем, что это плотность нужного нам распределения. При достаточно малом $a > 0$ функция q неотрицательна (из-за условия на логарифмическую плотность). При этом

$$\int_0^\infty e^{itx} q(x)dx = \varphi(t) + a \int e^{itx} p'(x)dx = \varphi(t) + a \int_0^\infty e^{itx} dp(x) =$$

$$= \varphi(t) - a \int_0^\infty ite^{itx} p(x)dx = \varphi(t) - ait\varphi(t)$$

   (в частности, $\int_0^\infty q(x)dx = 1$), что и требовалось.

9. (**3-6**) Пусть $X_1, \ldots, X_n$ - выборка из одномерного нормального распределения с неизвестными параметрами и $s^2 = \sum_{j=1}^n (X_i - \overline{X})^2/(n-1)$ – оценка дисперсии. Пусть $\mu_k = n^{-1} \sum_{j=1}^n (X_j - \overline{X})^k$, $k \in \{3, 4\}$ – выборочные центральные моменты. Выборочные асимметрия и эксцесс – это статистики μ_3/s^3 и $\mu_4/s^4 - 3$. Доказать, что каждая из них не зависит от s^2.

   Решение. Выберем какой-нибудь ортонормированный базис в пространстве векторов, сумма координат которых равна нулю, и пусть $Y_1, \ldots, Y_{n-1}$ — координаты случайного вектора $X_1 - \overline{X}, \ldots, X_n - \overline{X}$ в этом базисе. Известно, что этот вектор имеет среднее 0 и скалярную матрицу ковариаций $\sigma^2 I$ (так как он ортогональная проекция X на упомянутое пространство). Для такого вектора, случайная величина $\|Y\|$ и случайный вектор $Y/\|Y\|$ независимы. Действительно, так как для любого $x > 0$ и борелевского подмножества единичной сферы $B \subset S^{n-2}$

$$\mathsf{P}\left(\|Y\| \leq x, \frac{Y}{\|Y\|} \in B\right) = (2\pi)^{-d/2} \int\limits_{\|y\| \leq x, y/\|y\| \in B} e^{-\|y\|^2/2} dy = (2\pi)^{-d/2} \int\limits_0^x \int\limits_B r^{n-2} e^{-r^2/2} J(\theta) dr d\theta,$$

   где $J(\theta)$ – та часть якобиана перехода к сферическим координатам, которая не зависит от r; $d = n-1$. Подставляя $x = \infty$ и $B = S^{n-2}$ соответственно, имеем

$$\mathsf{P}\left(\frac{Y}{\|Y\|} \in B\right) = (2\pi)^{-d/2} \int\limits_0^\infty r^{n-2} e^{-r^2/2} dr \int\limits_B J(\theta) d\theta$$

$$\mathsf{P}(\|Y\| \leq x) = (2\pi)^{-d/2} \int\limits_0^x r^{n-2} e^{-r^2/2} dr \int\limits_{S^{n-2}} J(\theta) d\theta$$

$$1 = (2\pi)^{-d/2} \int\limits_0^\infty r^{n-2} e^{-r^2/2} dr \int\limits_{S^{n-2}} J(\theta) d\theta.$$

   Из четырех равенств следует что $\mathsf{P}(\|Y\| \leq x, \frac{Y}{\|Y\|} \in B) = \mathsf{P}(\|Y\| \leq x)\mathsf{P}(\frac{Y}{\|Y\|} \in B)$. Однако s^2 — функция от $\|Y\|$, а выборочная асимметрия и эксцесс — функции от $Y/\|Y\|$.

10. Источник излучения, расположенный на плоскости в начале координат, при нормальной работе испускает частицы равномерно и независимо одна от другой во всех направлениях. Если же он испорчен, то лучи, по которым летят частицы, концентрируются около одного (неизвестного) направления. Предложить способ определить по наблюдениям, исправен ли источник. Более точно, указать такие функции $f_n(X_1, ..., X_n)$ (где X_i –оставляемые частицами отметки на единичной окружности), что для исправного источника соотношение $f_n \to \xi$ по распределению выполняется, а для любого неисправного нет, причем ξ – известное распределение.

    Решение. Пусть S_n — векторная сумма всех $X_1, \ldots, X_n$. При исправном источнике, согласно ЦПТ, $n^{-1/2}S_n \to N(0, I/2)$, так как для $X_1 \sim [0, 2\pi]$ имеем

    $$\mathsf{D}((\cos X_1, \sin X_1)^T) = \frac{1}{2}I.$$

    В частности, $\|n^{-1/2}S_n\| \to \exp(1)$ по распределению при $n \to \infty$. Данное утверждение не зависит от выбора ортонормированного базиса в $\mathbb{R}^2$.

    Пусть теперь лучи собираются вокруг одного направления. Выбирая подходящий базис, можем считать, что это направление вектора $(1, 0)$. Тогда $\mathsf{E}\cos X_1 > 0$, так что $\|n^{-1/2}S_n\| \to \infty$.

11. (**3-6**) Пусть $W = \{W_t, t \geq 0\}$ – стандартное броуновское движение.

    Найти $\lim_{t\to\infty} \sqrt{t}\mathsf{P}(W_s \leq 1, s \leq t)$.

    Решение. По теореме Башелье

    $$\sqrt{t}\mathsf{P}(W_s \leq 1, s \leq t) = 2\sqrt{t}\mathsf{P}(0 \leq W_t \leq 1) = 2\sqrt{t}\mathsf{P}(0 \leq W_1 \leq 1/\sqrt{t}) = \sqrt{\frac{2t}{\pi}}\int_0^{1/\sqrt{t}} e^{-x^2/2}dx \to \sqrt{\frac{2}{\pi}}.$$

    Ответ: $\sqrt{\frac{2}{\pi}}$.

12. Последовательность случайных величин $X_0, X_1, X_2, \ldots$ такова, что $X_0 = X_1 = 1$, $X_{n+1} = |X_n + \varepsilon_n X_{n-1}|, n \geq 1$, где $\{\varepsilon_n\}$ н.о.р., $\mathsf{P}(\varepsilon_1 = \pm 1) = 1/2$. Найти $\mathsf{P}(\min\limits_{n\geq 1} X_n = 0)$.

    Решение. Будем рассматривать наш процесс как случайное блуждание на парах целых неотрицательных чисел $(x, y) = (X_n, X_{n+1}), n \geq 0$. Тогда это блуждание описывается следующим графом

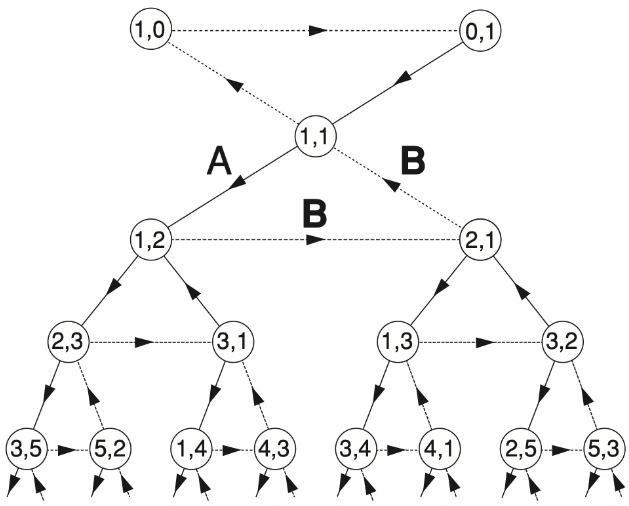

Figure 4. The Fibonacci graph.

    где искомая вероятность – это достижение точки $(1, 0)$.

    Движение "вниз" (увеличивающее максимум из двух чисел) всегда соответствует $\varepsilon_n = 1$, а движение "в сторону" и "вверх" — значению $\varepsilon_n = -1$. Действительно, из точки (x, y)

вниз переходим в точку $(x, x+y)$; после этого два перехода $(x, x+y) \to (x+y, x) \to (x, y)$ возвращают в исходную точку.

Пусть p — вероятность, находясь в корне любого из деревьев на картинке (те в любой из ненулевых вершин), когда-либо пройти из этого корня НЕ вниз (тогда в случае, если корень – точка (1,1), эта вероятность совпадает с той которую хотим найти). Тогда $p = 1/2 + p^3/2$ (если на первом шаге пошли вниз, то нужно, чтобы удалось пройти вправо из корня левого нижнего дерева, затем вверх из правого нижнего и наконец вверх из исходного). У этого уравнения два положительных корня: 1 и $(\sqrt{5}-1)/2$. Почему вероятность не может быть равна 1? Если рассмотрим наш процесс как случайное блуждание по уровням графа (самый верхний - уровень 0, точка (1,1) - уровень 1, далее по возрастанию), то видно, что после попадания на любой уровень, начиная со второго и ниже: вероятность переместиться на уровень с большим номером равна 3/4 (из нижней левой вершины треугольника спустимся либо сразу вниз, либо в правую нижнюю и оттуда вниз); вероятность переместиться на уровень с меньшим номером равна 1/4. Известно, что для такого блуждания вероятность возвращения в начало меньше единицы.