Содержание

1. Раздел 0. Стандартные задачи

Задача 1. Некто имеет N ключей, из которых только один от его двери. Какова вероятность, что, используя ключи в случайном порядке, он откроет дверь

а) первым ключом,

б) последним ключом?

Найти вероятность, что потребуется не менее k попыток, чтобы открыть дверь, если ключи, которые не подошли,

в) откладываются,

г) не откладываются.

Решение. Для каждого из N ключей, которые последовательно вынимаются, введем с.в. $\xi_i = \mathcal{I}($«подошел i-й ключ»$)$. По условию $\mathbb{P}(\xi_i = 1) = \frac{1}{N} = \mathbb{E}\xi_i$.

а) $p_1 = \mathbb{P}(\xi_1 = 1) = \frac{1}{N}$;

б) $p_2 = \mathbb{P}(\xi_N = 1) = \frac{1}{N}$;

в) $p_3 = 1 - \mathbb{P}(\xi_1 + \ldots + \xi_{k-1} = 1) = 1 - \mathbb{E}(\xi_1 + \ldots + \xi_{k-1}) = 1 - \frac{k-1}{N} = \frac{N-k+1}{N}$;

г) $p_4 = \big(\mathbb{P}(\xi_1 = 0)\big)^{k-1} = \big(\frac{N-1}{N}\big)^{k-1}$.

$\square$

Задача 2. Ребенок играет с десятью буквами разрезной азбуки: А, А, А, Е, И, К, М, М, Т, Т. Какова вероятность того, что при случайном расположении букв в ряд он получит слово «МАТЕМАТИКА»?

Решение. Для букв А, А, А есть 3 возможные позиции, которые они могут занять в произвольном порядке, т.е. для них имеется 3! перестановки. Аналогично для букв М, М и Т, Т есть по $2! = 2$ перестановки, для букв Е, И, К есть лишь один вариант расположения в слове. Всего находим $m = 3! \cdot 2^2 = 24$ перестановки, при которых ребенок получит слово «МАТЕМАТИКА». Общее число возможных расположений букв равно $N = 10!$. Искомая вероятность имеет вид

$$p = \frac{m}{N} = \frac{4!}{10!} = \frac{1}{151200} \approx 6.6 \cdot 10^{-6}.$$

$\square$

Задача 3. В лотерее участвует n билетов, из которых m билетов являются выигрышными. Какова вероятность хотя бы одного выигрыша для участника лотереи, имеющего k билетов ($m \leqslant n, k \leqslant n$).

Решение. Заметим, что если $k+m \geqslant n$, то у участника лотереи всегда найдется выигрышный билет, т.е. вероятность выигрыша равна единице. Поэтому считаем $k+m<n$. Число всевозможных наборов k билетов из n равно C_n^k, число всевозможных наборов k билетов из $(n-m)$ проигрышных билетов равно C_{n-m}^k. Отсюда искомая вероятность равна

$$p = 1 - \frac{C_{n-m}^k}{C_n^k}.$$

□

Задача 4. Из урны, содержащей a белых, b черных и c красных шаров (и только их), одновременно извлечены наугад три шара. Какова вероятность того, что все они разного цвета?

Решение. Число всевозможных наборов из трех шаров равно C_{a+b+c}^3. Набор состоит из шаров разного цвета, если взят произвольный шар из a белых, произвольный — из b черных и произвольный — из c красных шаров, т.е. (abc) наборов, удовлетворяющих условию. Искомая вероятность равна

$$p = \frac{abc}{C_{a+b+c}^3}.$$

□

Задача 5. Из урны, содержащей a белых, b черных и c красных шаров (и только их), последовательно извлекаются три шара. Найти вероятность следующих событий:

а) все три шара разного цвета;

б) шары извлечены в последовательности белый, черный, красный;

в) шары извлечены в обратной последовательности.

Решение. а) Из предыдущей задачи находим $p_1 = \frac{abc}{C_{a+b+c}^3}$.

б) Число всевозможных упорядоченных наборов из трех шаров равно A_{a+b+c}^3, а число наборов, удовлетворяющих условию, равно abc. Отсюда находим $p_2 = \frac{abc}{A_{a+b+c}^3}$.

в) Аналогично предыдущему пункту находим $p_3 = \frac{abc}{A_{a+b+c}^3}$.

□

Задача 6. Из урны, содержащей a белых и b черных шаров, извлекается наугад один шар и откладывается в сторону. Какова вероятность того, что извлеченный наугад второй шар окажется белым, если:

а) первый извлеченный шар белый;

б) цвет первого извлеченного шара остается неизвестным?

Решение. Введем с.в. $\xi_i = \mathcal{I}$(«на i-м шаге вынут белый шар»).

а) Искомая вероятность равна

$$p_1 = \mathbb{P}(\xi_2 = 1 \,|\, \xi_1 = 1) = \frac{\mathbb{P}(\xi_2 = 1, \xi_1 = 1)}{\mathbb{P}(\xi_1 = 1)} =$$

$$= \frac{a(a-1) \,/\, (a+b)(a+b-1)}{a \,/\, (a+b)} = \frac{a-1}{a+b-1}.$$

б) Искомая вероятность равна

$$p_2 = \mathbb{P}(\xi_2 = 1) = \mathbb{P}(\xi_2 = 1, \xi_1 = 1) + \mathbb{P}(\xi_2 = 1, \xi_1 = 0) =$$

$$= \frac{a(a-1)}{(a+b)(a+b-1)} + \frac{ab}{(a+b)(a+b-1)} = \frac{a}{a+b}.$$

□

Задача 7. Партия продукции состоит из десяти изделий, среди которых два изделия дефектные. Какова вероятность того, что из пяти отобранных наугад и проверенных изделий:

а) ровно одно изделие дефектное;

б) ровно два изделия дефектные;

в) хотя бы одно изделие дефектное?

Какое наименьшее число изделий необходимо проверить для того, чтобы среди них с вероятностью 0.9 содержалось хотя бы одно дефектное изделие?

Решение. а) Из пяти отобранных наугад и проверенных изделий ровно одно изделие дефектное, если 4 изделия взяты из 8 недефектных, одно — из двух дефектных. Искомая вероятность равна

$$p_1 = \frac{C_8^4\, C_2^1}{C_{10}^5} = \frac{5}{9}.$$

б) Из пяти отобранных наугад и проверенных изделий ровно два изделия дефектные, если 3 изделия взяты из 8 недефектных, остальные два — дефектные. Искомая вероятность равна

$$p_2 = \frac{C_8^3\, C_2^2}{C_{10}^5} = \frac{2}{9}.$$

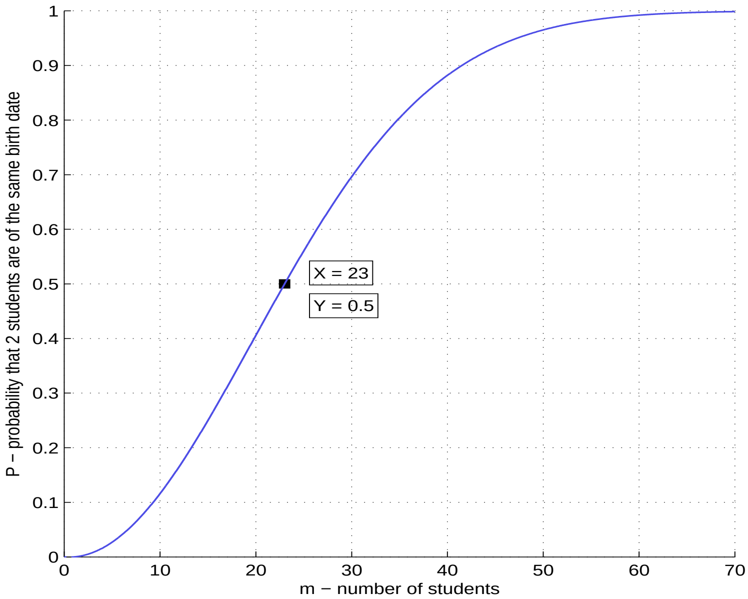
в) Рассмотрим противоположное событие: дефектных деталей нет, если все 5 изделий взяты из 8 недефектных. Искомая вероятность равна

$$p_3 = 1 - \frac{C_8^5}{C_{10}^5} = \frac{7}{9}.$$

Найдем наименьшее n, т.ч. среди n выбранных изделий с вероятностью 0.9 содержалось хотя бы одно дефектное изделие. Рассмотрим противоположное событие: дефектных деталей среди них нет, т.е. все n изделий взяты из 8 недефектных. Искомая вероятность равна

$$p_n = 1 - \frac{C_8^n}{C_{10}^n} = 1 - \frac{(10-n)(9-n)}{90} \geqslant 0.9 \iff n^2 - 19n + 81 \leqslant 0.$$

Из последнего неравенства подбором находим $\hat{n} = 7$. □



Задача 8. Найти вероятность того, что из 50 студентов, присутствующих на лекции, хотя бы двое имеют одну и ту же дату рождения.

Решение. $M = 365$, $m = 50$. Общее число равновозможных исходов равно $n = M^m$, $n_0 = M(M-1)(M-2)\ldots\big(M-(m-1)\big) = A_M^m$ — число исходов, в которых нет ни одного совпадения дат рождения, искомая вероятность равна $P = 1 - \frac{n_0}{n} = 1 - \left(1 - \frac{1}{M}\right)\left(1 - \frac{2}{M}\right)\ldots\left(1 - \frac{m-1}{M}\right)$. Вычислим это равенство приближенно.

$$\ln\frac{n_0}{n} = \sum_{k=1}^{m-1}\ln\left(1 - \frac{k}{M}\right) \approx \sum_{k=1}^{m-1}\left(-\frac{k}{M}\right) = -\frac{1}{M}\frac{m(m-1)}{2}$$

$$P(m) \approx 1 - e^{-\frac{1}{M}\frac{m(m-1)}{2}}$$

Выше приведен график вероятности того, что хотя бы двое из класса имеют одну и туже дату рождения, как функции от числа учеников.

Интересно отметить, что размер класса, где с вероятностью $1\,/\,2$ найдутся по крайней мере два человека с совпадающими днями рождения, не столь уж велик — $m = 23$ человека. □

Задача 9. Известно, что в результате бросания десяти игральных костей выпала по крайней мере одна «шестерка». Какова вероятность того, что число выпавших «шестерок» больше единицы?

Решение. A — выпадение хотя бы одной шестерки, B — выпадение более одной шестерки, $B \subset A$, $p = 1/6$ — вероятность выпадения шестерки в отдельном бросании; искомая вероятность равна

$$\mathbb{P}(B\,|\,A) = \frac{\mathbb{P}(AB)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B)}{\mathbb{P}(A)} = \frac{1 - \mathbb{P}(\overline{B})}{1 - \mathbb{P}(\overline{A})} =$$

$$= \frac{1 - (1-p)^{10} - C_{10}^{1} p (1-p)^9}{1 - (1-p)^{10}} \approx 0.6148.$$

□

Задача 10. Опыт состоит в подбрасывании монеты до тех пор, пока два раза подряд она не выпадет одной и той же стороной. Каждому возможному исходу опыта припишем вероятность $1/2$ (монета «правильная»). Построить пространство элементарных событий и найти вероятности следующих событий:

а) опыт окончится до шестого бросания;

б) для завершения опыта потребуется четное число бросаний.

Решение. В качестве элементарного события рассмотрим реализацию ограниченной последовательности из чередующихся нулей и единиц с двумя подряд идущими одинаковыми цифрами в конце. Для всякого $n > 1$ имеется два таких элементарных события ω_n^1, ω_n^2, соответствующих двум реализациям последовательности длины n, у которых цифры на каждой позиции взаимно противоположны. Их вероятности равны

$$\mathbb{P}(\{\omega_n^1\}) = \mathbb{P}(\{\omega_n^2\}) = \frac{1}{2^n}.$$

Таким образом, определено пространство элементарных исходов $\Omega = \{\omega_n^1, \omega_n^2, n \geq 2\}$. Получим дискретное вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$, где $\mathcal{F}$ — класс всех подмножеств множества Ω.

Рассмотрим событие A_i — завершение опыта на i–м шаге. Тогда A_i при $i > 1$ содержит два элементарных события ω_i^1 и ω_i^2. Отсюда

$$\mathbb{P}(A_1) = 0, \quad \mathbb{P}(A_i) = \frac{2}{2^i} = \frac{1}{2^{i-1}}, \quad i = 2, 3, \ldots.$$

а) Искомая вероятность равна

$$p_1 = \mathbb{P}\Big(\bigcup_{i=2}^{5} A_i\Big) = \sum_{i=2}^{5} \mathbb{P}(A_i) = 1 - \frac{1}{2^4} = 0.9375.$$

б) Искомая вероятность равна

$$p_2 = \mathbb{P}\Big(\bigcup_{i=1}^{\infty} A_{2i}\Big) = \sum_{i=1}^{\infty} \mathbb{P}(A_{2i}) = \sum_{i=1}^{\infty} \frac{1}{2^{2i-1}} = \frac{2}{3}.$$

□

Задача 11. Из колоды, содержащей 52 карты, наугад берется 10 карт. Найти математическое ожидание числа тузов?

Решение. Для каждой из выбранных карт введем случайную величину

$$X_i = \mathcal{I}(\text{«}i\text{-я карта является тузом»}),\ 1 \le i \le 10.$$

Тогда $X_i \in \mathrm{Be}(1/13)$, поскольку имеется 4 туза среди 52 карт. Находим искомое математическое ожидание

$$\mathbb{E}\sum_{i=1}^{10} X_i = \sum_{i=1}^{10} \mathbb{E}X_i = \frac{10}{13}.$$

□

Задача 12. В урне находится m шаров, из которых m_1 белых и m_2 черных ($m_1+m_2 = m$). Производится n извлечений одного шара с возвращением его (после определения его цвета) обратно в урну. Найти вероятность того, что ровно r раз из n будет извлечен белый шар.

Решение. Число раз, когда будет извлечен белый шар, среди n извлечений имеет биномиальное распределение с параметрами n и $\frac{m_1}{m}$. Поэтому искомая вероятность равна

$$p = C_n^r \left(\frac{m_1}{m}\right)^r \left(1 - \frac{m_1}{m}\right)^{n-r} = C_n^r \left(\frac{m_1}{m}\right)^r \left(\frac{m_2}{m}\right)^{n-r}.$$

□

Задача 13. Найти вероятность того, что при размещении n различных шаров по N ящикам заданный ящик будет содержать ровно k: $0 \leqslant k \leqslant n$, шаров (все различимые размещения равновероятны).

Решение. Шары различимы, и каждый из них может находиться в одном из N ящиков, поэтому число размещений n различных шаров по N ящикам равно N^n.

Если же заданный ящик содержит k шаров, то каждый из оставшихся $(n-k)$ шаров, которые можно отобрать из n шаров C_n^k способами, может находиться в одном из $N-1$ ящиков, и число нужных размещений равно $(N-1)^{n-k}C_n^k$. Находим искомую вероятность:

$$p = \frac{(N-1)^{n-k}C_n^k}{N^n}.$$

□

Задача 14. В урне находится m шаров, из которых m_1 — первого цвета, m_2 — второго цвета, …, m_s — s-го цвета $(m_1 + m_2 + \ldots + m_s = m)$. Производится n извлечений одного шара с возвращением его (после определения его цвета) обратно в урну. Найти вероятность того, что r_1 раз будет извлечен шар первого цвета, r_2 раз — шар второго цвета, …, r_s раз — шар s-го цвета $(r_1 + r_2 + \ldots + r_s = n)$.

Решение. Вероятность того, что извлеченный шар будет i-го цвета, равна $p_i = \frac{m_i}{m}$. Тогда вероятность искомого события описывается полиномиальным распределением (обобщение биномиального) с вероятностями p_i, $i = 1, \ldots, s$, то есть

$$p = \frac{n!}{r_1! \ldots r_s!}\left(\frac{m_1}{m}\right)^{r_1} \ldots \left(\frac{m_s}{m}\right)^{r_s}$$

(где $p(r_1, \ldots, r_s) = \frac{n!}{r_1! \ldots r_s!}$ — мультиномиальный коэффициент, получающийся при разложении $(x_1 + \ldots x_n)^n = \sum\limits_{r_1+\ldots+r_s=n} p(r_1, \ldots, r_s)x_1^{r_1} \ldots x_s^{r_s}$). □

Задача 15. В гардеробе все шляпы N посетителей оказались случайным образом перепутанными. Шляпы не имеют внешних отличительных признаков. Какова вероятность того, что хотя бы один посетитель получит свою шляпу (рассмотреть случаи $N = 4$, $N = 10000$)?

Решение. Обозначим A_i — совпадение шляпы у i-го посетителя, B — совпадение шляпы хотя бы у одного посетителя. Ввиду совместности событий A_i искомая вероятность равна

$$\mathbb{P}(B) = \mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_i \mathbb{P}(A_i) - \sum_{i_1<i_2} \mathbb{P}(A_{i_1}A_{i_2}) + \sum_{i_1<i_2<i_3} \mathbb{P}(A_{i_1}A_{i_2}A_{i_3}) - \ldots - (-1)^n \mathbb{P}(A_{i_1} \ldots A_{i_n}) =$$

$$= \sum_{k=1}^{n} S_k = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots - (-1)^n \frac{1}{n!} \xrightarrow{n\to\infty} e^{-1}$$

(поскольку каждая сумма $S_k = \sum\limits_{i_1<\ldots<i_k} \mathbb{P}(A_{i_1} \ldots A_{i_k})$ содержит C_n^k равнозначных слагаемых, каждое из которых представляет собой вероятность совпадения шляп у k посетителей $A_{i_1}, \ldots, A_{i_k}$ и равно $\frac{(n-k)!}{n!}$; отсюда $S_k = C_n^k \frac{(n-k)!}{n!} = \frac{1}{k!}$).

$$N = 4: \quad P = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} \approx 0.4167,$$

$$N = 10000: \quad P \approx e^{-1} \approx 0.4581.$$

□

Задача 16. На восьми одинаковых карточках написаны соответственно числа 2, 4, 6, 7, 8, 11, 12, 13. Наугад берутся две карточки. Определить вероятность того, что образованная из двух полученных чисел дробь сократима.

Решение. Число всевозможных пар карточек равно C_8^2. Заметим, что все нечетные числа на данных карточках являются простыми числами, которым некратно ни одно из других чисел. Поэтому любая пара чисел, в которой есть нечетное число, образует несократимую дробь. Соответственно, образованная из двух полученных чисел дробь сократима в том и только в том случае, когда оба числа являются четными. Четных чисел среди карточек равно 5, поэтому число пар, образующих сократимую дробь, равно C_5^2. Находим искомую вероятность

$$p = \frac{C_5^2}{C_8^2} = \frac{4 \cdot 5}{7 \cdot 8} = \frac{5}{14}.$$

□

Задача 17. Бросается n игральных костей. Найти вероятность события, состоящего в том, что на всех костях выпало одинаковое число очков.

Решение. Число всевозможных комбинаций очков, выпадаемых на n костях, равно 6^n. Число комбинаций, удовлетворяющих условию, равно 6: каждая комбинация соответствует тому числу очков, которое выпало на всех костях. Находим искомую вероятность

$$p = \frac{6}{6^n} = \frac{1}{6^{n-1}}.$$

□

Задача 18. Монета подбрасывается n раз. Найти вероятность того, что число выпадений герба нечетно.

Решение. Число всевозможных исходов в результате n подбрасываний монеты равно 2^n. Количество исходов, в которых число выпадений герба нечетно, равно $n_1 = C_n^1 + C_n^3 + \ldots$. Наряду с этим рассмотрим количество исходов, в которых число выпадений герба четно: $n_2 = C_n^0 + C_n^2 + \ldots$. Из формул бинома Ньютона

$$0 = (1-1)^n = C_n^0 - C_n^1 + C_n^2 - C_n^3 + \ldots = n_2 - n_1, \quad 2^n = (1+1)^n = C_n^0 + C_n^1 + C_n^2 + C_n^3 + \ldots = n_2 + n_1,$$

откуда $n_1 = n_2 = 2^{n-1}$. Находим искомую вероятность

$$p = \frac{n_1}{2^n} = \frac{1}{2}.$$

□

Задача 19. Брошены шесть игральных костей. Найти вероятность следующих событий:

а) на всех костях выпало разное число очков;

б) суммарное число выпавших очков равно 7.

Решение. а) Число всевозможных комбинаций очков, выпадаемых на шести костях, равно 6^6. Число комбинаций, удовлетворяющих условию, равно $6!$ — количество различных перестановок чисел от 1 до 6. Находим искомую вероятность

$$p = \frac{6!}{6^6} = \frac{5!}{6^5} \approx 0.0154.$$

б) Число всевозможных комбинаций очков, выпадаемых на шести костях, равно 6^6. В комбинациях, удовлетворяющих условию, на одной из костей выпало два очка, на остальных пяти костях выпало по одному очку. Число таких комбинаций равно 6. Находим искомую вероятность

$$p = \frac{6}{6^6} = \frac{1}{6^5} \approx 1.3 \cdot 10^{-4}.$$

□

Задача 20. Игральная кость бросается n раз. Какова вероятность того, что:

а) хотя бы один раз выпадает «шестерка»;

б) «шестерка» выпадает ровно один раз?

Решение. а) Число всевозможных комбинаций очков, выпадаемых на n костях, равно 6^n. Число комбинаций, в которых не выпадет ни одна «шестерка», равно 5^n. Находим искомую вероятность

$$p = 1 - \frac{5^n}{6^n}.$$

б) Число всевозможных комбинаций очков, выпадаемых на n костях, равно 6^n. Число комбинаций, в которых выпадет ровно одна «шестерка», равно $C_n^1 5^{n-1} = n\, 5^{n-1}$. Находим искомую вероятность

$$p = n\, \frac{5^{n-1}}{6^n}.$$

□

Задача 21. Несколько раз бросается игральная кость. Какое событие более вероятно:

а) сумма выпавших очков четна;

б) сумма выпавших очков нечетна?

Решение. Найдем вероятность первого события. Пусть игральная кость бросается n раз. Число всевозможных комбинаций очков, выпадаемых на n костях, равно 6^n. В комбинации, в которой сумма выпавших очков четна, количество костей с нечетными номерами четно. Число таких комбинаций равно $n_1 = 3^n\,(C_n^0 + C_n^2 + \ldots)$. В одной из предыдущих задач выяснялось, что $C_n^0 + C_n^2 + \ldots = 2^{n-1}$. Находим искомую вероятность

$$p_1 = \frac{n_1}{6^n} = \frac{2^{n-1}3^n}{6^n} = \frac{1}{2}.$$

Отсюда вероятность второго события $p_2 = \frac{1}{2}$, т.е. события равновероятны. □

Задача 22. Для уменьшения общего количества игр $2n$ команд спортсменов разбиваются на две подгруппы. Определить вероятности того, что две наиболее сильные команды окажутся:

а) в одной подгруппе;

б) в разных подгруппах.

Решение. Заметим, что события в двух пунктах задачи являются взаимоисключающими, поэтому для их вероятностей p_1 и p_2 выполнено $p_1 + p_2 = 1$.

Найдем вероятность первого события. Число различных способов выбрать n команд для первой подгруппы из $2n$ команд равно C_{2n}^n. Если две наиболее сильные команды окажутся в первой подгруппе, то остальные команды для первой подгруппы можно выбрать C_{2n-2}^{n-2} способами. Если же две наиболее сильные команды окажутся во второй подгруппе, то команды для первой подгруппы можно выбрать C_{2n-2}^n способами. Находим вероятность первого события

$$p_1 = \frac{C_{2n-2}^{n-2} + C_{2n-2}^n}{C_{2n}^n} = \frac{n-1}{2n-1}.$$

Отсюда вероятность второго события $p_2 = \frac{n}{2n-1}$. □

Задача 23. 40 участников турнира разбиваются на четыре равные группы. Определить вероятность того, что четыре сильнейших участника окажутся в разных группах.

Решение. Число различных способов разбиения 40 участников на четыре группы по 10 человек равно $N = \frac{40!}{(10!)^4}$. Если четыре сильнейших участника распределены по одному в каждой из групп (что можно сделать $4!$ способами), то остальные 36 участника можно разбить на четыре группы по 9 человек $n_1 = \frac{36!}{(9!)^4}$ способами. Находим искомую вероятность

$$p = \frac{n_1 \cdot 4!}{N} = \frac{4! \cdot 10^4}{37 \cdot 38 \cdot 39 \cdot 40} \approx 0.1094.$$

□

Задача 24. В урне находятся черные и белые шары, которые наугад по одному без возвращения извлекаются из урны до тех пор, пока урна не опустеет. Какое событие более вероятно:

а) первый извлеченный шар белый;

б) последний извлеченный шар белый?

Решение. Обозначим через n_1 число белых шаров, n_2 — число черных шаров, $N = n_1+n_2$.

а) Искомая вероятность равна $p = \frac{n_1}{N}$.

б) Считаем все шары пронумерованными и рассмотрим процесс изъятия всех шаров как последовательность из N чисел. Число всевозможных исходов равно $N!$. Если же последний шар белый, т.е. последнее число в последовательности принимает одно из n_1 значений, то число таких исходов равно $n_1 \cdot (N-1)!$. Находим искомую вероятность

$$p_2 = \frac{n_1 \cdot (N-1)!}{N!} = \frac{n_1}{N}.$$

Значит, рассматриваемые в задаче события равновероятны. □

Задача 25. В урне находятся черные и белые шары, причем отношение числа белых шаров к числу черных шаров равно L. Найти вероятность того, что при последовательном извлечении наугад всех шаров из урны последним окажется черный шар.

Решение. В обозначениях предыдущей задачи имеем $n_1 = Ln_2$, $N = (L+1)n_2$. Искомая вероятность так же равна вероятности извлечь черный шар на первом шаге, поэтому

$$p = \frac{n_2}{N} = \frac{1}{L+1}.$$

□

Задача 26. В урне находятся a белых и b черных шаров, Шары наугад по одному извлекаются из урны без возвращения. Найти вероятность того, что k-й вынутый шар оказался белым.

Решение. Считаем все шары пронумерованными и рассмотрим процесс изъятия всех шаров как последовательность из $a+b$ чисел. Число всевозможных исходов равно $(a+b)!$. Если же k-й вынутый шар оказался белым, т.е. число на k-м месте в последовательности принимает одно из a значений, то число таких исходов равно $a \cdot (a+b-1)!$. Находим искомую вероятность

$$p = \frac{a \cdot (a+b-1)!}{(a+b)!} = \frac{a}{a+b}.$$

□

Задача 27. 30 шаров размещаются по 8 ящикам так, что для каждого шара одинаково возможно попадание в любой ящик. Найти вероятность размещения, при котором будет 3 пустых ящика, 2 ящика — с тремя, 2 ящика — с шестью и 1 ящик — с двенадцатью шарами.

Решение. Число всевозможных размещений шаров по ящикам равно 8^{30}. Число способов выбрать ящики с указанными в условии количествами шаров в них равно мультиномиальному коэффициенту $p(3,2,2,1) = \dfrac{8!}{3!2!2!}$. Если число шаров для ящиков зафиксировано, то 30 шаров можно разместить по ящикам указанным в условии образом можно $p(12,6,6,3,3) = \dfrac{30!}{12!6!6!3!3!}$ способами. Находим искомую вероятность

$$p = \frac{p(12,6,6,3,3)\,p(3,2,2,1)}{8^{30}}.$$

□

Задача 28. N частиц случайно и независимо друг от друга размещаются в k ячейках так, что каждая из них попадает в i-ую ячейку с вероятностью p_i ($i = 1,\ldots,k, \sum\limits_{i=1}^{k} p_i = 1$). Найти вероятность того, что число частиц в ячейках примет заданные значения n_1, …, n_i, …, n_k (полиномиальное распределение).

Решение. N частиц можно разбить на k групп с числом частиц n_1, …, n_k в этих группах $\frac{N!}{n_1!\ldots n_k!}$ способами (мультиномиальный коэффициент). Каждое такое распределение частиц по ячейкам произойдет, учитывая независимость поведения частиц, с вероятностью $p_1^{n_1}\ldots p_k^{n_k}$. Находим искомую вероятность

$$p = \frac{N!}{n_1!\ldots n_k!}\,p_1^{n_1}\ldots p_k^{n_k}.$$

□

Задача 29. Некто обладает одной облигацией, которую намеревается продать в один из последующих четырех дней, в которых цена облигации принимает различные значения, априори неизвестные, но становящиеся известными в начале каждого дня продаж. Предполагается, что цены облигации независимы и их перестановки по торговым дням равновозможны. Какова стратегия продавца, состоящая в выборе дня продажи облигации и гарантирующая максимальную вероятность того, что он продаст облигацию в день ее наибольшей цены?

Решение. Выбор для продажи облигации первого дня (как, впрочем, любого из четырех дней) без какой либо информации об ее цене в различных днях, дает для вероятности выбора дня с наибольшей ценой значение $1/4$ (не зависящее, заметим, от масштаба цен). Вместо такой тривиальной стратегии рассмотрим две следующие возможные стратегии:

а) на первом шаге (в первый день торгов) запомним имевшую место цену облигации, не продавая ее, а затем продадим облигацию в тот день, когда ее цена окажется большей цены, зафиксированной в первый день, или (когда такого дня не окажется) в последний (четвертый) день, независимо от цены этого дня (стратегия S_1);

б) не продавая облигацию в первом и втором торговых днях, зафиксируем максимальную цену из двух, имевших место для этих дней, и продадим облигацию в третьем торговом дне, если цена облигации в нем будет выше, чем указанная зафиксированная максимальная цена, или, в противном случае, в четвертом дне (стратегия S_2).

Сравним эти стратегии по значению вероятности выбора дня продажи облигации, в котором ее цена имела наибольшее значение.

Поскольку решение нашей задачи не зависит от абсолютных значений цены, положим их равными 1, 2, 3, 4 и выпишем все $4! = 24$ равновозможных вариантов распределений значений цены облигации по торговым дням:

1234	2134	3124*+	4123
1243+	2143*+	3142*+	4132
1324+	2314+	3214*+	4213
1342+	2341+	3241*+	4231
1423*	2413*	3412*	4312
1432*	2431*	3421*	4321

Варианты, отмеченные звездочкой (*), благоприятствуют успеху при стратегии S_1, варианты, отмеченные плюсом (+), приводят к успеху при стратегии S_2. Тривиальным стратегиям, состоящим в безусловном выборе для продажи облигации какого–либо априори определенного дня, соответствуют наборы по шесть благоприятствующих вариантов. Следовательно, оптимальной является стратегия S_1, при которой вероятность продажи облигации в день ее наибольшей цены составляет $p^0 = P(S_1) = \frac{11}{24} \approx 0.458$. При стратегии S_2 эта вероятность оставляет $P(S_2) = \frac{10}{24} \approx 0.417$, а при тривиальной стратегии (безусловном выборе, например, первого торгового дня) — всего $p = \frac{6}{24} = 0.25$. □

Задача 30. По схеме случайного выбора с возвращением из множества чисел $\{1, 2, \dots, N\}$ выбираются числа X и Y. Найти

а) вероятность $\mathbb{P}\{X + Y < N\}$,

б) предел этой вероятности при $N \to \infty$.

Решение. $n_0 = N^2$ — общее число исходов, $n = \dfrac{(N-1)(N-2)}{2}$ — число исходов, благоприятствующих событию $\{X + Y < N\}$, отсюда

$$\mathbb{P}\{X + Y < N\} = \frac{(N-1)(N-2)}{2N^2} \xrightarrow{N\to\infty} 0.5.$$

□

Задача 31. Из n лотерейных билетов k — выигрышные ($n \geqslant 2k$). Какова вероятность, что среди k купленных билетов по крайней мере один будет выигрышным?

Решение. Вероятность того, что среди купленных билетов нет выигрышных: $\dfrac{C_{n-k}^k}{C_n^k}$. Тогда искомая вероятность $p = 1 - \dfrac{C_{n-k}^k}{C_n^k}$. □

Задача 32. Из совокупности всех подмножеств множества $\{1, 2, \ldots, N\}$ по схеме выбора с возвращением выбираются множества A и B. Найти вероятность, что A и B не пересекаются.

Решение. Имеется 2^N различных подмножеств множества $\{1, 2, \ldots, N\}$. Отсюда каждое из множеств A и B можно выбрать 2^N способами, а пару множеств A и B: 4^N способами.

Пусть A и B не пересекаются и $|A| = k$, $0 \leqslant k \leqslant N$. Таких множеств A имеется C_N^k, при этом множество B является подмножеством $\{1, 2, \ldots, N\} \setminus A$ и его можно выбрать 2^{N-k} способами. Число пар множеств A и B, которые не пересекаются, таким образом, равно $n_1 = \sum\limits_{k=0}^{N} C_N^k 2^{N-k}$. Исходя из формулы бинома Ньютона, находим искомую вероятность:

$$\sum_{k=0}^{N} C_N^k 2^{N-k} = (1+2)^N = 3^N \Rightarrow p = \frac{n_1}{4^N} = \left(\frac{3}{4}\right)^N.$$

□

Задача 33. Показать, что борелевская σ-алгебра в $\mathbb{R}^1$, содержащая все числовые промежутки вида $[a, b)$, содержит все промежутки вида (a, b), $(a, b]$, $[a, b]$ и отдельные точки прямой.

Решение. Каждый из данных промежутков получается из числового промежутка вида $[a, b)$ добавлением или вычитанием одной из точек a или b. Поскольку σ-алгебра замкнута относительно операций объединения, пересечения и вычитания, достаточно показать, что каждая отдельная точка $c \in \mathbb{R}^1$ принадлежит борелевской σ-алгебре: $\{c\} = \bigcap\limits_{n=1}^{\infty} [c, c+\frac{1}{n})$. □

Задача 34. Пусть $\Omega = [a, b]$, F — σ–алгебра, содержащая все отрезки $[a, b]$ ($a \leqslant \alpha < \beta \leqslant b$) с вероятностной мерой $\mathbb{P}\{\omega \in [\alpha, \beta]\} = \dfrac{\mathrm{mes}[\alpha, \beta]}{\mathrm{mes}[a, b]}$. Показать, что

a) $\mathbb{P}\{\omega = c = \mathrm{const}\} = 0$;

б) $\mathbb{P}\{\omega_1 = \omega_2\} = 0$.

Найти вероятность, что для трех исходов ω_1, ω_2, ω_3 третий лежит между первыми двумя.

Решение. а) $\{\omega = c\} = \bigcap\limits_{n=1}^{\infty} \Big\{|\omega - c| < \frac{1}{n}\Big\}$

$$\Rightarrow \mathbb{P}\{\omega = c\} = \lim_{n\to\infty} \mathbb{P}\Big\{|\omega - c| < \frac{1}{n}\Big\} = 0$$

(теорема о непрерывности вероятности);

б) решается аналогично при переходе к двумерному пространству исходов $\Omega = [a, b]\times[a, b]$, в котором $\mathbb{P}\{(\omega_1, \omega_2) \in S \subset \Omega\} = \frac{|S|}{|\Omega|}$.

В связи с предполагаемой независимостью исходов, шесть вариантов их расположения на отрезке $[a, b]$ равновозможны; из них в двух вариантах третий исход лежит между первыми двумя. Следовательно, искомая вероятность равна $1/3$. $\square$

Задача 35. Может ли число всех событий какого-либо вероятностного пространства быть равным 129; 130; 128?

Решение. Имеем дискретное вероятностное пространство, изоморфное пространству с конечным множеством элементарных исходов Ω: $|\Omega| = n$, класс событий которого является множеством всех подмножеств Ω, т.е. число всех событий равно 2^n. Исходя из условия задачи, таким может быть только число $128 = 2^7$. $\square$

Задача 36. Число элементарных событий некоторого вероятностного пространства равно n. Указать минимальное и максимальное возможные значения для числа событий.

Решение. Минимальное значений $N_{\min}$ для числа событий соответствует тривиальному классу событий: $\mathcal{F}_1 = \{\varnothing, \Omega\}$, т.е. $N_{\min} = 2$.

Максимальное значение $N_{\max}$ для числа событий соответствует максимальному классу событий: $\mathcal{F}_2 = 2^{\Omega}$ — множество всех подмножеств Ω, отсюда $N_{\max} = 2^n$. $\square$

Задача 37. Доказать справедливость равенства $\mathbb{P}(A \bigtriangleup B) = \mathbb{P}(A) + \mathbb{P}(B) - 2\mathbb{P}(A\,B)$.

Решение. Как уже выяснилось, $\mathbb{P}(A\overline{B}) = \mathbb{P}(A) - \mathbb{P}(AB)$, $\mathbb{P}(B\overline{A}) = \mathbb{P}(B) - \mathbb{P}(BA)$. Имеем

$$\mathbb{P}(A \bigtriangleup B) = \mathbb{P}(A\overline{B} \cup B\overline{A}) = \mathbb{P}(A\overline{B}) + \mathbb{P}(B\overline{A}) =$$

$$= \mathbb{P}(A) - \mathbb{P}(AB) + \mathbb{P}(B) - \mathbb{P}(BA) = \mathbb{P}(A) + \mathbb{P}(B) - 2\mathbb{P}(A\,B).$$

$\square$

Задача 38. Пусть A, B, C — заданные события. Доказать справедливость неравенств

а) $\mathbb{P}(AB) + \mathbb{P}(AC) + \mathbb{P}(BC) \geqslant \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - 1$;

б) $\mathbb{P}(AB) + \mathbb{P}(AC) - \mathbb{P}(BC) \leqslant \mathbb{P}(A)$;

в) $\mathbb{P}(A \triangle B) \leqslant \mathbb{P}(A \triangle C) + \mathbb{P}(C \triangle B)$.

Решение. а) По формуле включений и исключений

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(AB) - \mathbb{P}(BC) - \mathbb{P}(AC) + \mathbb{P}(ABC).$$

Отсюда имеем

$$1 \geqslant \mathbb{P}(A \cup B \cup C) - \mathbb{P}(ABC) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(AB) - \mathbb{P}(BC) - \mathbb{P}(AC)$$
$$\Rightarrow\ \mathbb{P}(AB) + \mathbb{P}(AC) + \mathbb{P}(BC) \geqslant \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - 1.$$

б) По формуле включений и исключений для множеств AB и AC

$$\mathbb{P}\big(A(B \cup C)\big) = \mathbb{P}(AB) + \mathbb{P}(AC) - \mathbb{P}(ABC).$$

Т.к. $A(B \cup C) \subset A$, $ABC \subset BC$, из монотонности вероятности следует цепочка неравенств, обосновывающая искомое неравенство:

$$\mathbb{P}(A) \geqslant \mathbb{P}\big(A(B \cup C)\big) = \mathbb{P}(AB) + \mathbb{P}(AC) - \mathbb{P}(ABC) \geqslant \mathbb{P}(AB) + \mathbb{P}(AC) - \mathbb{P}(BC).$$

в) Используя формулу из предыдущей задачи, перепишем исходное уравнение в следующем виде, после приведения подобных слагаемых:

$$\mathbb{P}(AC) + \mathbb{P}(BC) - \mathbb{P}(AB) \leqslant \mathbb{P}(C).$$

Данное неравенство доказано в предыдущем пункте (поменяв местами A и C).

$\square$

Задача 39. В каждую из n пронумерованных ячеек в случайном порядке помещается один из n так же пронумерованных шаров. Найти вероятность того, что ни в одной из ячеек номер шара не совпадет с номером ячейки.

Решение. Задача Монмора о совпадениях.

A_i — совпадение номеров в i-ой ячейке, B — совпадение номеров ячейки и шара хотя бы в одной из ячеек; искомая вероятность равна $\mathbb{P}(\bar{B}) = 1 - \mathbb{P}(B)$. Ввиду совместности событий A_i получим

$$\mathbb{P}(B) = \mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_i \mathbb{P}(A_i) - \sum_{i_1<i_2} \mathbb{P}(A_{i_1}A_{i_2}) + \sum_{i_1<i_2<i_3} \mathbb{P}(A_{i_1}A_{i_2}A_{i_3}) - \ldots - (-1)^n \mathbb{P}(A_{i_1}\ldots A_{i_n}) =$$
$$= \sum_{k=1}^{n} S_k = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots - (-1)^n \frac{1}{n!} \xrightarrow{n\to\infty} e^{-1}$$

(поскольку каждая сумма $S_k = \sum\limits_{i_1<\ldots<i_k} \mathbb{P}(A_{i_1}\ldots A_{i_k})$ содержит C_n^k равнозначных слагаемых, каждое из которых представляет собой вероятность совпадения шляп у k посетителей $A_{i_1}, \ldots, A_{i_k}$ и равно $\frac{(n-k)!}{n!}$; отсюда $S_k = C_n^k \frac{(n-k)!}{n!} = \frac{1}{k!}$). $\square$

Задача 40. Имеются две урны. В одной из них находится один белый шар, в другой — один черный шар (других шаров урны не содержат). Выбирается наугад одна урна. В нее добавляется один белый шар и после перемешивания один из шаров извлекается. Извлеченный шар оказался белым. Определить апостериорную вероятность того, что выбранной оказалась урна, которая первоначально содержала белый шар.

Решение. Обозначим H_1 и H_2 гипотезы о том, что в урне первоначально находился (соответственно) один белый и один черный шар; $\mathbb{P}(H_1) = \mathbb{P}(H_2) = 0.5$. Событие B — извлечение белого шара; $\mathbb{P}(B \mid H_1) = 1$, $\mathbb{P}(B \mid H_2) = 0.5$. Применим формулу Байеса:

$$\mathbb{P}(H_1 \mid B) = \frac{\mathbb{P}(B \mid H_1)\mathbb{P}(H_1)}{\mathbb{P}(B \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(B \mid H_2)\mathbb{P}(H_2)} = \frac{2}{3}.$$

□

Задача 41. В условиях предыдущей задачи добавление белого шара, перемешивание и извлечение одного шара производится n раз. Всегда извлеченный шар оказывался белым. Определить ту же вероятность, что и в предыдущей задаче.

Решение. Обозначим H_1 и H_2 гипотезы о том, что в урне первоначально находился (соответственно) один белый и один черный шар; $\mathbb{P}(H_1) = \mathbb{P}(H_2) = 0.5$. Событие B — извлечение белого шара n раз; $\mathbb{P}(B \mid H_1) = 1$, $\mathbb{P}(B \mid H_2) = 0.5^n$. Применим формулу Байеса:

$$\mathbb{P}(H_1 \mid B) = \frac{\mathbb{P}(B \mid H_1)\mathbb{P}(H_1)}{\mathbb{P}(B \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(B \mid H_2)\mathbb{P}(H_2)} = \frac{2^n}{1 + 2^n}.$$

□

Задача 42. В первой урне содержится a белых и b черных шаров (и только они), во второй — c белых и d черных шаров (и только они). Из выбранной наугад урны извлекается один шар, который обратно не возвращается. Извлеченный шар оказался белым. Найти вероятность того, что и второй шар, извлеченный из той же урны, окажется белым.

Решение. Обозначим H_1 и H_2 гипотезы о выборе первой и второй урны соответственно, $\mathbb{P}(H_1) = \mathbb{P}(H_2) = 0.5$. При извлечении первого белого шара (событие A_1) априорные вероятности гипотез согласно формуле Байеса следует принять равными

$$\mathbb{P}(H_1 \mid A_1) = \frac{\mathbb{P}(A_1 \mid H_1)\mathbb{P}(H_1)}{\mathbb{P}(A_1 \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(A_1 \mid H_2)\mathbb{P}(H_2)}; \quad \mathbb{P}(H_2 \mid A_1) = 1 - \mathbb{P}(H_1 \mid A_1),$$

где $\mathbb{P}(A_1 \mid H_1) = \dfrac{a}{a+b}$, $\mathbb{P}(A_1 \mid H_2) = \dfrac{c}{c+d}$. Вероятность извлечение вторым белого шара (событие A_2) находится из формулы полной вероятности

$$\mathbb{P}(A_2 \mid A_1) = \mathbb{P}(A_2 \mid H_1 A_1)\mathbb{P}(H_1 \mid A_1) + \mathbb{P}(A_2 \mid H_2 A_1)\mathbb{P}(H_2 \mid A_1),$$

где $\mathbb{P}(A_2 \mid H_1 A_1) = \dfrac{a-1}{a+b-1}$, $\mathbb{P}(A_2 \mid H_2 A_1) = \dfrac{c-1}{c+d-1}$. □

Задача 43. Два стрелка стреляют по мишени. Один из них попадает в цель в среднем в 50% случаев, а второй — в 80% случаев. Перед выстрелом они бросают правильную монету для определения очередности. Посторонний наблюдатель знает показатели меткости стрелков, но не знает, кто из них в данный момент стреляет. Наблюдатель видит, что стрелок попал в цель. Какова вероятность того, что стрелял первый стрелок?

Решение. Обозначим гипотезы о том, что стрелял первый или второй стрелок соответственно H_1 и H_2, $\mathbb{P}(H_1) = \mathbb{P}(H_2) = 0.5$. Событие B — попадание стрелка в цель, $\mathbb{P}(B \mid H_1) = 0.5$, $\mathbb{P}(B \mid H_2) = 0.8$. Применим формулу Байеса

$$\mathbb{P}(H_1 \mid B) = \frac{\mathbb{P}(B \mid H_1)\mathbb{P}(H_1)}{\mathbb{P}(B \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(B \mid H_2)\mathbb{P}(H_2)} = \frac{0.5}{1.3} \approx 0.38.$$

□

Задача 44. В урне 7 белых и 3 черных шара. Без возвращения одновременно извлекаются 3 шара. Известно, что среди них есть хотя бы один черный шар. Какова вероятность того, что другие два шара белые?

Решение. Число возможных исходов при извлечении трех шаров равно C_{10}^3. Число исходов, при которых извлечен один черный и два белых шара, равно $C_3^1\, C_7^2$. Находим искомую вероятность $p = \dfrac{C_3^1\, C_7^2}{C_{10}^3} = \dfrac{21}{40} = 0.525$. □

Задача 45. Известно, что 96% выпускаемой продукции соответствует стандарту. Упрощенная схема контроля признает годным с вероятностью 0.98 каждый стандартный экземпляр аппаратуры и с вероятностью 0.05 — каждый нестандартной экземпляр аппаратуры. Найти вероятность, что изделие, прошедшее контроль, соответствует стандарту.

Решение. H_1 и H_2 — гипотезы о соответствии и несоответствии изделия стандарту, $\mathbb{P}(H_1) = 0.96$, $\mathbb{P}(H_2) = 0.04$. Событие B — прохождение контроля, $\mathbb{P}(B \mid H_1) = 0.98$, $\mathbb{P}(B \mid H_2) = 0.05$. Применим формулу Байеса

$$\mathbb{P}(H_1 \mid B) = \frac{\mathbb{P}(B \mid H_1)\mathbb{P}(H_1)}{\mathbb{P}(B \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(B \mid H_2)\mathbb{P}(H_2)} \approx 0.998.$$

□

Задача 46. В $m+1$ урне содержится по m шаров, причем урна с номером n содержит n белых и $m-n$ черных шаров ($n = 0, 1, \ldots, m$). Случайным образом выбирается урна и из нее k раз с возвращением извлекаются шары. Найти а) вероятность, что следующим также будет извлечен белый шар, при условии, что все k шаров оказались белыми,

б) ее предел при $m \to \infty$.

Решение. Обозначим: H_n — гипотеза о n-ом номере урны (о нахождении в ней n белых шаров), событие A — извлечение из урны k белых шаров при k извлечениях, событие B — извлечение из урны белого шара в $k+1$-м извлечении. Имеем $\mathbb{P}(B \,|\, H_1) = 0.98$, $\mathbb{P}(B \,|\, H_2) = 0.05$. Применим формулу Байеса

$$\mathbb{P}(A \,|\, H_n) = \left(\frac{n}{m}\right)^k, \quad \mathbb{P}(H_n \,|\, A) = \frac{\mathbb{P}(A \,|\, H_n)\mathbb{P}(H_n)}{\mathbb{P}(A)} = \frac{\left(\frac{n}{m}\right)^k}{\sum\limits_{n=1}^{m}\left(\frac{n}{m}\right)^k}$$

(формула Байеса). По формуле полной вероятности находим

$$\mathbb{P}(B \,|\, A) = \sum_{n=1}^{m} \mathbb{P}(B \,|\, H_n A)\mathbb{P}(H_n \,|\, A) = \frac{\sum\limits_{n=1}^{m}\left(\frac{n}{m}\right)^{k+1}}{\sum\limits_{n=1}^{m}\left(\frac{n}{m}\right)^k} \xrightarrow{m\to\infty} \frac{k+1}{k+2}.$$

□

Задача 47. Показать, что из независимости событий A и B следует независимость событий A и $\overline{B}$, $\overline{A}$ и B, $\overline{A}$ и $\overline{B}$.

Решение. Достаточно показать, что из независимости событий A и B следует независимость событий A и $\overline{B}$:

$$\mathbb{P}(A \cdot \overline{B}) = \mathbb{P}(\overline{B} \,|\, A)\mathbb{P}(A) = (1 - \mathbb{P}(B \,|\, A))\mathbb{P}(A) = (1 - \mathbb{P}(B))\mathbb{P}(A) = \mathbb{P}(\overline{B})\mathbb{P}(A).$$

□

Задача 48. Показать, что из равенства $\mathbb{P}(A \,|\, B) = \mathbb{P}(A \,|\, \overline{B})$ для ненулевых событий A и B следует равенство $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$, т.е. их независимость.

Решение. Справедливы импликации:

$$\mathbb{P}(AB) + \mathbb{P}(A\overline{B}) = \mathbb{P}(A) \;\Rightarrow\; \mathbb{P}(A \,|\, B)\mathbb{P}(B) + \mathbb{P}(A \,|\, \overline{B})\mathbb{P}(\overline{B}) = \mathbb{P}(A) \;\Rightarrow$$

$$\Rightarrow\; \mathbb{P}(A \,|\, B)(\mathbb{P}(B) + \mathbb{P}(\overline{B})) = \mathbb{P}(A) \;\Rightarrow\; \mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B).$$

□

Задача 49. Подбрасываются три игральные кости. События A, B и C означают выпадение одинакового числа очков (соответственно) на первой и второй, на второй и третьей, на первой и третьей костях. Являются ли эти события независимыми

а) попарно,

б) в совокупности?

Решение. $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \dfrac{1}{6}$, $\mathbb{P}(AB) = \mathbb{P}(BC) = \mathbb{P}(AC) = \dfrac{1}{36}$ (попарная независимость). Но

$$\mathbb{P}(ABC) = \frac{1}{36} \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{216}$$

(зависимость в совокупности). □

Задача 50. Допустим, что вероятность столкновения молекулы с другими молекулами в промежутке времени $[t, t+\Delta t)$ равна $p = \lambda\Delta t + \overline{o}(\Delta t)$ и не зависит от времени, прошедшего после предыдущего столкновения ($\lambda = \mathrm{const}$). Найти распределение времени свободного пробега молекулы и вероятность того, что это время превысит заданную величину t^*.

Решение. Разобьем интервал $\Delta = [0, t)$ на n отрезков равной длины: $\Delta = \bigcup\limits_{i=1}^{n} \Delta_i$, $|\Delta_i| = \Delta t_i = \frac{t}{n}$. Пусть A_i — событие, означающее, что на отрезке Δ_i молекула претерпит столкновение с другими молекулами. По условию задачи

$$\mathbb{P}(A_i) = \lambda\Delta t_i + \overline{o}(\Delta t_i) = \frac{\lambda t}{n} + \overline{o}\Big(\frac{1}{n}\Big)$$

и, далее, $B = \bigcap\limits_{i=1}^{n} \overline{A}_i$, $\mathbb{P}(B) = \left(1 - \left(\dfrac{\lambda t}{n} + \overline{o}\Big(\dfrac{1}{n}\Big)\right)\right)^n \xrightarrow{n\to\infty} e^{-\lambda t}$. □

Задача 51. В каждую i-ую единицу времени живая клетка получает случайную дозу облучения X_i, причем $\{X_i\}_{i=1}^{t}$ имеют одинаковую функцию распределения $F_X(x)$ и независимы в совокупности $\forall t$. Получив интегральную дозу облучения, равную ν, клетка погибает. Оценить среднее время жизни клетки $\mathbb{E}T$.

Решение. Условие гибели клетки $S_T = \sum\limits_{i=1}^{T} X_i \geqslant \nu$. Используя тождество Вальда: $\mathbb{E}S_T = \mathbb{E}T \cdot \mathbb{E}X$, получаем

$$\mathbb{E}T \geqslant \frac{\nu}{\mathbb{E}X}$$

Тождество Вальда: введем Y_j

$$Y_j = \begin{cases} 1, & \text{если} \quad X_1 + \ldots + X_{j-1} = S_{j-1} < \nu, \\ 0, & \text{в остальных случаях} . \end{cases}$$

(X_j и Y_j независимы). Далее

$$S_T = \sum_{j=1}^{\infty} Y_j X_j \Rightarrow \mathbb{E}S_T = \mathbb{E}X \sum_{j=1}^{\infty} \mathbb{E}Y_j = \mathbb{E}X \sum_{j=1}^{\infty} \mathbb{P}\{Y_j = 1\} = \mathbb{E}X \sum_{j=1}^{\infty} \mathbb{P}\{T \geqslant j\} = \mathbb{E}X \cdot \mathbb{E}T.$$

□

Задача 52. Урна содержит N шаров с номерами от 1 до N. Пусть K — наибольший номер, полученный при n их поштучных извлечениях с возвращением. Найти

а) распределение K,

б) асимптотику математического ожидания $\mathbb{E}K$ при $N \to \infty$.

Решение. $F_K(k) = \mathbb{P}\{K < k\} = \left(\frac{k-1}{N}\right)^n$ (вероятность того, что номера всех n извлеченных шаров меньше k). Отсюда при $N \to \infty$

$$\mathbb{E}K = \sum_{j=0}^{N} \mathbb{P}\{K > j\} = \sum_{j=0}^{N} (1 - \mathbb{P}\{K \leqslant j\}) = \sum_{j=0}^{N} (1 - \mathbb{P}\{K < j+1\}) =$$

$$= \sum_{j=0}^{N} \left(1 - \left(\frac{j}{N}\right)^n\right) \approx N + 1 - N \int_0^1 x^n\,dx = N + 1 - \frac{N}{n+1} \approx \frac{n}{n+1} N.$$

$$\mathbb{P}\{K = k\} = F_K(k+1) - F_K(k) = \left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n,$$

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

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

□

Задача 53. Компоненты случайного вектора имеют нормальное распределение. Следует ли из этого, что вектор имеет нормальное распределение?

Решение. Нет. Контрпример: пусть вектор (X, Y) имеет плотность распределения

$$f(x, y) = \begin{cases} \frac{1}{\pi} e^{-\frac{x^2+y^2}{2}}, & \text{если } xy \neq 0, \\ 0, & \text{в прочих случаях }, \end{cases}$$

т.е. не является нормальным. Однако его компоненты X и Y нормальны:

$$f(x) = \int\limits_{-\infty}^{\infty} f(x,y)\,dy = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \quad \text{и} \quad f(y) = \int\limits_{-\infty}^{\infty} f(x,y)\,dx = \frac{1}{\sqrt{2\pi}}e^{-y^2/2}.$$

$\square$

Задача 54. Пусть случайная величина X_n принимает значения 2^n и -2^n с вероятностями $1/2$. Выполняется ли для последовательности независимых случайных величин X_1, X_2, … закон больших чисел?

Решение. Пусть $\varepsilon < 1$.

$$\mathbb{P}\left\{\frac{\left|\sum\limits_{i=1}^{n} X_i\right|}{n} > \varepsilon\right\} \geqslant \mathbb{P}\left\{\frac{\left|\sum\limits_{i=1}^{n-2} X_i + 2^{n-1} + 2^n\right|}{n} > \varepsilon\right\} \cdot \mathbb{P}\{X_{n-1} = 2^{n-1};\ X_n = 2^n\} =$$

$$= \frac{1}{4}\mathbb{P}\left\{\frac{\left|\sum\limits_{i=1}^{n-2} X_i + 2^{n-1} + 2^n\right|}{n} > \varepsilon\right\} \geqslant \frac{1}{4}\mathbb{P}\left\{\frac{-\left|\sum\limits_{i=1}^{n-2} X_i\right| + 2^{n-1} + 2^n}{n} > \varepsilon\right\} = \frac{1}{4},$$

поскольку

$$-\left|\sum_{i=1}^{n-2} X_i\right| + 2^{n-1} + 2^n \geqslant -\sum_{i=1}^{n-2} 2^i + 2^{n-1} + 2^n > 2^n > n.$$

То есть ЗБЧ не выполняется. $\square$

Задача 55. Пусть случайная величина X_n принимает значения n, 0 и $-n$ с вероятностями $1/4$, $1/2$, $1/4$. Выполняется ли для последовательности независимых случайных величин X_1, X_2, … закон больших чисел?

Решение. Имеем $\mathbb{E}X_n = 0$. Рассмотрим характеристическую функцию сл.в. $S_n = \frac{X_1+\ldots+X_n}{n}$ при $|t| < 1$:

$$\varphi_{X_k}(t) = e^{itX_k} = \cos^2\frac{kt}{2} \;\Rightarrow\; \varphi_{S_n}(t) = \varphi_{X_1}(t/n)\ldots\varphi_{X_n}(t/n) = \left[\cos\frac{t}{2n}\cos\frac{t}{n}\ldots\cos\frac{t}{2}\right]^2.$$

Отсюда (при четном n)

$$\varphi_{S_n}(t) < 1\cdot\ldots\cdot 1\cdot\left[\cos\frac{t}{4}\cos\frac{t(n+2)}{4n}\ldots\cos\frac{t}{2}\right]^2 < \left|\cos\frac{t}{4}\right|^n \xrightarrow{n\to\infty} 0,$$

т.к. $\left|\cos\frac{t}{4}\right| < 1$.

При этом выполнение ЗБЧ означало бы, что

$$S_n \xrightarrow{p} 0 \Leftrightarrow S_n \xrightarrow{D} 0 \Leftrightarrow \varphi_{S_n}(t) \xrightarrow{n\to\infty} 1.$$

Поэтому ЗБЧ не выполняется. $\square$

Задача 56. Пусть $\{X_n\}_{n=1}^{\infty}$ — последовательность независимых случайных величин, причем X_n принимает значения $-\sqrt{n}$, $\sqrt{n}$ с вероятностями $1/2$ каждое. Выполняется для этой последовательности закон больших чисел?

Решение. $\mathbb{E}X_n = 0$, $\operatorname{Var} X_n = n$, т.е. достаточные условия для существования ЗБЧ в форме Чебышева ($\operatorname{Var} X_n \leqslant c < \infty$) не выполнены и необходимо непосредственно проверить выполнение ЗБЧ. Применим для этого метод характеристических функций.

$$\varphi_{X_n}(t) = \frac{e^{-it\sqrt{n}} + e^{it\sqrt{n}}}{2} = \cos t\sqrt{n}; \quad S_n = \frac{\sum\limits_{i=1}^{n} X_i}{n}; \quad \varphi_{S_n}(t) = \cos\frac{t}{n}\cos\frac{t\sqrt{2}}{n}\ldots\cos\frac{t\sqrt{n}}{n}.$$

Выполнение ЗБЧ означает:

$$S_n \xrightarrow{p} 0 \Leftrightarrow S_n \xrightarrow{D} 0 \Leftrightarrow \varphi_{S_n}(t) \xrightarrow{n\to\infty} 1.$$

Последнее соотношение не выполняется: пусть $0 < t < \pi/2$ ($0 < \cos(t/n) < 1$) и n четно ($n = 2k$); тогда

$$\varphi_{S_n}(t) = \cos\frac{t}{2k}\ldots\cos\frac{t\sqrt{k}}{2k}\ldots\cos\frac{t\sqrt{2k}}{2k} < 1\cdot\ldots\cdot 1\cdot\Big(\cos\frac{t\sqrt{k}}{2k}\Big)^{k+1} < \Big(\cos\frac{t\sqrt{k}}{2k}\Big)^{k} =$$

$$= \Big(1 - \frac{t^2}{4k} + \overline{o}\Big(\frac{1}{k}\Big)\Big)^k \xrightarrow{k\to\infty} e^{-t^2/4} < 1.$$

То есть ЗБЧ не выполняется. $\square$

Задача 57. При каких значениях $\alpha > 0$ к последовательности независимых случайных величин $\{X_n\}_{n=1}^{\infty}$, таких что $\mathbb{P}\{X_n = n^\alpha\} = \mathbb{P}\{X_n = -n^\alpha\} = 1/2$, применим закон больших чисел?

Решение. В предыдущей задаче показано, что при $\alpha = \frac{1}{2}$ ЗБЧ не выполняется. Положим теперь $\alpha = \frac{1}{2} - \varepsilon$. Тогда $\operatorname{Var} S_n = \frac{\sum_{i=1}^{n} i^{2\alpha}}{n^2} = \frac{\sum_{i=1}^{n} i^{1-2\varepsilon}}{n^2} \xrightarrow{n\to\infty} 0$, что приводит (ввиду неравенства Чебышева) к выполнению ЗБЧ (заметим, что при этом условие равномерной ограниченности дисперсии X_n ($\operatorname{Var} X_n = n^{2\alpha}$), являясь достаточным, может не выполняться). $\square$

Задача 58. Пусть $\{X_n\}_{n=1}^{\infty}$ — последовательность случайных величин с дисперсиями σ_i^2. Доказать, что если все корреляционные моменты (ковариации) R_{ij} случайных величин X_i и X_j неположительны и при $n \to \infty$ $\frac{\sum_{i=1}^{n}\sigma_i^2}{n^2} \to 0$, то для последовательности $\{X_n\}_{n=1}^{\infty}$ выполняется закон больших чисел.

Решение.

$$\operatorname{Var} S_n = \frac{\operatorname{Var}\bigl(\sum\limits_{i=1}^{n} X_i\bigr)}{n^2} = \frac{\mathbb{E}\bigl(\sum\limits_{i=1}^{n} \overset{\circ}{X}_i\bigr)^2}{n^2} = \frac{\sum\limits_{i=1}^{n} \sigma_i^2 + 2\sum\limits_{i<j} R_{ij}}{n^2} \leqslant \frac{\sum\limits_{i=1}^{n} \sigma_i^2}{n^2} \xrightarrow{n\to\infty} 0,$$

что приводит (ввиду неравенства Чебышева) к выполнению ЗБЧ. □

Задача 59. Пусть $\{X_n\}_{n=1}^{\infty}$ — последовательность случайных величин с равномерно ограниченными дисперсиями, причем каждая случайная величина X_n зависит только от X_{n-1} и X_{n+1}, но не зависит от остальных X_i. Доказать выполнение для этой последовательности закона больших чисел.

Решение.

$$\forall i,\ \operatorname{Var} S_n = \frac{\mathbb{E}\bigl(\sum\limits_{i=1}^{n} \overset{\circ}{X}_i\bigr)^2}{n^2} = \frac{\sum\limits_{i=1}^{n} \operatorname{Var} X_i + 2\sum\limits_{i<j} R_{ij}}{n^2} \leqslant \frac{3c}{n} \xrightarrow{n\to\infty} 0,$$

что приводит (в виду неравенства Чебышева) к выполнению ЗБЧ. □

Задача 60. Книга объемом 500 страниц содержит 50 опечаток. Оценить вероятность того, что на случайно выбранной странице имеется не менее трех опечаток. (Использовать нормальное и пуассоновское приближения, сравнить результаты).

Решение. X — число ошибок на случайно выбранной странице, с биномиальным распределением $X \in \mathrm{Bi}(n,p)$, $n = 50$, $p = 1/500$; точный расчет: $\mathbb{P}\{X \geqslant 3\} \approx 154 \cdot 10^{-6}$; нормальная аппроксимация: $X \in \mathcal{N}(m, \sigma^2)$, $m = np = 0.1$, $\sigma^2 = np(1-p) \approx 0.1$, $\mathbb{P}\{X \geqslant 3\} < 10^{-6}$; пуассоновская аппроксимация: $X \in \mathrm{Po}(\lambda)$, $\lambda = np = 0.1$, $\mathbb{P}\{X \geqslant 3\} \approx 155 \cdot 10^{-6}$. Пуассоновская аппроксимация предпочтительней. □

Задача 61. В тесто для выпечки булок с изюмом замешано N изюмин. Всего из данного теста выпечено K булок. Оценить вероятность того, что в случайно выбранной булке число изюмин находится в пределах от a до b.

Решение. Решаем аналогично предыдущей задаче. X — число изюмин в случайно выбранной булке, с биномиальным распределением $X \in \mathrm{Bi}\bigl(N, \frac{1}{K}\bigr)$; точный расчет:

$$\mathbb{P}\{a \leqslant X \leqslant b\} = \sum_{i=a}^{b} C_N^i \frac{1}{K^i}\Bigl(1 - \frac{1}{K}\Bigr)^{N-i} = \frac{1}{K^N}\sum_{i=a}^{b} C_N^i (K-1)^{N-i}.$$

□

Задача 62. Юноша собирается сыграть три теннисных матча со своими родителями, и он должен победить два раза подряд. Порядок матчей может быть следующим: отец–мать–отец, мать–отец–мать. Юноше нужно решить, какой порядок для него предпочтительней, учитывая, что отец играет лучше матери.

Решение. Введем следующие величины:

- p — вероятность того, что юноша обыграет отца в одном матче;
- q — вероятность того, что юноша обыграет мать в одном матче.

По условию задачи $p < q$. Найдем вероятности того, что юноша победит два раза подряд:

- с порядком матчей отец–мать–отец:

$$p_1 = \mathbb{P}(\text{«победа в первых двух матчах»}) + \mathbb{P}(\text{«победа в последних двух матчах»}) =$$
$$= p \cdot q + (1-p) \cdot qp = 2pq - p^2q;$$

- с порядком матчей мать–отец–мать:

$$p_2 = \mathbb{P}(\text{«победа в первых двух матчах»}) + \mathbb{P}(\text{«победа в последних двух матчах»}) =$$
$$= q \cdot p + (1-q) \cdot pq = 2pq - pq^2.$$

Условие $p < q$ дает соотношение $p_1 > p_2$, поэтому для юноши предпочтительней порядок матчей отец–мать–отец.

□

Задача 63. Пусть X и Y — независимые случайные величины, равномерно распределенные на $(-b, b)$. Найдите вероятность q_b того, что уравнение $t^2 + tX + Y = 0$ имеет действительные корни. Доказать, что существует $\lim\limits_{b\to\infty} q_b = q$. Найдите q.

Решение. Заданное в условии квадратное уравнение имеет действительные корни тогда и только тогда, когда $D = X^2 - 4Y \geqslant 0$, т.е.

$$q_b = \mathbb{P}(X^2 - 4Y \geqslant 0).$$

Функция распределения случайной величины $\xi_1 = X^2$:

$$F_{\xi_1}(x) = \mathbb{P}(X^2 < x) = \mathbb{P}(-\sqrt{x} < X < \sqrt{x}) = \frac{\sqrt{x}}{b}, \quad x \in [0, b^2];$$

случайной величины $\xi_2 = 4Y$:

$$F_{\xi_2}(y) = \mathbb{P}(4Y < y) = \mathbb{P}(Y < \frac{y}{4}) = \frac{b + y/4}{2b} = \frac{y + 4b}{8b}, \quad y \in [-4b, 4b].$$

Соответственно находим плотности распределений:

$$p_{\xi_1}(x) = \frac{1}{2b\sqrt{x}}, \quad x \in [0, b^2]; \quad p_{\xi_2}(y) = \frac{1}{8b}, \quad y \in [-4b, 4b].$$

1) $0 < b < 4$. Имеем

$$q_b = \mathbb{P}(\xi_1 - \xi_2 \geqslant 0) = \int\limits_{x-y\geqslant 0} p_{\xi_1}(x)\cdot p_{\xi_2}(y)\,dx\,dy =$$

$$= \int\limits_{-4b}^{0} p_{\xi_2}(y)\,dy \int\limits_{0}^{b^2} p_{\xi_1}(x)\,dx + \int\limits_{0}^{b^2} p_{\xi_2}(y)\,dy \int\limits_{y}^{b^2} p_{\xi_1}(x)\,dx =$$

$$= \frac{1}{2} + \int\limits_{0}^{b^2} \frac{1}{8b^2}(b-\sqrt{y})\,dy = \frac{1}{2} + \frac{b}{24}.$$

2) $b \geqslant 4$. Имеем

$$q_b = \int\limits_{x-y\geqslant 0} p_{\xi_1}(x)\cdot p_{\xi_2}(y)\,dx\,dy =$$

$$= \int\limits_{-4b}^{0} p_{\xi_2}(y)\,dy \int\limits_{0}^{b^2} p_{\xi_1}(x)\,dx + \int\limits_{0}^{4b} p_{\xi_2}(y)\,dy \int\limits_{y}^{b^2} p_{\xi_1}(x)\,dx =$$

$$= \frac{1}{2} + \int\limits_{0}^{4b} \frac{1}{8b^2}(b-\sqrt{y})\,dy = 1 - \frac{2}{3\sqrt{b}}.$$

Наконец, получим

$$\lim_{b\to\infty} q_b = \lim_{b\to\infty}\left(1 - \frac{2}{3\sqrt{b}}\right) = 1.$$

□

Задача 64. Найдите вероятность q_n того, что случайная $(0,1)$-матрица размера $n\times n$ является невырожденной над полем $GF_2 = \{0,1\}$. Доказать, что существует $\lim\limits_{n\to\infty} q_n = q > 0$.

Решение. Количество всевозможных матриц размера $n\times n$ равно 2^{n^2}. Рассмотрим невырожденную матрицу A размера $n\times n$.

1) Число всевозможных первых строк a_1 матрицы A равно $2^n - 1$ — любая строка из $\{0,1\}^n$, кроме нулевой.

2) Число всевозможных вторых строк a_2 матрицы A равно $2^n - 2$ — любая строка из $\{0,1\}^n$, кроме нулевой и строки a_1.

3) Число всевозможных третьих строк a_3 матрицы A равно $2^n - 4$ — любая строка из $\{0,1\}^n$, кроме нулевой и строки a_1, a_2 и $a_1 + a_2$.

k) Число всевозможных k-х строк a_k матрицы A равно $2^n - 2^{k-1}$ — любая строка из $\{0,1\}^n$, кроме любой линейной комбинации первых $k-1$ строк $a_1, a_2, \ldots, a_{k-1}$ (такая комбинация определяется набором коэффициентов $(c_1, \ldots, c_{k-1}) \in \{0,1\}^{k-1}$).

Получим число всевозможных невырожденных $(0,1)$-матриц размера $n \times n$:

$$K_n = a_1 \cdot \ldots \cdot a_n = (2^n - 1) \cdot (2^n - 2) \cdot \ldots \cdot (2^n - 2^{n-1}),$$

отсюда

$$q_n = \frac{K_n}{2^{n^2}} = \left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{2^2}\right) \cdot \ldots \cdot \left(1 - \frac{1}{2^n}\right).$$

Данное бесконечное произведение сходится: $\lim\limits_{n\to\infty} q_n = q > 0$, поскольку сходится соответствующий ряд

$$\sum_{i=1}^{\infty} \ln\left(1 - \frac{1}{2^i}\right),$$

так как его члены при $n \to \infty$ эквивалентны бесконечно убывающей геометрической прогрессии $\sum\limits_{i=1}^{\infty} \frac{1}{2^i} = 1$.

$\square$

Задача 65. Известно, что 90% выпускаемой продукции соответствует стандарту. Упрощенная схема контроля признает годным с вероятностью $0,88$ каждый стандартный экземпляр аппаратуры и с вероятностью $0,05$ — каждый нестандартной экземпляр аппаратуры. Найдите вероятность, что изделие, прошедшее контроль, соответствует стандарту.

Решение. Для простоты введем для каждого изделия случайные величины С («стандарт») и К («контроль»), принимающие значения 0 и 1. По условию задачи

$$\mathbb{P}(\mathrm{C} = 1) = 0,9 \quad \Rightarrow \quad \mathbb{P}(\mathrm{C} = 0) = 0,1$$

$$\mathbb{P}(\mathrm{K} = 1 | \mathrm{C} = 1) = 0,88, \quad \mathbb{P}(\mathrm{K} = 1 | \mathrm{C} = 0) = 0,05$$

Требуется найти

$$\mathbb{P}(\mathrm{C} = 1 | \mathrm{K} = 1) = \frac{\mathbb{P}(\mathrm{C} = 1, \mathrm{K} = 1)}{\mathbb{P}(\mathrm{K} = 1)} =$$

$$= \frac{\mathbb{P}(\mathrm{K} = 1 | \mathrm{C} = 1) \cdot \mathbb{P}(\mathrm{C} = 1)}{\mathbb{P}(\mathrm{K} = 1 | \mathrm{C} = 1) \cdot \mathbb{P}(\mathrm{C} = 1) + \mathbb{P}(\mathrm{K} = 1 | \mathrm{C} = 0) \cdot \mathbb{P}(\mathrm{C} = 0)} =$$

$$= \frac{0,88 \cdot 0,9}{0,88 \cdot 0,9 + 0,05 \cdot 0,1} \approx 0,9937.$$

$\square$

Задача 66. Вероятности конъюнкций событий A, B и C приведены в таблице истинности:

A	B	C	$\mathbb{P}(ABC)$
0	0	0	20/36
0	0	1	5/36
0	1	0	5/36
0	1	1	0
1	0	0	5/36
1	0	1	0
1	1	0	0
1	1	1	1/36

Можно ли наблюдаемые события A и B использовать как признаки, используемые для обнаружения события C?

Решение. Из таблицы находим

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{6}, \quad \mathbb{P}(AB) = \mathbb{P}(BC) = \mathbb{P}(AC) = \frac{1}{36},$$

откуда следует попарная независимость событий A, B и C. Поэтому каждое событие A и B в отдельности не может служить признаком события C.

Однако можно использовать их вместе:

$$\mathbb{P}(C|AB) = \frac{\mathbb{P}(ABC)}{\mathbb{P}(AB)} = 1, \quad \mathbb{P}(C|A\bar{B}) = \frac{\mathbb{P}(A\bar{B}C)}{\mathbb{P}(A\bar{B})} = 0,$$

$$\mathbb{P}(C|\bar{A}B) = \frac{\mathbb{P}(\bar{A}BC)}{\mathbb{P}(\bar{A}B)} = 0, \quad \mathbb{P}(C|\bar{A}\bar{B}) = \frac{\mathbb{P}(\bar{A}\bar{B}C)}{\mathbb{P}(\bar{A}\bar{B})} = \frac{1}{5}.$$

Значит,

- при $\{A = B = 1\}$ событие C произойдет;
- при $\{A = 1, B = 0\}$ или $\{A = 0, B = 1\}$ событие C не произойдет;
- при $\{A = B = 0\}$ вероятность события C повышается: $\mathbb{P}(C|\bar{A}\bar{B}) > \mathbb{P}(C)$.

$\square$

Задача 67. Пусть $(\Omega, \Sigma, \mathbb{P})$ — некоторое вероятностное пространство и A — алгебра подмножеств Ω такая, что $\sigma(A) = \Sigma$ ($\sigma(A)$ — наименьшая σ-алгебра, содержащая алгебру A). Доказать, что

$$\forall \varepsilon > 0,\ B \in \Sigma \quad \exists A_\varepsilon \in A: \quad \mathbb{P}(A_\varepsilon \triangle B) \leqslant \varepsilon.$$

Решение. Рассмотрим совокупность множеств

$$\mathcal{B} = \big\{ B \in \Sigma \,|\, \forall \varepsilon > 0\ \exists A_B \in A:\ \mathbb{P}(A_B \triangle B) \leqslant \varepsilon \big\}.$$

$\mathcal{B}$ является σ-алгеброй. Действительно,

1) $\Omega \in \mathcal{B}$, т.к. $\Omega \in A \quad \Rightarrow \mathbb{P}(\Omega \triangle \Omega) = 0$.

2) $B \in \mathcal{B} \quad \Rightarrow \forall \varepsilon > 0\, \exists A_B \in A: \ \mathbb{P}(A_B \triangle B) \leqslant \varepsilon$

$$\Rightarrow \mathbb{P}(\bar{A}_B \triangle \bar{B}) = \mathbb{P}(A_B \triangle B) \leqslant \varepsilon,$$

где $\bar{B} = \Omega \setminus B$. Отсюда по определению $\bar{B} \in \mathcal{B}$.

3) $\forall B_1, B_2, \ldots \in \mathcal{B}$ для $B = \bigcup\limits_{i=1}^{\infty} B_i$ также выполнено условие:

$$\forall \varepsilon > 0\ \exists A_B \in A: \ \mathbb{P}(A_B \triangle B) \leqslant \varepsilon.$$

Действительно, при фиксированном $\varepsilon > 0$ для каждого $B_k \in \mathcal{B}$ подбираем

$$A_{B_k} \in A: \quad \mathbb{P}(A_{B_k} \triangle B) \leqslant \frac{\varepsilon}{2^k}.$$

Тогда для $A_B = \bigcup\limits_{i=1}^{\infty} A_{B_i}$ имеем

$$\mathbb{P}(A_B \triangle B) \leqslant \mathbb{P}\Big(\bigcup_{i=1}^{\infty} (A_{B_i} \triangle B_i)\Big) \leqslant \sum_{i=1}^{\infty} \mathbb{P}(A_{B_i} \triangle B_i) \leqslant \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon.$$

Отсюда по определению $B = \bigcup\limits_{i=1}^{\infty} B_i \in \mathcal{B}$.

Из данных свойств следует, что совокупность множеств $\mathcal{B}$ является σ-алгеброй. По определению $\Sigma \supseteq \mathcal{B}$, $\mathcal{B}$ содержит алгебру A и является σ-алгеброй $\quad \Rightarrow \mathcal{B} \supseteq \sigma(A)$. По условию задачи $\sigma(A) = \Sigma$. Имеем

$$\Sigma \supseteq \mathcal{B} \supseteq \Sigma \quad \Rightarrow \mathcal{B} = \Sigma.$$

Из определения $\mathcal{B}$ вытекает требуемое утверждение задачи.

$\square$

Задача 68. Двумерный случайный вектор $X = (X_1, X_2)$ имеет следующую функцию плотности распределения:

$$f(x_1, x_2) = \begin{cases} \dfrac{c}{\sqrt{x_1^2 + x_2^2}}, & \text{при } x_1^2 + x_2^2 \leqslant 1, \\ 0, & \text{иначе }. \end{cases}$$

1) Найдите c.

2) Найдите частные и условные распределения его компонент.

3) Являются ли они

   a) стохастически зависимыми;

б) коррелированными?

Решение. 1) Имеем

$$1 = \int\limits_{x_1^2+x_2^2\leqslant 1} \frac{c}{\sqrt{x_1^2+x_2^2}}\,dx_1\,dx_2 = \Big[\text{ замена } \begin{cases} x_1 = \rho\cos\varphi \\ x_2 = \rho\sin\varphi \end{cases}, \ \frac{D(x_1,x_2)}{D(\rho,\varphi)} = \rho\Big] =$$

$$= c\int\limits_0^{2\pi} d\varphi \int\limits_0^1 d\rho = c\cdot 2\pi,$$

откуда $c = \frac{1}{2\pi}$.

2) Ищем плотности частных распределений:

$$p_1(x_1) = \int\limits_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} f(x_1,x_2)\,dx_2 = \frac{2}{2\pi}\int\limits_0^{\sqrt{1-x_1^2}} \frac{dx_2}{\sqrt{x_1^2+x_2^2}} = \frac{1}{\pi}\ln\big(\tfrac{1+\sqrt{1-x_1^2}}{|x_1|}\big), \quad x_1\in[-1,1].$$

Аналогично

$$p_2(x_2) = \int\limits_{-\sqrt{1-x_2^2}}^{\sqrt{1-x_2^2}} f(x_1,x_2)\,dx_1 = \frac{1}{\pi}\ln\big(\tfrac{1+\sqrt{1-x_2^2}}{|x_2|}\big), \quad x_2\in[-1,1].$$

Функция условного распределения

$$F_1(x) = \mathbb{P}(X_1 \leqslant x | X_2 = x_2) = \int\limits_{-\sqrt{1-x_2^2}}^{x} \frac{c}{\sqrt{y^2+x_2^2}}\,dy.$$

Отсюда находим условную плотность:

$$p_1(x|X_2=x_2) = F_1'(x) = \frac{1}{2\pi\sqrt{x^2+x_2^2}}, \quad x\in\big(-\sqrt{1-x_2^2}, \sqrt{1-x_2^2}\big).$$

Аналогично

$$p_2(y|X_1=x_1) = \frac{1}{2\pi\sqrt{x_1^2+y^2}}, \quad y\in\big(-\sqrt{1-x_1^2}, \sqrt{1-x_1^2}\big).$$

3) а) Поскольку

$$p_1(x_1)\cdot p_2(x_2) \neq f(x_1,x_2),$$

компоненты X_1 и X_2 случайного вектора стохастически зависимы.

б) Частные распределения симметричны относительны нуля, поэтому $\mathbb{E}X_1 = \mathbb{E}X_2 = 0$. Так же из-за симметричности распределения $f(x_1, x_2)$

$$\mathbb{E}X_1X_2 = \int\limits_{x_1^2+x_2^2 \leqslant 1} x_1 \cdot x_2 \cdot f(x_1, x_2)\, dx_1\, dx_2 = 0.$$

Поэтому

$$\operatorname{cov}(X_1, X_2) = \mathbb{E}X_1X_2 - \mathbb{E}X_1 \cdot \mathbb{E}X_2 = 0,$$

т.е. компоненты X_1 и X_2 некоррелированы.

$\square$

Задача 69. Предположим, что с.в. $X \in L_2$, это означает $\mathbb{E}X^2 < \infty$. Докажите, что

$$\|X - \mathbb{E}(X|Y_1, \ldots, Y_n)\|_{L_2} = \min_{\varphi \in H} \|X - \varphi(Y_1, \ldots, Y_n)\|_{L_2}, \tag{1.1}$$

где H — подпространство пространства L_2 всевозможных борелевских функций $\varphi(Y_1, \ldots, Y_n) \in L_2$; $\mathbb{E}(X|Y_1, \ldots, Y_n)$ — условное математическое ожидание с.в. X относительно σ-алгебры, порожденной с.в. $Y_1, \ldots, Y_n$, часто говорят просто относительно с.в. $Y_1, \ldots, Y_n$;

$$\|X\|_{L_2} = \sqrt{\langle X, X\rangle_{L_2}} = \sqrt{\mathbb{E}(X \cdot X)} = \sqrt{\mathbb{E}(X^2)}.$$

Решение. Через $\mathcal{A}$ обозначим σ-алгебру, порожденную с.в. $Y_1, \ldots, Y_n$, кратко будем писать

$$\mathbb{E}^{\mathcal{A}}X = \mathbb{E}(X|Y_1, \ldots, Y_n).$$

Рассмотрим с.в. $\xi \in H$, т.е. $\xi = \varphi(Y_1, \ldots, Y_n)$ для некоторого φ. Тогда

$$\langle \xi, X - \mathbb{E}^{\mathcal{A}}X\rangle = \langle \xi, X\rangle - \langle \xi, \mathbb{E}^{\mathcal{A}}X\rangle =$$

$$= \mathbb{E}(\xi \cdot X) - \mathbb{E}(\xi \cdot \mathbb{E}^{\mathcal{A}}X) = \mathbb{E}(\xi \cdot X) - \mathbb{E}(\mathbb{E}^{\mathcal{A}}(\xi \cdot X)) = \mathbb{E}(\xi \cdot X) - \mathbb{E}(\xi \cdot X) = 0,$$

где было использовано свойство:

$$\eta \in H \quad \Rightarrow \quad \mathbb{E}^{\mathcal{A}}(\xi\eta) = \eta\mathbb{E}^{\mathcal{A}}\xi.$$

Значит, $X - \mathbb{E}^{\mathcal{A}}X \perp \xi, \quad \forall \xi \in H$, т.е. $\mathbb{E}^{\mathcal{A}}$ является проектором на подпространство H в L_2. Поэтому

$$\|X - \varphi(Y_1, \ldots, Y_n)\|_{L_2}^2 = \|X - \xi\|^2 = \|(X - \mathbb{E}^{\mathcal{A}}X) + (\mathbb{E}^{\mathcal{A}}X - \xi)\|^2 =$$

$$= \|X - \mathbb{E}^{\mathcal{A}}X\|^2 + \|\mathbb{E}^{\mathcal{A}}X - \xi\|^2 \geqslant \|X - \mathbb{E}^{\mathcal{A}}X\|^2.$$

Отсюда и получаем, где достигается минимум в (1.1).

$\square$

Задача 70. Докажите, что если в условиях предыдущей задачи $(X, Y_1, \ldots, Y_n)^T$ — является нормальным случайным вектором (без ограничения общности можно также считать, что $(Y_1, \ldots, Y_n)^T$ — невырожденный нормальный случайный вектор), то в качестве H можно взять подпространство всевозможных линейных комбинаций с.в. $Y_1, \ldots, Y_n$. Т.е. мы можем более конкретно сказать, на каком именно классе борелевских функций достигается минимум в (1.1).

Решение. Без ограничения общности считаем, что все с.в. $X, Y_1, \ldots, Y_n$ имеют нулевое математическое ожидание. Будем искать $\mathbb{E}(X|Y_1, \ldots, Y_n)$ в виде

$$\mathbb{E}(X|Y_1, \ldots, Y_n) = c_1 Y_1 + \ldots + c_n Y_n. \tag{1.2}$$

Вектор коэффициентов $(c_1, \ldots, c_n)^T$ однозначно определяется из системы

$$\langle X - c_1 Y_1 - \ldots - c_n Y_n, Y_k \rangle_{L_2} = \mathbb{E}((X - c_1 Y_1 - \ldots - c_n Y_n) \cdot Y_k) = 0, \quad k = 1, \ldots, n,$$

так как матрица данной линейной системы является ковариационной матрицей случайного вектора $(Y_1, \ldots, Y_n)^T$, которая невырождена.

Рассмотрим случайный вектор $\xi = (X - c_1 Y_1 - \ldots - c_n Y_n, Y_1, \ldots, Y_n)^T$, который является линейным преобразованием от гауссовского вектора $(X, Y_1, \ldots, Y_n)^T$, и, следовательно, тоже является гауссовским. Ковариационная матрица случайного вектора ξ блочно-диагональна, т.к. $\mathbb{E}((X - c_1 Y_1 - \ldots - c_n Y_n) \cdot Y_k) = 0$, поэтому из вида характеристических функций для гауссовских векторов получим, что случайные величины

$$X - c_1 Y_1 - \ldots - c_n Y_n, Y_1, \ldots, Y_n$$

независимы.

Поэтому для всякой борелевской функции φ независимы с.в. $X - c_1 Y_1 - \ldots - c_n Y_n$ и $\varphi(Y_1, \ldots, Y_n)$. откуда

$$\mathbb{E}\big((X - c_1 Y_1 - \ldots - c_n Y_n) \cdot \varphi(Y_1, \ldots, Y_n)\big) = \mathbb{E}(X - c_1 Y_1 - \ldots - c_n Y_n) \cdot \mathbb{E}\varphi(Y_1, \ldots, Y_n) = 0.$$

Значит, $X - c_1 Y_1 - \ldots - c_n Y_n$ ортогонален подпространству H пространства L_2 всевозможных борелевских функций $\varphi(Y_1, \ldots, Y_n) \in L_2$. Следовательно, $c_1 Y_1 + \ldots + c_n Y_n$ является минимизирующим элементом в (1.1), который единственен в H и совпадает с условным математическим ожиданием от с.в. X, т.е. выполнено (1.2).

$\square$

Задача 71.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in N\left(\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \begin{Vmatrix} 5 & 2 & 7 \\ 2 & 5 & 7 \\ 7 & 7 & 14 \end{Vmatrix}\right).$$

а) Найдите распределение случайной величины $y_1 = x_1 + x_2 - x_3$.

б) Найдите распределение случайной величины $y_2 = x_1 + x_2 + x_3$.

в) Найдите $\mathbb{E}(y_2 \,|\, x_1 = 5, x_2 = 3)$.

г) Найдите $\mathbb{E}(y_2 \mid x_1 = 5, x_2 < 3)$.

д) Найдите $\mathbb{P}(y_2 < 10 \mid x_1 = 5, x_2 < 3)$.

Решение. а) $y_1 = C_1 x$, где $C_1 = (1,\ 1,\ -1)$. Поэтому

$$\mathbb{E}y_1 = C_1 \mathbb{E}x = 4,$$

$$\operatorname{Var} y_1 = C_1 (\operatorname{Var} x)\, C_1^T = 0.$$

Отсюда получим, что $y_1 \equiv 4$ — вырожденная случайная величина.

б) $y_2 = C_2 x$, где $C_2 = (1,\ 1,\ 1)$. Поэтому

$$\mathbb{E}y_2 = C_2 \mathbb{E}x = 6,$$

$$\operatorname{Var} y_2 = C_2 (\operatorname{Var} x)\, C_2^T = 56.$$

y_2 является гауссовской случайной величиной(как линейное преобразование гауссовского вектора x), поэтому $y_2 \in N(6, 56)$.

в) Из п.а) получим $x_1 + x_2 - x_3 \equiv 4$, поэтому $y_2 = 2(x_1 + x_2) - 4$. Имеем

$$\mathbb{E}(y_2 \mid x_1 = 5, x_2 = 3) = \mathbb{E}(2(x_1 + x_2) \mid x_1 = 5, x_2 = 3) - 4 = 2 \cdot (5 + 3) - 4 = 12.$$

г)

$$\mathbb{E}(y_2 \mid x_1 = 5, x_2 < 3) = \mathbb{E}(2(x_1 + x_2) \mid x_1 = 5, x_2 < 3) - 4 = 2\mathbb{E}(x_2 \mid x_2 < 3) + 6 =$$

$$= \Big[x_2 = y + 3,\ y \in N(0, 5)\Big] = 2\mathbb{E}(y \mid y < 0) + 12 = 12 + 4 \int\limits_{-\infty}^{0} \frac{y\, e^{-\frac{y^2}{10}}}{\sqrt{10\pi}}\, dy =$$

$$= \Big[u = -\tfrac{y^2}{10}\Big] = 12 - \tfrac{20}{\sqrt{10\pi}} \int\limits_{-\infty}^{0} e^u\, du = 12 - 2\sqrt{\tfrac{10}{\pi}}.$$

д)

$$\mathbb{P}(y_2 < 10 \mid x_1 = 5, x_2 < 3) = \mathbb{P}(2(x_1 + x_2) - 4 < 10 \mid x_1 = 5, x_2 < 3) = \mathbb{P}(x_2 < 2 \mid x_2 < 3) =$$

$$= \frac{\mathbb{P}(x_2 < 2)}{\mathbb{P}(x_2 < 3)} = \Big[\mathbb{P}(x_2 < 3) = \frac{1}{2}\Big] = 2 \int\limits_{-\infty}^{-1} \frac{e^{-\frac{y^2}{10}}}{\sqrt{10\pi}}\, dy =$$

$$= \Big[u = \tfrac{y}{\sqrt{5}}\Big] = 2 \int\limits_{-\infty}^{-1/\sqrt{5}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\, du = 2\Phi(-\tfrac{1}{\sqrt{5}}),$$

где $\Phi(x) = \int\limits_{-\infty}^{x} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\, dt$.

□

Задача 72. Может ли функция $\varphi(t) = \begin{cases} 1, & t \in [-T, T] \\ 0, & t \notin [-T, T] \end{cases}$ — быть характеристической функцией некоторой с.в.? Изменится ли ответ, если «чуть-чуть» размазать (сгладить) разрывы функции $\varphi(t)$ в точках $t = \pm T$?

Решение. 1. Пусть $\varphi(t)$ является характеристической функцией, т.е. для некоторой с.в. ξ, $\varphi(t) = \mathbb{E}e^{it\xi}$.

$\varphi(t)$ непрерывна в нуле, поэтому в каждой точке t

$$|\varphi(t+h) - \varphi(t)| = |\mathbb{E}e^{it\xi}(e^{ih\xi} - 1)| \leqslant \mathbb{E}|(e^{ih\xi} - 1)| \xrightarrow{h \to 0} 0.$$

Значит, функция $\varphi(t)$ должна быть непрерывна в каждой точке, что неверно, поскольку имеются разрывы в точках $t = \pm T$.

Поэтому $\varphi(t)$ не может быть характеристической функцией некоторой с.в.

2. Пусть $\hat{\varphi}(t)$ — сглаженная в окрестностях точек $t = \pm T$ функция, на остальных промежутках совпадающая с $\varphi(t)$, и пусть она является характеристической функцией некоторой с.в. Тогда, т.к. $\int_{-\infty}^{+\infty} |\hat{\varphi}(t)|\, dt < \infty$, по формуле обращения для характеристических функций плотность распределения соответствующей с.в. равна

$$f(x) = \tfrac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \hat{\varphi}(t)\, dt.$$

Но $\int_{-\infty}^{+\infty} e^{-itx} \varphi(t)\, dt = 2\frac{\sin Tx}{x}$ — знакопеременная функция, поэтому для «чуть-чуть» сглаженной функции $\hat{\varphi}(t)$ плотность распределения с.в. $f(x)$ также оказывалась бы знакопеременной функцией, что невозможно.

Поэтому при дополнительном сглаживании $\varphi(t)$ вновь не может быть характеристической функцией некоторой с.в.

□

Задача 73. Докажите, что при $n \to \infty$

$$X_n \xrightarrow{L_2} X \Rightarrow X_n \xrightarrow{L_1} X \Rightarrow X_n \xrightarrow{P} X \Leftarrow X_n \xrightarrow{\text{п.н.}} X,$$

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X.$$

С помощью контрпримеров покажите, что никакие другие стрелки импликации в эту схему в общем случае добавить нельзя. При каких дополнительных условиях можно утверждать, что

$$X_n \xrightarrow{\text{п.н.}} X \Rightarrow X_n \xrightarrow{L_1} X?$$

Кроме того, показать, что

$$X_n \xrightarrow{P} X \ (n \to \infty) \Leftrightarrow \rho_P(X_n, X) = \mathbb{E}\Big(\frac{|X_n - X|}{1 + |X_n - X|}\Big) \xrightarrow{n \to \infty} 0.$$

Также показать, что

$$X_n \xrightarrow{d} c \quad \Rightarrow \quad X_n \xrightarrow{P} c, \text{ где } c = \text{const} \ \ (\text{не с.в.})$$

Решение. 1. $X_n \xrightarrow{\text{п.н.}} X \Rightarrow X_n \xrightarrow{P} X$.

Это следует из теоремы из функционального анализа: из сходимости почти наверное последовательности измеримых функций $\{X_n\}$ следует сходимость по мере, в частности, по вероятностной мере.

2. $X_n \xrightarrow{L_1} X \Rightarrow X_n \xrightarrow{P} X$.

Это следует из неравенства Маркова:

$$\mathbb{P}(|X_n - X| > \varepsilon) \leqslant \frac{\mathbb{E}|X_n - X|}{\varepsilon}.$$

3. $X_n \xrightarrow{L_2} X \Rightarrow X_n \xrightarrow{L_1} X$.

Это следует из неравенства $\big(\mathbb{E}|X_n - X|\big)^2 \leqslant \mathbb{E}|X_n - X|^2$, являющееся неравенством Йенсена вида

$$g(\mathbb{E}Y) \leqslant \mathbb{E}g(Y)$$

для выпуклой функции $g(x) = x^2$ при $Y = |X_n - X|$.

4. $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$.

Как доказывается в курсе теории вероятности, сходимость по распределению эквивалентна слабой сходимости, поэтому достаточно доказать, что для любой ограниченной непрерывной функции $\varphi(\cdot)$ выполняется

$$\lim_{n\to\infty} \mathbb{E}\varphi(X_n) = \mathbb{E}\varphi(X). \tag{1.3}$$

Имеем

$$\forall\varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0.$$

$$|\mathbb{E}\varphi(X_n) - \mathbb{E}\varphi(X)| \leqslant \mathbb{E}|\varphi(X_n) - \varphi(X)| = \int\limits_\Omega |\varphi(X_n(\omega)) - \varphi(X(\omega))|\, P(d\omega) =$$

$$= \int\limits_{|\varphi(X_n(\omega))-\varphi(X(\omega))|>\varepsilon} |\varphi(X_n(\omega))-\varphi(X(\omega))|\, P(d\omega) + \int\limits_{|\varphi(X_n(\omega))-\varphi(X(\omega))|\leqslant\varepsilon} |\varphi(X_n(\omega))-\varphi(X(\omega))|\, P(d\omega).$$

Из ограниченности φ имеем $|\varphi(x)| < C$, $|\varphi(x) - \varphi(y)| < 2C$, поэтому

$$\int\limits_\Omega |\varphi(X_n(\omega)) - \varphi(X(\omega))|\, P(d\omega) < 2C \cdot \mathbb{P}(|X_n - X| > \varepsilon) +$$

$$+ \int\limits_{|\varphi(X_n(\omega))-\varphi(X(\omega))|\leqslant\varepsilon} |\varphi(X_n(\omega)) - \varphi(X(\omega))|\, P(d\omega).$$

Из непрерывности φ следует, что

$$\forall\delta > 0\ \exists\varepsilon > 0 : \quad |X_n(\omega) - X(\omega)| \leqslant \varepsilon\ |\varphi(X_n(\omega)) - \varphi(X(\omega))| < \delta \ \text{ и}$$

$$2C\cdot\mathbb{P}(|X_n-X| > \varepsilon) + \int\limits_{|\varphi(X_n(\omega))-\varphi(X(\omega))|\leqslant\varepsilon} |\varphi(X_n(\omega))-\varphi(X(\omega))|\, P(d\omega) < 2C\cdot\mathbb{P}(|X_n-X| > \varepsilon)+\delta.$$

Из сходимости $X_n \xrightarrow{P} X$ следует, что $\exists N: \ \forall n \geqslant N \ \Rightarrow \ \mathbb{P}(|X_n - X| > \varepsilon) < \frac{\delta}{2C}$, откуда

$$|\mathbb{E}\varphi(X_n) - \mathbb{E}\varphi(X)| < 2\delta.$$

Свойство (1.3) доказано.

5. Из сходимости по распределению не следует сходимость по вероятности.

   Рассматривается вероятностная тройка $(\Omega, \mathcal{F}, \mathbb{P})$, где $\Omega = [0,1]$, $\mathcal{F}$ — σ-алгебра борелевских множеств из Ω, $\mathbb{P}$ — мера Лебега. Случайные величины $X(\omega) = \omega$ и $Y(\omega) = 1 - \omega$ имеют одну и ту же функцию распределения

   $$F_X(x) = F_Y(x) = x \cdot \mathbb{I}_{\{x \in [0,1]\}}.$$

   Поэтому последовательность с.в. $X, Y, X, Y, \ldots$ сходится по распределению, т.к. каждый элемент последовательности имеет одну и ту же функцию распределения. При этом, очевидно, сходимости по вероятности нет, т.к. например,

   $$\mathbb{P}(|X - Y| > 0.5) = 0.5.$$

   Также заметим, что из данного примера следует, что из сходимости по распределению не следует сходимость в L_1, в L_2 и почти наверное, поскольку, как уже доказано, каждая из них влечет сходимость по вероятности.

6. Из сходимости по вероятности не следует сходимость почти наверное.

   Рассматривается та же вероятностная тройка $(\Omega, \mathcal{F}, \mathbb{P})$. Рассмотрим в качестве последовательности с.в. серию бегущих импульсов:

   $$X_1 = \mathbb{I}_{[0,1/2]},\ X_2 = \mathbb{I}_{[1/2,1]},\ X_3 = \mathbb{I}_{[0,1/4]},\ X_4 = \mathbb{I}_{[1/4,1/2]}, \ldots,$$

   в общем виде $n = 2^k - 2 + i \Rightarrow \ X_n = \mathbb{I}_{[\frac{i-1}{2^k}, \frac{i}{2^k}]}$.

   Последовательность $\{X_n\}$ не сходится ни в одной точке ω, т.к. при сколь угодно больших n с.в. X_n равна как 0, так и 1. При этом $X_n \xrightarrow{P} X \equiv 0$, т.к. при малых $\varepsilon > 0$

   $$\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{P}(|X_n| = 1) = \frac{1}{2^{[\log n]}} \xrightarrow{n \to \infty} 0.$$

7. Из сходимости в L_1 не следует сходимость почти наверное.

   Последовательность $\{X_n\}$ из предыдущего пункта сходится к $X \equiv 0$ в L_1:

   $$\mathbb{E}|X_n - X| = \mathbb{E}X_n = \frac{1}{2^{[\log n]}} \xrightarrow{n \to \infty} 0,$$

   но, как отмечалось, не сходится почти наверное.

8. Пример, когда из сходимости почти наверное не следует сходимость в L_1 или в L_2. Из данного примера также следует, что из сходимости по вероятности не следует сходимость в L_1 или в L_2.

   Рассматривается та же вероятностная тройка $(\Omega, \mathcal{F}, \mathbb{P})$. Строим последовательность с.в.

   $$X_n(\omega) = n^2 \cdot \mathbb{I}_{[0, \frac{1}{n}]}.$$

Из данной формулы следует, что последовательность $\{X_n\}$ сходится к нулю почти наверное и по вероятности. Однако

$$\mathbb{E}|X_n| = n^2 \cdot \frac{1}{n} = n \xrightarrow{n\to\infty} \infty,$$
$$\mathbb{E}|X_n|^2 = n^4 \cdot \frac{1}{n} = n^3 \xrightarrow{n\to\infty} \infty.$$

Поэтому последовательность $\{X_n\}$ не сходится в L_1 и в L_2.

9. Выполнена импликация
$$X_n \xrightarrow{\text{п.н.}} X \;\Rightarrow\; X_n \xrightarrow{L_1} X,$$
т.е. возможен предельный переход под знаком математического ожидания, если семейство с.в. $\{X_n\}$ является равномерно интегрируемым:
$$\sup_n \mathbb{E}\big[|X_n| \cdot \mathbb{I}_{\{|X_n|>c\}}\big] \xrightarrow{c\to+\infty} 0.$$

10. Покажем, что
$$X_n \xrightarrow{P} X \ (n\to\infty) \Leftrightarrow \rho_P(X_n, X) = \mathbb{E}\Big(\frac{|X_n - X|}{1+|X_n - X|}\Big) \xrightarrow{n\to\infty} 0.$$

Необходимость. Очевидно, $\frac{|X_n-X|}{1+|X_n-X|} < 1$, $\frac{|X_n-X|}{1+|X_n-X|} \leqslant |X_n - X|$. Зафиксируем произвольное $\varepsilon > 0$. Имеем
$$\mathbb{E}\Big(\frac{|X_n - X|}{1+|X_n - X|}\Big) = \int\limits_{|X_n-X|\leqslant\varepsilon} \frac{|X_n(\omega) - X(\omega)|}{1+|X_n(\omega)-X(\omega)|}\,P(d\omega) +$$
$$+ \int\limits_{|X_n-X|>\varepsilon} \frac{|X_n(\omega) - X(\omega)|}{1+|X_n(\omega)-X(\omega)|}\,P(d\omega) \leqslant \varepsilon + \mathbb{P}(|X_n - X| > \varepsilon).$$

$X_n \xrightarrow{P} X$, поэтому $\mathbb{P}(|X_n - X| > \varepsilon) \xrightarrow{n\to\infty} 0$ и при $n > N_\varepsilon$:
$$\mathbb{E}\Big(\frac{|X_n - X|}{1+|X_n - X|}\Big) \leqslant \varepsilon + \mathbb{P}(|X_n - X| > \varepsilon) \leqslant 2\varepsilon.$$

Это и означает, что
$$\mathbb{E}\Big(\frac{|X_n - X|}{1+|X_n - X|}\Big) \xrightarrow{n\to\infty} 0.$$

Достаточность. Доказываем от противного: пусть нет сходимости последовательности $\{X_n\}$ по вероятности, тогда
$$\exists\varepsilon_0 > 0,\ \delta > 0,\ \text{т.ч. } \forall N,\ \exists n > N:\ \mathbb{P}(|X_n - X| > \varepsilon_0) > \delta.$$

Т.к. функция $g(y) = \frac{y}{1+y}$ монотонно возрастает, для таких n получим
$$\mathbb{E}\Big(\frac{|X_n - X|}{1+|X_n - X|}\Big) \geqslant \int\limits_{|X_n-X|>\varepsilon_0} \frac{|X_n(\omega) - X(\omega)|}{1+|X_n(\omega)-X(\omega)|}\,P(d\omega) > \delta \cdot \frac{\varepsilon_0}{1+\varepsilon_0}.$$

Это противоречит условию $\mathbb{E}\Big(\frac{|X_n-X|}{1+|X_n-X|}\Big) \xrightarrow{n\to\infty} 0$, поэтому исходное предположение неверно, что завершает доказательство.

11. Итак, пусть $X_n \xrightarrow{d} c$, $c = \text{const}$. Это означает, что $\forall \varepsilon > 0$,

$$\mathbb{P}(|X_n - c| < \varepsilon) \xrightarrow{n\to\infty} \mathbb{P}(|c - c| < \varepsilon) = 1.$$

Это по определению и означает, что $X_n \xrightarrow{P} c$ (при $n \to \infty$).

$\square$

Задача 74. Пусть $x_1, \ldots, x_n$ — независимые одинаково распределенные с.в.. Пусть также характеристическая функция с.в. x_k представляется в окрестности $t = 0$ в виде

$$\varphi_{x_k}(t) = \mathbb{E}(e^{itx_k}) = 1 + imt + o(t).$$

Используя то, что

$$S_n \xrightarrow{d} c \quad \Rightarrow \quad S_n \xrightarrow{P} c, \text{ где } c = \text{const} \text{ (не с.в.)}$$

и

$$S_n \xrightarrow{d} S \quad \Leftrightarrow \quad \varphi_{S_n}(t) \to \varphi_S(t), \text{ равномерно по } t \text{ в окрестности } t = 0\ ,$$

найдите

$$S_n = \frac{1}{n} \sum_{i=1}^{n} x_i \xrightarrow{P} ?$$

Решение. Имеем

$$\varphi_{S_n}(t) = \left[\varphi_{x_k}\left(\tfrac{t}{n}\right)\right]^n = \left[1 + \frac{imt}{n} + o\left(\tfrac{1}{n}\right)\right]^n \xrightarrow{n\to\infty} e^{imt}, \text{ равномерно по } t \text{ в окрестности } t = 0\ .$$

Т.к. $\varphi(t) = e^{imt}$ является характеристической функцией для вырожденной с.в. $m = \text{const}$, из соответствия между характеристическими функциями и распределениями получим $S_n \xrightarrow{d} m$.

В таком случае выполнено $S_n \xrightarrow{P} m$, т.е. m — искомый предел по вероятности.

$\square$

Задача 75. Пусть $x_1, x_2, x_3, \ldots$ — последовательность независимых одинаково распределенных с.в.. Положим $S_n = \sum\limits_{k=1}^{n} x_k$. Покажите, что

1) (з.б.ч.) если $\mathbb{E}(|x_k|) < \infty$, то $S_n/n \xrightarrow{P} m$ при $n \to \infty$, где $m = \mathbb{E}(x_k)$;

2) (ц.п.т.) если $\mathbb{E}(x_k^2) < \infty$, то $(S_n - m \cdot n)/\sqrt{n \cdot D} \xrightarrow{d} N(0,1)$ при $n \to \infty$, где $D = \operatorname{Var} x_k$.

3) (задача математической статистики) Предположим, что независимо n раз кидается монетка с вероятностью выпадения орла в каждом опыте равной p (точного значения p мы не знаем, а знаем лишь то, что $0.1 \leqslant p \leqslant 0.9$), т.е. $x_k \in \mathrm{Be}(p)$. Сколько раз нужно кинуть монетку (оцените p), чтобы оценка $\bar{p}(x) = \frac{\sum\limits_{k=1}^{n} x_k}{n}$ с вероятностью $\gamma \geqslant 0.95$ отличалась от истинного значения p не более, чем на величину $\delta = 0.01$? Применить неравенство Чебышева и предельную теорему (точность, которую дает ц.п.т., оцените с помощью неравенства Берри – Эссена). Сравнить результаты.

Решение. 1) Поскольку $\mathbb{E}(|x_k|) < \infty$, из свойств характеристических функций

$$\exists \varphi_{x_k}^{(1)}(t) = i\mathbb{E}x_k = im,$$

$$\varphi_{x_k}(t) = 1 + t\varphi_{x_k}^{(1)}(t) + o(t) = 1 + imt + o(t).$$

Поэтому

$$\varphi_{S_n/n}(t) = \Big[\varphi_{x_k}\big(\tfrac{t}{n}\big)\Big]^n = \Big(1 + \frac{imt}{n} + o\big(\tfrac{1}{n}\big)\Big)^n \xrightarrow{n\to\infty} \varphi(t) = e^{imt}.$$

$\varphi(t)$ является характеристической функцией вырожденной с.в. $m = \mathrm{const}$, поэтому по теореме непрерывности для характеристических функций и распределений

$$S_n/n \xrightarrow{d} m.$$

В данном случае можно утверждать, что $S_n/n \xrightarrow{P} m$. Действительно, имеем $\forall \varepsilon > 0$

$$\mathbb{P}(|S_n/n - m| < \varepsilon) \xrightarrow{n\to\infty} \mathbb{P}(|m - m| < \varepsilon) = 1,$$

что по определению и означает $S_n/n \xrightarrow{P} m$.

2) Из свойств характеристических функций для с.в. $\xi_k = x_k - m$ имеем

$$\varphi_{\xi_k}(t) = 1 - \frac{t^2\sigma^2}{2} + o(t^2),$$

где $\sigma^2 = \mathbb{E}\xi_k^2 = \operatorname{Var} x_k$. Для с.в. $T_n = (S_n - m\cdot n)/\sqrt{n\cdot D}$ имеем

$$\varphi_{T_n}(t) = \Big[\varphi_{\xi_k}\Big(\frac{t}{\sigma\sqrt{n}}\Big)\Big]^n =$$

$$= \Big[1 - \frac{t^2}{2n} + o\Big(\frac{1}{n}\Big)\Big]^n \xrightarrow{n\to\infty} \varphi(t) = e^{-t^2/2}.$$

$\varphi(t)$ является характеристической функцией с.в. с распределением $N(0,1)$. Из теоремы непрерывности для характеристических функций и распределений вытекает ц.п.т.:

$$T_n = (S_n - m\cdot n)/\sqrt{n\cdot D} \xrightarrow{d} N(0,1).$$

3) а) Согласно неравенству Чебышева для с.в. $\bar p(x)$

$$\mathbb{P}\big(|\bar p(x) - p| \leqslant \delta\big) \geqslant 1 - \frac{\operatorname{Var}\bar p(x)}{\delta^2}.$$

Из независимости с.в. $x_1, \dots, x_n$ имеем $\operatorname{Var}\bar p(x) = \frac{\operatorname{Var}x_k}{n} = \frac{p(1-p)}{n}$. Оценка величины $\bar p(x)$ с вероятностью $\gamma \geqslant 0.95$ гарантируется при

$$\gamma \leqslant 1 - \frac{\operatorname{Var}\bar p(x)}{\delta^2} = 1 - \frac{p(1-p)}{n\delta^2}$$

$$\Rightarrow \quad n \geqslant \frac{p(1-p)}{\delta^2(1-\gamma)}.$$

Т.к. величина $p(1-p)$ достигает максимума при $p = 0.5$, данное неравенство заведомо выполнено при числе бросаний монетки

$$n = \frac{0.5^2}{\delta^2(1-\gamma)} = 5 \cdot 10^4.$$

б) Из ц.п.т. и неравенства Берри – Эссена получаем

$$\sup_x \Big|\mathbb{P}\Big(\sum_{k=1}^n \frac{x_k - np}{\sigma\sqrt{n}} < x\Big) - \Phi(x)\Big| \leqslant \frac{C_0\mu^3}{\sigma^3\sqrt{n}}.$$

Здесь $C_0 < 0.7056$, $\sigma^2 = \operatorname{Var}x_k = p(1-p)$, $\Phi(x) = \int\limits_{-\infty}^{x} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt$,

$$\mu^3 = \mathbb{E}|x_k - p|^3 = p(1-p)^3 + (1-p)p^3 = \sigma^2(1 - 2\sigma^2).$$

Т.к. $\sum\limits_{k=1}^n \frac{x_k - np}{\sigma\sqrt{n}} = \frac{\sqrt{n}}{\sigma}\big(\bar p(x) - p\big)$, условие $|\bar p(x) - p| \leqslant \delta$ выполняется с вероятностью не менее

$$\Phi\Big(\frac{\delta\sqrt{n}}{\sigma}\Big) - \Phi\Big(-\frac{\delta\sqrt{n}}{\sigma}\Big) - \frac{2C_0\mu^3}{\sigma^3\sqrt{n}}.$$

Разность $\Phi(\frac{\delta\sqrt{n}}{\sigma}) - \Phi(-\frac{\delta\sqrt{n}}{\sigma})$ принимает наименьшее значение при максимальном $\sigma^2 = p(1-p)$, которое достигается при $p = 0.5$. Кроме того, наибольшее значение величины

$$\frac{\mu^3}{\sigma^3} = \frac{\sigma^2(1-2\sigma^2)}{\sigma^3} = \frac{1}{\sigma} - 2\sigma$$

достигается при наименьшем возможном σ, которое дает $p = 0.1$, откуда $\sigma = 0.3$. Поэтому оценка для $\bar p(x)$ с вероятностью $\gamma \geqslant 0.95$ гарантируется при

$$\gamma \leqslant \Phi(2\delta\sqrt{n}) - \Phi(-2\delta\sqrt{n}) - \frac{2 \cdot 0.7}{\sqrt{n}}\big(1/0.3 - 0.6\big).$$

Например, $\Phi(2\delta\sqrt{n}) - \Phi(-2\delta\sqrt{n}) = 0.985$ при $n = 14971$, при таком n

$$\frac{2 \cdot 0.7}{\sqrt{n}}\big(1/0.3 - 0.6\big) = 0.0315,$$

поэтому $|\bar{p}(x)-p| \leqslant \delta$ выполняется с вероятностью $\gamma \geqslant 0.985-0.0315 = 0.9535 > 0.95$.

Сравнивая результаты для методов неравенства Чебышева и предельной теоремы, получаем, что предельная теорема дает более точные оценки и гарантирует нужную точность оценки для p уже при при числе бросаний монетки $n = 14971$, тогда как неравенство Чебышева — лишь при $n = 50000$.

□

Задача 76. 1. Пусть при каждом $n \geqslant 1$ независимые с.в. $x_{1n}, x_{2n}, \ldots, x_{nn}$ таковы, что $x_{kn} \in \mathrm{Be}(p_{kn})$, где $\max\limits_{1 \leqslant k \leqslant n} p_{kn} \xrightarrow{n\to\infty} 0$, $\sum\limits_{k=1}^{n} p_{kn} \xrightarrow{n\to\infty} \lambda$. Тогда

$$\mathbb{P}(S_n = m) \xrightarrow{n\to\infty} e^{-\lambda}\frac{\lambda^m}{m!}, \quad m = 0, 1, 2, \ldots, \quad \text{где } S_n = \sum_{k=1}^{n} x_{kn}. \tag{1.4}$$

2. В течение дня вы играете в казино и участвуете в $N = 100$ независимых розыгрышах. В каждом розыгрыше вы выигрываете с вероятностью $p = 0.01$. Оцените вероятность того, что вам не удастся ни разу выиграть. Оцените вероятности того, что вы выиграете ровно один раз и ровно три раза.

   Предположим, что вы ходите играть в казино в течении $n = 100$ дней (количество розыгрышей в день и вероятность выиграть не менялись). Оцените вероятность того, что за эти 100 дней вы в общей сложности выиграете не менее 100 раз, не менее 300 раз.

Решение. 1. Учитывая, что $p_{kn} = o(1)$ (при $n \to \infty$), характеристическая функция для суммы S_n независимых с.в. имеет вид

$$\varphi_{S_n}(t) = (1 + p_{1n}(e^{it}-1)) \cdot (1 + p_{2n}(e^{it}-1)) \cdot \ldots \cdot (1 + p_{nn}(e^{it}-1))$$

$$\Rightarrow \quad \ln \varphi_{S_n}(t) = \sum_{k=1}^{n} \ln(1 + p_{kn}(e^{it}-1)) = \sum_{k=1}^{n} (p_{kn} + o(p_{kn}))(e^{it}-1) =$$
$$= (\lambda + o(1))(e^{it}-1).$$

Поэтому

$$\varphi_{S_n}(t) = \exp\bigl\{(\lambda + o(1))(e^{it}-1)\bigr\} \xrightarrow{n\to\infty} \exp\bigl\{\lambda(e^{it}-1)\bigr\}.$$

Из теоремы непрерывности для характеристических функций и распределений следует

$$S_n \xrightarrow{d} \mathrm{Po}(\lambda),$$

откуда по определению получаем (1.4).

Согласно результатам Ю.В.Прохорова(1953), сходимость в (1.4) равномерная и скорость сходимости для случая $p_{kn} = p_n = \dfrac{\lambda}{n}$ описывается формулой

$$\sum_{m=0}^{+\infty} \Bigl|\mathbb{P}(S_n = m) - e^{-\lambda}\frac{\lambda^m}{m!}\Bigr| \leqslant 2\,\frac{\lambda}{n} \cdot \min(2, \lambda).$$

2. Количество выигрышей в течении дня $T_N = \sum\limits_{k=1}^{N} x_k$, где $x_k \in \mathrm{Be}(p)$. Из доказанного свойства распределение T_N близко к $\mathrm{Po}(\lambda)$, где $\lambda = Np = 1$. Поэтому

$$\mathbb{P}(\text{«не удастся ни разу выиграть»}) = \mathbb{P}(T_N = 0) \approx e^{-\lambda} = e^{-1} \approx 0.368,$$

$$\mathbb{P}(\text{«вы выиграете ровно один раз»}) = \mathbb{P}(T_N = 1) \approx e^{-1} \approx 0.368,$$

$$\mathbb{P}(\text{«вы выиграете ровно три раза»}) = \mathbb{P}(T_N = 3) \approx e^{-\lambda}\frac{\lambda^3}{3!} = \frac{e^{-1}}{6} \approx 0.061.$$

Количество выигрышей в течении $n = 100$ дней

$$S_m = \sum_{k=1}^{m} x_k, \quad \text{где } x_k \in \mathrm{Be}(p),\ m = Nn = 10^4.$$

Применим для S_m ц.п.т.:

$$\mathbb{P}\Big(\frac{S_m - \mathbb{E}S_m}{\sqrt{\mathrm{Var}\, S_m}} < x\Big) = \mathbb{P}\Big(\frac{S_m - 100}{100\sqrt{p(1-p)}} < x\Big) \approx \Phi(x).$$

Отсюда $\mathbb{P}\big(S_m < 100 + 100x \cdot \sqrt{p(1-p)}\big) \approx \mathbb{P}\big(S_m < 100 + 10x\big) \approx \Phi(x)$. Имеем

$$\mathbb{P}(\text{«вы выиграете не менее 100 раз»}) \approx 1 - \Phi(0) = 0.5,$$

$$\mathbb{P}(\text{«вы выиграете не менее 300 раз»}) \approx 1 - \Phi(20) \approx 0.$$

□

Задача 77. До проведения схемы испытаний Бернулли разыгрывается с.в. p, имеющая равномерное распределение на отрезке $[0.1, 0.9]$ (результаты розыгрыша нам неизвестны). После того как эта с.в. была разыграна, начинают проводиться опыты по схеме Бернулли (независимо $n = 1000$ раз подкидывается монетка) с вероятность успеха (выпадения «орла») в каждом опыте равной p (после того как с.в. p была разыграна, она уже приняла какое-то значения из отрезка $[0.1, 0.9]$ и рассматривается в серии опытов Бернулли уже как число, причем не меняющееся от опыта к опыту). В результате опыта было посчитано значение числа успехов $r = 777$. Определите апостериорное распределение с.в. p, т.е. найдите условную плотность распределения $p(x|r = 777)$. Оцените, как изменится ответ, если точное значение числа успехов нам неизвестно. Известно только, что $r \in [750, 790]$. Т.е. посчитайте условную плотность вероятности $p(x|r \in [750, 790])$.

Решение. 1) Обозначим $X_k = \mathcal{I}(\text{«выпал орел»}) \in \mathrm{Be}(p)$, $S_n = X_1 + \ldots + X_n$. Условная функция распределения

$$F(x) = \mathbb{P}(p < x | S_n = r) = \frac{\mathbb{P}(p < x, S_n = r)}{\mathbb{P}(S_n = r)} =$$

$$= \frac{1/0.8 \cdot \int\limits_{0.1}^{x} C_n^r y^r (1-y)^{n-r}\, dy}{1/0.8 \cdot \int\limits_{0.1}^{0.9} C_n^r y^r (1-y)^{n-r}\, dy}$$

Применяя формулу Стирлинга для биномиальных коэффициентов с известными значениями $r = 777$, $n = 1000$, обосновывается следующее приближение:

$$\int_{0.1}^{0.9} C_n^r\, y^r (1-y)^{n-r}\, dy \approx \int_{0}^{1} C_n^r\, y^r (1-y)^{n-r}\, dy = C_n^r\, B(r+1, n-r+1) = \frac{1}{n+1},$$

где $B(a,b)$ — бета-функция, откуда имеем

$$F(x) \approx (n+1) \cdot \int_{0.1}^{x} C_n^r\, y^r (1-y)^{n-r}\, dy.$$

Условная плотность равна

$$p(x|r = 777) = F'(x) \approx (n+1) \cdot C_n^r x^r (1-x)^{n-r}.$$

2) Аналогично выписываем условную функцию распределения

$$F(x) = \mathbb{P}(p < x | S_n \in [r_1, r_2]) = \frac{\mathbb{P}(p < x, S_n \in [r_1, r_2])}{\mathbb{P}(S_n \in [r_1, r_2])}$$

При всяком $r \in [r_1, r_2]$ можно считать $\mathbb{P}(S_n = r) \approx \frac{1}{n+1}$, поэтому

$$F(x) \approx (n+1)\, \frac{\sum\limits_{r \in [r_1, r_2]} \int\limits_{0.1}^{x} C_n^r\, y^r (1-y)^{n-r}\, dy}{r_2 - r_1 + 1}.$$

Условная плотность равна

$$p(x|r \in [750, 790]) = F'(x) \approx (n+1)\, \frac{\sum\limits_{r=750}^{790} C_n^r\, x^r (1-x)^{n-r}}{41}.$$

Исходя из ц.п.т. для $S_n = np + \xi\sqrt{np(1-p)}$, где $\xi \in \mathcal{N}(0,1)$, последнюю формулу можно переписать в виде

$$p(x|r \in [r_1, r_2]) = F'(x) \approx \frac{n+1}{r_2 - r_1 + 1}\Big[\Phi\Big(\frac{r_2 - nx}{\sqrt{nx(1-x)}}\Big) - \Phi\Big(\frac{r_1 - nx}{\sqrt{nx(1-x)}}\Big)\Big].$$

□

Задача 78. Показать, что при бросании симметричной монеты n раз отношение числа выпадений герба к числу выпадений решки почти наверное стремится к 1 при $n \to \infty$, а вероятность того, что число выпадений герба в точности равняется числу выпадений решки, при четном числе бросаний, стремится к 0 при $n \to \infty$.

Решение. 1. Введем с.в. $X_k = \mathcal{I}$(«выпал герб»). Из условия задачи следует, что $X_k \in \mathrm{Be}(1/2)$. По теореме Колмогорова об у.з.б.ч. для $S_n = X_1 + \ldots + X_n$ с вероятностью 1 выполнено

$$\frac{S_n}{n} \xrightarrow{n\to\infty} \frac{1}{2}.$$

Поэтому для отношения числа выпадений герба к числу выпадений решки почти наверное выполнено

$$\frac{S_n}{n - S_n} \xrightarrow{n\to\infty} 1.$$

2. При четном числе бросаний $n = 2m$ вероятность того, что число выпадений герба в точности равняется числу выпадений решки, равна

$$p_m = \frac{C_{2m}^m}{2^{2m}} = \frac{(2m)!}{2^{2m} \cdot (m!)^2}.$$

По формуле Стирлинга $m! = (1 + o(1)) \cdot \sqrt{2\pi m}\big(\frac{m}{e}\big)^m$(при $m \to \infty$). Подставив эту формулу в выражение для p_m, получим требуемое свойство:

$$p_m = \frac{1 + o(1)}{\sqrt{\pi m}} \xrightarrow{m\to\infty} 0.$$

□

Задача 79. Пусть с.в. $x_n \in \Gamma(\lambda, n)$. Покажите, что из ц.п.т. следует

$$\frac{x_n - m(\lambda) \cdot n}{\sigma(\lambda) \cdot \sqrt{n}} \xrightarrow{d} N(0,1) \text{ при } n \to \infty.$$

Найдите $m(\lambda)$, $\sigma(\lambda)$.

Решение. Характеристическая функция с.в. $x_n \in \Gamma(\lambda, n)$ имеет вид

$$\varphi_{x_n}(t) = \big(1 - \tfrac{it}{\lambda}\big)^{-n}.$$

Из свойств характеристических функций следует равенство распределению с.в.

$$x_n \text{ и } y_1 + \ldots + y_n,$$

где $y_k \in \Gamma(\lambda, 1)$, $k = 1, \ldots, n$. Из ц.п.т. для $y_1 + \ldots + y_n$ следует

$$\frac{x_n - \mathbb{E}x_n}{\sqrt{\operatorname{Var} x_n}} \xrightarrow{d} N(0,1) \text{ при } n \to \infty.$$

Поскольку $\mathbb{E}x_n = \frac{n}{\lambda}$, $\operatorname{Var} x_n = \frac{n}{\lambda^2}$, получим

$$m(\lambda) = \mathbb{E}y_k = \frac{1}{\lambda}, \quad \sigma(\lambda) = \operatorname{Var} y_k = \frac{1}{\lambda}.$$

□

Задача 80. Пусть $(\Omega, \Xi, \mathbb{P})$ — вероятностное пространство, $\xi_1, \xi_2, \ldots$ — некоторая последовательность с.в.. Обозначим $\Xi_n^\infty = \sigma(\xi_n, \xi_{n+1}, \ldots)$ — σ-алгебру, порожденную с.в. $\xi_n, \xi_{n+1}, \ldots$ и пусть

$$\mathcal{X} = \bigcap_{n=1}^{\infty} \Xi_n^\infty.$$

Поскольку пересечение σ-алгебр есть снова σ-алгебра, то $\mathcal{X}$ — есть σ-алгебра. Эту σ-алгебру будем называть «хвостовой» или «остаточной», в связи с тем, что всякое событие $A \in \mathcal{X}$ не зависит от значений с.в. $\xi_1, \xi_2, \ldots, \xi_n$ при любом конечном n, а определяется лишь «поведением бесконечно далеких значений последовательности $\xi_1, \xi_2, \ldots$ ».

С помощью задачи 67 докажите справедливость следующего утверждения:

Пусть $\xi_1, \xi_2, \ldots$ — последовательность независимых в совокупности с.в. и $A \in \mathcal{X}$ (событие A принадлежит «хвостовой» σ-алгебре). Тогда $\mathbb{P}(A)$ может принимать лишь два значения 0 или 1.

Решение. Идея доказательства состоит в том, чтобы показать, что каждое «хвостовое» событие A не зависит от самого себя и, значит, $\mathbb{P}(A \cap A) = \mathbb{P}(A) \cdot \mathbb{P}(A)$, т.е. $\mathbb{P}(A) = \mathbb{P}^2(A)$, откуда $\mathbb{P}(A) = 0$ или 1.

Если $A \in \mathcal{X}$, то $A \in \Xi_1^\infty = \sigma(\bigcup\limits_n \Xi_1^n)$, где

$$\Xi_1^n = \sigma(\xi_1, \ldots, \xi_n),$$

и согласно задаче 67, найдутся множества $A_n \in \Xi_1^n$, $n \geqslant 1$, такие что $\mathbb{P}(A \bigtriangleup A_n) \to 0$, $n \to \infty$. Отсюда получим

$$\mathbb{P}(A_n) \to \mathbb{P}(A), \quad \mathbb{P}(A_n \cap A) \to \mathbb{P}(A). \tag{1.5}$$

Поскольку $A \in \mathcal{X}$, для каждого $n \geqslant 1$ события A_n и A независимы:

$$\mathbb{P}(A_n \cap A) = \mathbb{P}(A) \cdot \mathbb{P}(A_n),$$

откуда в силу (1.5) имеем $\mathbb{P}(A) = \mathbb{P}^2(A)$, т.е. $\mathbb{P}(A) = 0$ или 1.

□

Задача 81. Пусть X_n — последовательность независимых с.в., сходящаяся по вероятности к с.в. $X: \ X_n \xrightarrow{P} X$. Докажите, что с.в. X вырождена, т.е. $X \equiv x$, где x — некоторое число.

Решение.

1. Из курса функционального анализа известно, что из любой сходящейся по мере (в частности, по вероятностной) последовательности измеримых функций (в частности, с.в.) можно выделить подпоследовательность, сходящуюся почти всюду (п.н.):

   $$X_{n_k} \xrightarrow{\text{п.н.}} X \text{ при } k \to \infty.$$

   Если ввести σ-алгебры $\Xi_n^\infty = \sigma(X_n, X_{n+1}, \ldots)$ и хвостовую σ-алгебру $\mathcal{X} = \bigcap\limits_{n=1}^{\infty} \Xi_n^\infty$, то событие

   $$A = \{\omega : \ X = \lim_{k\to\infty} X_{n_k} \in B\} \in \mathcal{X},$$

   где B — некоторое борелевское множество на $\mathbb{R}$.

2. Из закона нуля и единицы Колмогорова для всякого разбиения прямой $\mathbb{R}$ на борелевские множества $\{B_m\}_{m\geqslant 1}$ ровно для одного $m = m_0: \quad \mathbb{P}(A_{B_{m_0}}) = 1$, для остальных $m: \quad \mathbb{P}(A_{B_m}) = 0$, где
$$A_{B_m} = \{\omega : X = \lim_{k\to\infty} X_{n_k} \in B_m\}.$$
Рассматривая для определенности $B_m = [m, m+1)$, далее для разбиения $\{B_m^2\}_{m=1}^N$ множества B_{m_0}, $\mu(B_m^2) = \frac{1}{N}$, при некотором $m = m_2$ получим $\mathbb{P}(A_{B^2_{m_2}}) = 1$.
Рассуждая так же для разбиения множества $B^2_{m_2}$ и так далее, получим систему вложенных стягивающихся отрезков $\{B_i\}$, $\mu(B_i) \xrightarrow{i\to\infty} 0$, так что
$$\mathbb{P}\big(X = \lim_{k\to\infty} X_{n_k} \in B_i\big) = 1.$$
Отсюда для единственной общей точки $\{x\} = \bigcap\limits_{i=1}^{\infty} B_i$ получим
$$\mathbb{P}\big(X = \lim_{k\to\infty} X_{n_k} = x\big) = 1,$$
т.е. $X = \lim\limits_{k\to\infty} X_{n_k} \equiv x$ — вырожденная с.в.

$\square$

Задача 82. Число α из отрезка $[0,1]$ назовем нормально приближаемым рациональными числами, если найдутся $c, \varepsilon > 0$ такие, что при любом натуральном q
$$\min_{p\in\mathbb{Z}} \left|\alpha - \frac{p}{q}\right| \geqslant \frac{c}{q^{2+\varepsilon}}. \tag{1.6}$$
Используя лемму Бореля-Кантелли, докажите, что множество нормально приближаемых чисел на отрезке $[0,1]$ имеет Лебегову меру 1.

Решение. 1. Зафиксируем $c, \varepsilon > 0$ и рассмотрим множество
$$A_q = \left\{\alpha \in [0,1] \mid \min_{p\in\mathbb{Z}} \left|\alpha - \frac{p}{q}\right| < \frac{c}{q^{2+\varepsilon}}\right\}.$$
Множество A_q вложено в окрестности точек $\frac{1}{q}, \frac{2}{q}, \ldots, \frac{q-1}{q}$ и односторонние окрестности точек 0 и 1 радиуса $\dfrac{c}{q^{2+\varepsilon}}$. Поэтому
$$\mu(A_q) \leqslant q \cdot \frac{2c}{q^{2+\varepsilon}} = \frac{2c}{q^{1+\varepsilon}}$$
и ряд $\sum\limits_q \mu(A_q)$ сходится. По лемме Бореля-Кантелли
$$\mu\{\overline{\lim}\, A_q\} = 0.$$
Событие $\overline{\lim}\, A_q$ означает, что для бесконечно большого числа $q \in \mathbb{N}: \quad \min\limits_{p\in\mathbb{Z}} \left|\alpha - \frac{p}{q}\right| < \frac{c}{q^{2+\varepsilon}}$.
Дополнение $A = [0,1] \setminus \overline{\lim}\, A_q$ состоит из тех точек, для которых неравенство (1.6) нарушается лишь для конечного числа $q \in \mathbb{N}$. Очевидно, $\mu(A) = 1$.

2. Отбросив из множества A все рациональные точки (мера которых равна нулю), получим множество $\hat{A}$: $\mu(\hat{A}) = 1$, состоящее из точек

$$\alpha \in [0,1]: \qquad \text{для конечного числа } q \in \mathbb{N} \quad 0 < \min_{p \in \mathbb{Z}} \left|\alpha - \frac{p}{q}\right| < \frac{c}{q^{2+\varepsilon}}.$$

3. Поскольку для каждого числа $\alpha \in \hat{A}$ неравенство (1.6) нарушается лишь для конечного числа $q \in Q_\alpha \subset \mathbb{N}$, из иррациональности α следует, что

$$c_\alpha = \min_{q \in Q_\alpha} \min_{p \in \mathbb{Z}} \left|\alpha - \frac{p}{q}\right| > 0.$$

   Поэтому для всякого $\alpha \in \hat{A}$ найдутся $c = c_\alpha > 0$, $\varepsilon > 0$, такие что неравенство (1.6) выполнено при всех $q \in \mathbb{N}$. Множество $\hat{A}$ таких чисел, как уже было получено, имеет Лебегову меру 1.

□

Задача 83. В некотором Вузе проходит экзамен. Количество экзаменационных билетов N. Перед экзаменационной аудиторией выстроилась очередь из студентов, которые не знают, чему равно N. Согласно этой очереди студенты вызываются на экзамен (второй студент заходит в аудиторию после того как из нее выйдет первый и т.д.). Каждый студент с равной вероятностью может выбрать любой из N билетов (в независимости от других студентов). Проэкзаменованные студенты, выходя из аудитории, сообщают оставшейся очереди номера своих билетов. Оцените (сверху), сколько студентов должно быть проэкзаменовано, чтобы оставшаяся к этому моменту очередь смогла оценить число экзаменационных билетов с точностью 10% с вероятностью, не меньшей 0.95.

Решение. Пусть X_i, $i = 1, \ldots, n$ — номер билета, который вытащил i-й студент. Из условия задачи вытекает, что с.в. X_i имеет дискретное равномерное распределение на отрезке $[1, \ldots, N]$: $\mathbb{P}(X_i \leqslant k) = \frac{k}{N}$. Функция распределения с.в. $X_{(n)} = \max\limits_{1 \leqslant i \leqslant n} X_i$ равна

$$\mathbb{P}(X_{(n)} \leqslant k) = \mathbb{P}(X_1 \leqslant k, \ldots, X_n \leqslant k) = \big(\mathbb{P}(X_1 \leqslant k)\big)^n = \left(\tfrac{k}{N}\right)^n.$$

Асимптотически несмещенная оценка для числа билетов N имеет вид

$$\hat{N} = (n+1)/n \cdot X_{(n)}.$$

Действительно,

$$\mathbb{E}X_{(n)} = \sum_{k=1}^{N} k\,\frac{k^n - (k-1)^n}{N^n} = \sum_{k=1}^{N} \frac{k^{n+1} - (k-1)^{n+1} - (k-1)^n}{N^n} =$$

$$= N - N\sum_{k=1}^{N-1} \frac{k^n}{N^n} \cdot \frac{1}{N} \approx N - N\int\limits_0^1 x^n\,dx = \frac{n}{n+1} \cdot N.$$

В задаче требуется оценить число N с помощью его оценки $\hat{N}$ с точностью $\alpha = 0.1$ с вероятностью не меньшей $\gamma = 0.95$:

$$\mathbb{P}(\hat{N} > N(1-\alpha)) \geqslant \gamma$$

$$\Leftrightarrow\ 1 - (1-\alpha)^n \Big(\frac{n}{n+1}\Big)^n \geqslant \gamma.$$

Данное неравенство выполнено при $n \geqslant 20$. □

Задача 84. Пусть $\xi \in \mathrm{Po}(\lambda)$, $\lambda \gg 1$. Покажите, что

$$\sqrt{\xi} \approx \sqrt{\lambda} + N(0,\, 1/\,4).$$

Решение. 1. Сначала покажем, что при $\lambda \to \infty$: $\frac{\xi-\lambda}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$.

Ищем характеристическую функцию с.в. $\frac{\xi-\lambda}{\sqrt{\lambda}}$:

$$\varphi_\lambda(t) = \exp\{-it\sqrt{\lambda} + \lambda(e^{it/\sqrt{\lambda}} - 1)\}.$$

Применим разложение по Тейлору для функции e^z в окрестности нуля:

$$\varphi_\lambda(t) = \exp\{-it\sqrt{\lambda} + \lambda(it/\,\sqrt{\lambda} - t^2/\,2\lambda + o(1/\,\lambda))\} = \exp\{-\,t^2/\,2 + o(1)\}.$$

Поэтому $\varphi_\lambda(t) \xrightarrow{\lambda\to\infty} \exp\{-\,t^2/\,2\}$ и по теореме непрерывности для характеристических функций $\frac{\xi-\lambda}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$.

2. Имеем

$$\xi \approx \lambda + \eta\sqrt{\lambda}, \quad \eta \in N(0,1)$$

$$\Rightarrow\ \xi \approx \Big(\sqrt{\lambda} + \frac{\eta}{2}\Big)^2 \Big[1 - \frac{\eta^2/\,4}{(\sqrt{\lambda} + \eta/\,2)^2}\Big].$$

Т.к. $\mathbb{E}\eta^2 = 1$, при больших λ можно считать

$$\xi \approx \Big(\sqrt{\lambda} + \frac{\eta}{2}\Big)^2,$$

откуда вытекает искомое приближение $\sqrt{\xi} \approx \sqrt{\lambda} + N(0,\, 1/\,4)$.

□

Задача 85. В течение года фирма осуществляет $K \in \mathrm{Po}(\lambda)$ сделок (K – с.в., имеющая распределение Пуассона с параметром $\lambda = 100000$ [сделок]). Каждая сделка приносит фирме прибыль $V_n \in R[a,b]$ (V_n – с.в., имеющая равномерное распределение на отрезке $[a,b] = [-50\$, 100\$]$, n – номер сделки). Считая, что K, V_1, V_2, … — независимые в совокупности с.в., оцените

$$\mathbb{P}\Big(\sum_{n=1}^{K} V_n \leqslant 0\Big)\Big/\,\mathbb{P}\Big(\sum_{n=1}^{K} V_n > 0\Big). \tag{1.7}$$

Решение. Из предыдущей задачи воспользуемся оценкой для $K \in \mathrm{Po}(\lambda)$, т.к. λ достаточно велико:
$$K \approx \lambda + \eta\sqrt{\lambda}, \quad \eta \in N(0,1).$$
Используем для K доверительный интервал $[\lambda - 3\sqrt{\lambda}, \lambda + 3\sqrt{\lambda}]$, чей доверительный уровень высок:
$$\gamma = \mathbb{P}(K \in [\lambda - 3\sqrt{\lambda}, \lambda + 3\sqrt{\lambda}]) = 2 \cdot \Phi(3) - 1 \approx 0.9973.$$
Обозначим $S_K = \sum\limits_{n=1}^{K} V_n$ и перепишем (1.7):
$$\frac{\mathbb{P}\big(S_K \leqslant 0\big)}{\mathbb{P}\big(S_K > 0\big)} = \frac{1}{\mathbb{P}\big(S_K > 0\big)} - 1$$
и будем оценивать $\mathbb{P}\big(S_K > 0\big)$. Имеем
$$\mathbb{P}\big(S_K > 0\big) = \mathbb{P}\big(S_K > 0 \,|\, K \in [\lambda - 3\sqrt{\lambda}, \lambda + 3\sqrt{\lambda}]\big) \cdot \gamma + \\ + \mathbb{P}\big(S_K > 0 \,|\, K \notin [\lambda - 3\sqrt{\lambda}, \lambda + 3\sqrt{\lambda}]\big) \cdot (1 - \gamma).$$
В силу теоремы Натана для S_K выполнено приближение (поскольку значение управляющего параметра λ достаточно велико)
$$S_K \approx \frac{a+b}{2} \cdot K + \frac{b-a}{2\sqrt{3}} \cdot \xi\sqrt{K}, \quad \xi \in N(0,1).$$
Тогда условие $S_K > 0$ примет вид
$$\xi > -\sqrt{K}\frac{(a+b)\sqrt{3}}{b-a}.$$
В итоге получим следующие оценки:
$$\gamma \cdot \mathbb{P}\Big(\xi > -(\lambda - 3\sqrt{\lambda})^{1/2}\frac{(a+b)\sqrt{3}}{b-a}\Big) < \mathbb{P}\big(S_K > 0\big) \leqslant \gamma \cdot \mathbb{P}\Big(\xi > -(\lambda + 3\sqrt{\lambda})^{1/2}\frac{(a+b)\sqrt{3}}{b-a}\Big) + (1 - \gamma).$$
Подставляя значения a, b, λ, выясняется, что можно считать
$$\mathbb{P}\Big(\xi > -(\lambda - 3\sqrt{\lambda})^{1/2}\frac{(a+b)\sqrt{3}}{b-a}\Big) \approx 1, \ \mathbb{P}\Big(\xi > -(\lambda + 3\sqrt{\lambda})^{1/2}\frac{(a+b)\sqrt{3}}{b-a}\Big) \approx 1.$$
Поэтому можно оценить $\mathbb{P}\big(S_K > 0\big)$ в виде
$$\gamma < \mathbb{P}\big(S_K > 0\big) \leqslant 1,$$
откуда
$$\frac{\mathbb{P}\big(S_K \leqslant 0\big)}{\mathbb{P}\big(S_K > 0\big)} \leqslant \frac{1}{\gamma} - 1 \approx 0.0027.$$
□

Задача 86. 1) Имеется монетка (несимметричная). Несимметричность монетки заключается в том, что либо орел выпадает в два раза чаще решки; либо наоборот (априорно (до проведения опытов) оба варианта считаются равновероятными). Монетку бросили 10 раз. Орел выпал 7 раз. Определите апостериорную вероятность того, что орел выпадает в два раза чаще решки (апостериорная вероятность считается с учетом проведенных опытов (иначе говоря, это просто условная вероятность)).

2) Определите апостериорную вероятность того, что орел выпадает не менее чем в два раза чаще решки. Если несимметричность монетки заключается в том, что либо орел выпадает не менее чем в два раза чаще решки; либо наоборот (априорно оба варианта считаются равновероятными).

Решение. 1) Введем события $B = \{p = 1/3\}$, $\overline{B} = \{p = 2/3\}$, где p — вероятность выпадения орла в одном опыте. По условию задачи $\mathbb{P}(B) = \mathbb{P}(\overline{B}) = 1/2$. Требуется найти $\mathbb{P}(\overline{B}\,|\,r_{10} = 7)$, где $r_{10} \in \mathrm{Bi}(p, 10)$. По формуле Байеса получим

$$\mathbb{P}(\overline{B}\,|\,r_{10} = 7) = \frac{\mathbb{P}(r_{10} = 7\,|\,\overline{B}) \cdot \mathbb{P}(\overline{B})}{\mathbb{P}(r_{10} = 7\,|\,\overline{B}) \cdot \mathbb{P}(\overline{B}) + \mathbb{P}(r_{10} = 7\,|\,B) \cdot \mathbb{P}(B)} =$$

$$= \frac{C_{10}^7 \cdot (2/3)^7 (1/3)^3}{C_{10}^7 \cdot (2/3)^7 \cdot (1/3)^3 + C_{10}^7 \cdot (1/3)^7 \cdot (2/3)^3} = \frac{2^7}{2^7 + 2^3} = \frac{16}{17} \approx 0.9412.$$

2) Аналогично предыдущему пункту введем события $A = \{p \in R[0, 1/3]\}$, $\overline{A} = \{p \in R[2/3, 1]\}$, где p — вероятность выпадения орла в одном опыте. По условию задачи $\mathbb{P}(A) = \mathbb{P}(\overline{A}) = 1/2$. Требуется найти $\mathbb{P}(\overline{A}\,|\,r_{10} = 7)$, где $r_{10} \in \mathrm{Bi}(p, 10)$. По формуле Байеса получим

$$\mathbb{P}(\overline{A}\,|\,r_{10} = 7) = \frac{\mathbb{P}(r_{10} = 7\,|\,\overline{A}) \cdot \mathbb{P}(\overline{A})}{\mathbb{P}(r_{10} = 7\,|\,\overline{A}) \cdot \mathbb{P}(\overline{A}) + \mathbb{P}(r_{10} = 7\,|\,A) \cdot \mathbb{P}(A)} =$$

$$= \frac{\int\limits_{2/3}^{1} \mathbb{P}(r_{10} = 7\,|\,p) P(dp|\overline{A})}{\int\limits_{2/3}^{1} \mathbb{P}(r_{10} = 7\,|\,p) P(dp|\overline{A}) + \int\limits_{0}^{1/3} \mathbb{P}(r_{10} = 7\,|\,p) P(dp|A)} =$$

$$= \frac{3\int\limits_{2/3}^{1} \mathbb{P}(r_{10} = 7\,|\,p)\,dp}{3\int\limits_{2/3}^{1} \mathbb{P}(r_{10} = 7\,|\,p)\,dp + 3\int\limits_{0}^{1/3} \mathbb{P}(r_{10} = 7\,|\,p)\,dp} =$$

$$= \frac{3 \cdot C_{10}^7 \int\limits_{2/3}^{1} p^7 \cdot (1-p)^3\,dp}{3 \cdot C_{10}^7 \int\limits_{2/3}^{1} p^7 \cdot (1-p)^3\,dp + 3 \cdot C_{10}^7 \int\limits_{0}^{1/3} p^7 \cdot (1-p)^3\,dp} \approx \frac{0.1438}{0.1438 + 0.0024} \approx 0.9836.$$

$\square$

Задача 87. Восемь мальчиков и семь девочек купили билеты в кинотеатр на 15 подряд идущих сидячих мест. Предположим, что все 15! возможных способов сесть равновероятны. Вычислите среднее число пар рядом сидящих мальчика и девочки. Например, м, м, м, м, м, м, ж, м, ж, ж, ж, ж, ж, ж содержит три такие пары.

Решение. Пусть X_n, $n = 1, \ldots, 14$ — индикатор события, что на n-ом и $n+1$-ом месте сидят мальчик и девочка (неважно в каком порядке). Тогда $\mathbb{E}X_n = \mathbb{P}(X_n = 1)$ и искомое среднее число пар равно $r = \mathbb{E}(X_1 + \ldots + X_{14})$.

Найдем $\mathbb{P}(X_n = 1)$:

$$\mathbb{P}(X_n = 1) = \frac{(8 \cdot 7 + 7 \cdot 8) \cdot 13!}{15!} = \frac{2 \cdot 8 \cdot 7}{14 \cdot 15} = \frac{8}{15},$$

откуда $r = 14 \cdot \mathbb{E}X_n = \frac{112}{15} = 7\frac{7}{15}$.

□

Задача 88. На первом этаже семнадцатиэтажного общежития в лифт вошли десять человек. Предполагая, что каждый из вошедших может с равной вероятностью жить на любом из шестнадцати этажей (со 2-го по 17-ый), найдите среднее число остановок лифта.

Решение. Введем с.в. X_n, $n = 2, \ldots, 17$:

$$X_n = \mathcal{I}(\text{«на } n\text{-м этаже была остановка»}),$$

и с.в. ξ_k = «номер этажа, где живет k-й человек», $k = 1, \ldots, 10$. Из условия задачи следует, что

$$\mathbb{P}(\xi_k = 2) = \mathbb{P}(\xi_k = 3) = \ldots = \mathbb{P}(\xi_k = 17) = 1/\,16.$$

Среднее число остановок лифта равно $N = \mathbb{E}(X_2 + X_3 + \ldots + X_{17})$.

Найдем $\mathbb{E}X_n = \mathbb{P}(X_n = 1)$:

$$\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0) = 1 - \mathbb{P}(\xi_1 \neq n, \xi_2 \neq n, \ldots, \xi_{10} \neq n) =$$

$$= 1 - \mathbb{P}(\xi_1 \neq n) \cdot \ldots \cdot \mathbb{P}(\xi_{10} \neq n) = 1 - \big(1 - 1/\,16\big)^{10} = 1 - \big(15/\,16\big)^{10}$$

Отсюда находим среднее число остановок лифта

$$N = 16 \cdot \mathbb{E}X_n = 16 - 15 \cdot \big(15/\,16\big)^9 \approx 7.6.$$

□

Задача 89. В самолете n мест. Есть n пассажиров, выстроившихся друг за другом в очередь. Во главе очереди – «заяц». У всех, кроме «зайца», есть билет, на котором указан номер посадочного места. Так как «заяц» входит первым, он случайным образом занимает некоторое место. Каждый следующий пассажир, входящий в салон самолета, действует по такому принципу: если его место свободно, то садится на него, если занято, то занимает с равной вероятностью любое свободное. Найдите вероятность того, что последний пассажир сядет на свое место.

Решение. Для определенности считаем, что k-й пассажир должен сесть на k-е место, $k = 1, \dots, n$. Значит, в задаче нужно найти вероятность p_n того, что первые $n-1$ пассажиров займут первые $n-1$ мест.

По индукции докажем, что $\forall n > 1 : \ p_n = 1/2$.

1) $n = 2 \Rightarrow \quad p_2 = 1/2$ — вероятность того, что «заяц» займет свое 1-е место.

2)
$$n = 3 \ \Rightarrow \quad p_3 = \frac{1}{3} \cdot \big(p_2 + 1\big) = 1/2$$

— либо «заяц» займет $(n-1) = 2$-е место, и тогда получим подзадачу, эквивалентную исходной с числом пассажиров $(n-1) = 2$, где роль «зайца» играет 2-й пассажир,

— либо «заяц» займет свое 1-е место.

n)
$$p_n = \frac{1}{n} \cdot \big(p_2 + p_3 + \ldots + p_{n-1} + 1\big) = \frac{1}{n} \cdot \Big(\frac{n-2}{2} + 1\Big) = 1/2. \qquad (1.8)$$

Формула (1.8) для нахождения p_n представляет собой формулу полной вероятности.

Здесь получаем, что условная вероятность исходного события (когда последний пассажир сядет на свое место) при условии, что «заяц» занял k-е место, $2 \leqslant k \leqslant n-1$, равна p_{n-k+1}, поскольку тогда получаем подзадачу, эквивалентную исходной с числом пассажиров $n - k + 1$, где роль «зайца» играет k-й пассажир: пассажиры с номерами $i = 2, \dots, k-1$ займут свои места, а для k-го пассажира «своим» будет 1-е место, иначе он займет чье-нибудь место из оставшихся $n - k$ пассажиров.

Если же «заяц» займет свое 1-е место, то все пассажиры займут свои места, т.е. и последний пассажир всегда сядет на свое место.

□

Задача 90. В некотором городе прошел второй тур выборов. Выбор был между двумя кандидатами A и B (графы «против всех» на этих выборах не было). Сколько человек надо опросить на выходе с избирательных участков, чтобы исходя из ответов можно было определить долю проголосовавших за кандидата A с точностью 5% и с вероятностью не меньшей 0.99. Считайте, что исходя из голосования в первом туре, известно, что каждый из кандидатов наберет не меньше 30% голосов избирателей.

Решение. Обозначим долю проголосовавших за кандидата A через p. По условию задачи $p \in [0.3, 0.7]$. Результат опроса одного человека является с.в. $X \in \mathrm{Be}(p)$, где

$$X = \mathcal{I}(\text{«опрошенный проголосовал за кандидата } A\text{»}).$$

Пусть опрошено n человек, $S_n = X_1 + \ldots + X_n$, $X_k \in \mathrm{Be}(p)$. В задаче требуется найти n, при котором эмпирическая доля $\frac{S_n}{n}$ проголосовавших за кандидата A оценивает p с точностью $\alpha = 0.05$ с вероятностью $\gamma = 0.99$:

$$\mathbb{P}\Big(\frac{S_n}{n} \in [p - \alpha, p + \alpha]\Big) \geqslant \gamma.$$

Имеем

$$\mathbb{P}\Big(\frac{S_n}{n} \in [p-\alpha, p+\alpha]\Big) = \mathbb{P}\Big(\Big|\frac{S_n - np}{n}\Big| \leqslant \alpha\Big) =$$

$$= \mathbb{P}\Big(\Big|\frac{S_n - np}{\sqrt{n \cdot p(1-p)}}\Big| \leqslant \frac{\sqrt{n}\alpha}{\sqrt{p(1-p)}}\Big) = \mathbb{P}\Big(\Big|\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}}\Big| \leqslant \frac{\sqrt{n}\alpha}{\sqrt{p(1-p)}}\Big).$$

Согласно ц.п.т. при больших n можно считать $\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} \in N(0,1)$, поэтому

$$\mathbb{P}\Big(\Big|\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}}\Big| \leqslant \frac{\sqrt{n}\alpha}{\sqrt{p(1-p)}}\Big) \approx \Phi\Big(\frac{\sqrt{n}\alpha}{\sqrt{p(1-p)}}\Big) - \Phi\Big(-\frac{\sqrt{n}\alpha}{\sqrt{p(1-p)}}\Big) = 2\Phi\Big(\frac{\sqrt{n}\alpha}{\sqrt{p(1-p)}}\Big) - 1.$$

Чтобы условие $2\Phi\Big(\frac{\sqrt{n}\alpha}{\sqrt{p(1-p)}}\Big) - 1 \geqslant \gamma$ выполнялось при всех $p \in [0.3, 0.7]$, достаточно чтобы оно выполнялось при наименьшем возможном $\frac{\sqrt{n}\alpha}{\sqrt{p(1-p)}}$. Это достигается при $\bar{p} = 0.5$. Имеем

$$\Phi\Big(2\alpha\sqrt{n}\Big) \geqslant \frac{\gamma+1}{2} \quad \Rightarrow \qquad n \geqslant \Big(\tfrac{1}{2\alpha}\Phi^{-1}\big(\tfrac{\gamma+1}{2}\big)\Big)^2.$$

Подставляя значения α и γ, находим $\quad n \geqslant 665.$

$\square$

Задача 91. На множестве $n!$ перестановок n различных элементов задано равномерное распределение. Обозначим через ξ_k случайную величину, равную числу инверсий, образованных элементом с номером k, т.е. равную числу элементов с номерами меньшими чем k, которые стоят в перестановке правее элемента с номером k. Покажите, что

$$\frac{\sum\limits_{k=1}^{n} \xi_k - n^2/4}{n^{3/2}/6} \xrightarrow{d} N(0,1) \qquad \text{при } n \to \infty.$$

Решение. 1. Введем с.в. $\xi_{k,i}$, $k > i$:

$$\xi_{k,i} = \mathcal{I}(\text{«}k \text{ находится левее числа } i\text{»}).$$

Если позициии всех чисел, кроме k и i, фиксированы, то в одной из двух расстановок k находится левее числа i, поэтому $\mathbb{E}\xi_{k,i} = 1/2$.

При $i < j < k$, если позиции всех чисел, кроме k, i и j, фиксированы, то в двух из 6 расстановок число k находится левее чисел i и j, а в одной расстановке k левее j, и j находится левее i, отсюда $\mathbb{E}(\xi_{k,i}\xi_{k,j}) = 1/3$, $\mathbb{E}(\xi_{k,j}\xi_{j,i}) = 1/6$.

Наконец, если числа i, j, k, l попарно различны, $i < j$, $k < l$, то с.в. $\xi_{j,i}$ и $\xi_{l,k}$ независимы и $\mathbb{E}(\xi_{j,i}\xi_{l,k}) = \mathbb{E}\xi_{j,i} \cdot \mathbb{E}\xi_{l,k} = 1/4$.

Поскольку по определению

$$\xi_k = \xi_{k,1} + \xi_{k,2} + \ldots + \xi_{k,k-1},$$

имеем $\mathbb{E}\xi_k = (k-1)/\,2$. Далее

$$\mathbb{E}\xi_k^2 = \mathbb{E}\xi_{k,1}^2 + \ldots + \mathbb{E}\xi_{k,k-1}^2 + 2\sum_{i<j<k} \mathbb{E}(\xi_{k,i}\xi_{k,j}) =$$

$$= \frac{k-1}{2} + 2\cdot\frac{(k-1)(k-2)}{2}\cdot\frac{1}{3} = \frac{2k^2-3k+1}{6},$$

откуда $\operatorname{Var}\xi_k = \mathbb{E}\xi_k^2 - \bigl(\mathbb{E}\xi_k\bigr)^2 = \frac{k^2-1}{12}$.

Также вычислим при $k > m$

$$\mathbb{E}\xi_k\xi_m = \mathbb{E}(\xi_{k,1}+\xi_{k,2}+\ldots+\xi_{k,k-1})(\xi_{m,1}+\xi_{m,2}+\ldots+\xi_{m,m-1}) =$$

$$= \mathbb{E}\bigl(\xi_{k,1}\xi_{m,1}+\ldots+\xi_{k,m-1}\xi_{m,m-1}\bigr) + \mathbb{E}\xi_{k,m}(\xi_{m,1}+\ldots+\xi_{m,m-1}) +$$

$$+ \sum_{0<i<k,0<j<m,i\neq j,i\neq m} \mathbb{E}\xi_{k,i}\xi_{m,j} = \frac{m-1}{3} + \frac{m-1}{6} + \frac{(k-3)(m-1)}{4} = \frac{(k-1)(m-1)}{4} = \mathbb{E}\xi_k\cdot\mathbb{E}\xi_m.$$

Это означает, что с.в. ξ_k и ξ_m некоррелированы, $k \neq m$.

2. Введем с.в. $T_n = \sum\limits_{k=1}^{n} \xi_k$ — общее число инверсий в перестановке. Имеем

$$\mathbb{E}T_n = \sum_{k=1}^{n} \mathbb{E}\xi_k = \sum_{k=1}^{n} \frac{k-1}{2} = \frac{1}{4}n(n-1) = \frac{n^2}{4}\bigl(1+O(1/\,n)\bigr).$$

Из некоррелированности с.в. ξ_k и ξ_m, $k \neq m$, следует

$$\operatorname{Var} T_n = \sum_{k=1}^{n} \operatorname{Var}\xi_k = \sum_{k=1}^{n} \frac{k^2-1}{12} = -\frac{n}{12} + \frac{n(n+1)(2n+1)}{12\cdot 6} = \frac{n^3}{36}\bigl(1+O(1/\,n)\bigr).$$

3. Для всякой сл.в. $X_k = \xi_k - \mathbb{E}\xi_k$ характеристическая функция имеет вид

$$\varphi_{X_k}(t) = 1 - \frac{t^2\operatorname{Var}\xi_k}{2} + \overline{o}(t^2) = 1 - \frac{t^2(k^2-1)}{24} + \overline{o}(t^2).$$

Для сл.в. $Y_n = \frac{T_n - \mathbb{E}T_n}{\sqrt{\operatorname{Var} T_n}}$, исходя из независимости сл.в. $\xi_1, \ldots, \xi_n$ и свойств характеристических функций, получим

$$\varphi_{Y_n}(t) = \varphi_{X_1}\Bigl(\frac{t}{\sqrt{\operatorname{Var} T_n}}\Bigr)\cdot\ldots\cdot\varphi_{X_n}\Bigl(\frac{t}{\sqrt{\operatorname{Var} T_n}}\Bigr).$$

Поскольку

$$\varphi_{X_k}\Bigl(\frac{t}{\sqrt{\operatorname{Var} T_n}}\Bigr) = 1 - \frac{3(k^2-1)t^2}{2n^3} + \overline{o}\Bigl(\frac{1}{n^3}\Bigr) = \exp\Bigl(-\frac{3(k^2-1)t^2}{2n^3} + \overline{o}\Bigl(\frac{k^2}{n^3}\Bigr)\Bigr),$$

находим

$$\varphi_{Y_n}(t) = \exp\Bigl(-\sum_{k=1}^{n}\frac{3(k^2-1)t^2}{2n^3} + \sum_{k=1}^{n}\overline{o}\Bigl(\frac{k^2}{n^3}\Bigr)\Bigr) = \exp\Bigl(-\frac{t^2}{2} + \overline{o}(1)\Bigr) \xrightarrow{n\to\infty} \exp\Bigl(-\frac{t^2}{2}\Bigr),$$

откуда из предельных свойств для характеристических функций вытекает

$$Y_n = \frac{T_n - \mathbb{E}T_n}{\sqrt{\operatorname{Var} T_n}} \xrightarrow{d} N(0,1),$$

что равносильно

$$\frac{T_n - n^2/4}{n^{3/2}/6} \xrightarrow{d} N(0,1).$$

□

Задача 92. Найдите вероятность того, что пара случайно выбранных из $E^n = \{0,1\}^n$ векторов является ортогональной

а) над полем $F_2 = \{0,1\}$;

б) над полем действительных чисел.

Решение. а) Пусть

$$\xi_{x,y} = \begin{cases} 1, & \text{если } (x,y) = 0, \\ 0, & \text{если } (x,y) = 1. \end{cases}$$

Для искомой вероятности тогда имеем

$$P = \frac{1}{2^n \cdot 2^n} \sum_{x,y} \xi_{x,y} = \frac{1}{2^{2n}} \sum_x \sum_y \xi_{x,y}.$$

Поскольку при $x \neq 0$ с.в.

$$\xi_{x,y} = 1 + x_1 y_1 + \ldots + x_n y_n$$

является невырожденной линейной булевой функцией переменной $y \in E^n$, внутренняя сумма $\sum\limits_y \xi_{x,y}$ равна числу нулей функции $\xi_{x,y}$, т.е. равна 2^{n-1}. При $x = 0$, очевидно, $\sum\limits_y \xi_{x,y} = 2^n$. Отсюда

$$P = \frac{1}{2^{2n}} \Big(\sum_{x \neq 0} 2^{n-1} + 2^n \Big) = \frac{(2^n - 1)\cdot 2^{n-1} + 2^n}{2^{2n}} = \frac{1}{2} + \frac{1}{2^{n+1}}.$$

б) Аналогично получим

$$P = \frac{1}{2^{2n}} \sum_{x,y} \xi_{x,y} = \frac{1}{4^n} \sum_{k=0}^{n} \sum_{x\colon \|x\| = k} \sum_y \xi_{x,y} =$$

$$= \frac{1}{4^n} \sum_{k=0}^{n} \binom{n}{k} 2^{n-k} = \frac{1}{4^n} \cdot \big(1+2\big)^n = \Big(\frac{3}{4}\Big)^n.$$

□

Задача 93. Частица находится в начальный момент в вершине треугольника (Паскаля). Затем частица начинает двигаться (с вероятностью p вправо и с вероятностью $q = 1 - p$ — влево). С помощью аппарата производящих функций определите вероятность того, что частица за n шагов двигалась вправо ровно k раз.

Решение. Рассмотрим с.в. $X = \mathcal{I}$(«частица двигалась вправо»). Тогда $X \in \mathrm{Be}(p)$.

В задаче необходимо найти $\mathbb{P}(S_n = k)$, где $S_n = X_1 + \ldots + X_n$, случайные величины $X_i \in \mathrm{Be}(p)$, $i = 1, \ldots, n$ независимы. Из вида производящей функции с.в. S_n

$$\psi(z) = \mathbb{E}z^{S_n} = \sum_{k=0}^{n} z^k \, \mathbb{P}(S_n = k)$$

видно, что требуется найти коэффициент при z^k в разложении $\psi(z)$.

Из свойств производящих функций имеем

$$\mathbb{E}z^X = (pz + q) \quad \Rightarrow \psi(z) = \mathbb{E}z^{S_n} = (pz + q)^n.$$

По формуле бинома Ньютона находим

$$\mathbb{P}(S_n = k) = C_n^k p^k q^{n-k}.$$

□

Задача 94. Покажите, что если независимые случайные величины $X_1, \ldots, X_n$ имеют показательное распределение, т.е.

$$f_{X_i}(x) = \begin{cases} \lambda_i \exp(-\lambda_i x), \ x \geqslant 0 \\ 0, \ x < 0 \end{cases}$$

(часто пишут $X_i \in \mathrm{Exp}(\lambda_i)$), то

$$\min\{X_1, \ldots, X_n\} \in \mathrm{Exp}\Big(\sum_{i=1}^{n} \lambda_i\Big).$$

Решение. Из условия задачи следует

$$\mathbb{P}(X_i > x) = e^{-\lambda_i x}.$$

Обозначим с.в. $T = \min\{X_1, \ldots, X_n\}$. Найдем функцию распределения данной с.в.:

$$F_T(x) = \mathbb{P}(T \leqslant x) = 1 - \mathbb{P}(T > x) = 1 - \mathbb{P}(X_1 > x, \ldots, X_n > x) =$$

$$= 1 - \mathbb{P}(X_1 > x) \cdot \ldots \cdot \mathbb{P}(X_n > x) = 1 - \exp\Big(-x \sum_{i=1}^{n} \lambda_i\Big),$$

откуда по определению $T \in \mathrm{Exp}\big(\sum\limits_{i=1}^{n} \lambda_i\big)$.

□

Задача 95. Имеется n пронумерованных писем и n пронумерованных конвертов. Письма случайным образом раскладываются по конвертам (т.е. все $n!$ способов распределения n писем по n конвертам, так чтобы в каждом конверте было ровно по одному письму, считаются равновероятными). Найдите математическое ожидание случайной величины, равной числу совпадений (числу писем, лежащих в конвертах с теми же номерами).

Решение. Пронумеруем конверты 1, 2, ..., n и в качестве элементарного события рассмотрим $\omega = (\omega_1, \ldots, \omega_n)$, где ω_i — номер письма, попавшего в i-й конверт, $1 \le \omega_i \le n$.

Для каждого i-го письма введем с.в.

$$\xi_i = \mathcal{I}(\text{«}i\text{-е письмо попало в } i\text{-й конверт»}) = \mathcal{I}\{\omega_i = i\}.$$

Тогда $\mathbb{E}\xi_i = \mathbb{P}\{\omega_i = i\}$ и математическое ожидание общего числа совпадений равно $r_n = \mathbb{E}\Big(\sum\limits_{i=1}^{n} \xi_i\Big) = \sum\limits_{i=1}^{n} \mathbb{P}\{\omega_i = i\}$.

Каждое из n писем с одинаковой вероятностью может попасть в любой из n конвертов (т.к. по условию все способы распределения писем равновероятны), в частности, в i-й конверт, поэтому

$$\mathbb{P}\{\omega_i = i\} = \frac{1}{n} \Rightarrow r_n = \sum_{i=1}^{n} \mathbb{P}\{\omega_i = i\} = 1.$$

□

Задача 96. (Задача игрока де Мере) Что более вероятно: при одновременном бросании четырех игральных костей получить хотя бы одну единицу или при 24 бросаниях по две игральные кости одновременно получить хотя бы один раз две единицы? Найти вероятности указанных событий.

Решение. При четырех бросаниях правильной игральной кости число возможных (равновероятных) исходов равно 6^4, при этом в 5^4 случаях не выпадет единицы. Значит, вероятность выпадения хотя бы одной единицы при четырех бросаниях правильной игральной кости равно $\frac{6^4-5^4}{6^4} \approx 0.5177$. Аналогично находим, что вероятность выпадения хотя бы один раз двух единиц при 24-кратном бросании двух правильных игральных костей равна $\frac{36^{24}-35^{24}}{36^{24}} \approx 0.4914$. То есть первое событие происходит с вероятностью большей $\frac{1}{2}$, а второе — меньшей $\frac{1}{2}$. Это кажется удивительным, так как шансы получить единицу при бросании одной кости в шесть раз превосходят шансы получить две единицы при бросании двух костей, а 24 в шесть раз больше 4.

Посчитаем математическое ожидание числа бросков правильной игральной кости до первого выпадения единицы: $\sum\limits_{k=1}^{\infty} k\,\frac{1}{6}\big(\frac{5}{6}\big)^{k-1} = 6$. Напоминание:

$$f(x) = \sum_{k=1}^{\infty} x^k = \frac{1}{1-x}, \quad f'(x) = \sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}, \quad (1-x)f'(x) = \sum_{k=1}^{\infty} k(1-x)x^{k-1} = \frac{1}{1-x}$$

$$\Rightarrow \sum_{k=1}^{\infty} k\,\frac{1}{6}\Big(\frac{5}{6}\Big)^{k-1} = 6.$$

Посчитаем математическое ожидание числа бросков двух правильных игральных костей до первого выпадения двух единиц: $\sum\limits_{k=1}^{\infty} k\,\frac{1}{36}\left(\frac{35}{36}\right)^{k-1} = 36$. Таким образом, эти математические ожидания подчиняются «правилу пропорций». Если p — вероятность некоторого события, то среднее число испытаний, необходимое до первого выполнения этого события, равно $\frac{1}{p}$. □

Задача 97. n человек разного роста случайным образом выстраиваются в шеренгу. Найти вероятность того, что:

а) самый низкий окажется i-м слева;

б) самый высокий окажется первым слева, а самый низкий — последним слева;

в) самый высокий и самый низкий окажутся рядом;

г) между самым высоким и самым низким расположатся более k человек.

Решение. Пронумеровав людей от 1 до n, рассматриваем в качестве элементарных событий перестановки чисел 1, 2, ..., n, имеющие, согласно условию, одинаковую вероятность (равную $\frac{1}{n!}$).

а) На i-й позиции слева с одинаковой вероятностью может оказаться каждый из n человек, поэтому искомая вероятность равна $p_1 = \frac{1}{n}$.

б) Самый высокий окажется первым слева, как и в предыдущем пункте, с вероятностью $p = \frac{1}{n}$. Если первая позиция слева занята, на последней позиции слева с одинаковой вероятностью может оказаться каждый из остальных $n-1$ человек, поэтому искомая вероятность равна
$$p_2 = p\,\frac{1}{n-1} = \frac{1}{n(n-1)}.$$

в) Число способов построить шеренгу из n человек равно $n!$. Для нужного построения шеренги $n-2$ человека (кроме самого высокого и самого низкого) строятся произвольным способом, далее $n-1$ способом находится место для самого высокого и самого низкого, с учетом их взаимного расположения умножаем на 2 полученное число способов. Итого, искомая вероятность равна $\frac{2(n-1)(n-2)!}{n!} = \frac{2}{n}$;

Заметим, что при расположении людей по кругу ответ будет $\frac{2(n-2)!}{(n-1)!} = \frac{2}{n-1}$;

г) Аналогично, ищем число способов такой расстановки, что между самым высоким и самым низким ровно j человек: $2(n-2)!(n-j-1)$. (Заметим, что в случае расположения людей по кругу эта вероятность равна так же, как и в предыдущем пункте $\frac{2}{n-1}$) Тогда искомая вероятность:
$$\frac{2(n-2)!}{n!}\sum_{j=k+1}^{n-2}(n-j-1) = \frac{(n-k-1)(n-k-2)}{n(n-1)}.$$

□

Задача 98. Каждая из n палок одинаковой длины разламывается на две части — «длинную» и «короткую» так, что все «длинные» («короткие») обломки одинаковы по своей длине. $2n$ полученных обломков случайным образом объединяют в пары, каждая из которых образует новую «палку». Найти вероятность того, что:

а) все обломки объединяются в первоначальном порядке, образуя исходные палки;

б) все «длинные» обломки объединяются с «короткими».

Решение. Число всевозможных наборов палок, получаемых после объединения в пары, равно $N = \frac{(2n)!}{2^n \cdot n!}$ — число всевозможных цепочек длины $2n$ нужно поделить на количество перестановок пар в них, образующих палки, т.е. $n!$, и на число способов расположения обломков в каждой из n пар, т.е. 2^n.

а) Из всевозможных наборов палок, получаемых после объединения в пары, лишь в одном наборе обломки объединяются в первоначальном порядке, поэтому искомая вероятность равна
$$p_1 = \frac{1}{N} = \frac{2^n \cdot n!}{(2n)!}.$$

б) Если все «длинные» обломки объединяются с «короткими», то для первого «длинного» обломка имеется n вариантов выбора «короткого» обломка, для следующего «длинного» обломка имеется $(n-1)$ вариантов выбора «короткого» обломка и т.д. Всего имеем $n!$ всевозможных наборов палок, в которых все «длинные» обломки объединяются с «короткими», поэтому искомая вероятность равна
$$p_2 = \frac{n!}{N} = \frac{2^n \cdot (n!)^2}{(2n)!}.$$

□

Задача 99. В урне находится 3 белых и 2 черных шара (и только они). Эксперимент состоит в последовательном извлечении из урны всех шаров по одному наугад без возвращения. Построить вероятностное пространство. Описать σ-алгебру, порожденную случайной величиной X, если:

а) X — число белых шаров, предшествующих первому черному шару;

б) X — число черных шаров среди извлеченных;

в) $X = X_1 + X_2$, где X_1 — число белых шаров, предшествующих первому черному шару, X_2 — число черных шаров, предшествующих белому шару.

Решение. Будем каждому вынутому белому шару ставить в соответствие 0, вынутому черному шару — число 1. Тогда каждый элементарный исход (последовательное извлечение из урны всех шаров) представляет собой упорядоченную совокупность из 5 чисел: три нуля и две единицы. Всего имеем $C_5^2 = 10$ различных элементарных исходов, составляющих $\Omega = \{\omega_i\}_{i=1}^{10}$. В качестве класса событий $\mathcal{F}$ имеет смысл брать множество всех подмножеств Ω, а вероятностная мера однозначно задается значениями $p_i = \mathbb{P}(\omega_i)$, $\sum\limits_{i=1}^{10} p_i = 1$.

а) X — число белых шаров, предшествующих первому черному шару. В случае дискретной случайной величины в качестве событий, образующий порожденную данной с.в. σ-алгебру, можно брать $A_k = \{\omega : X(\omega) = k\}$, $k \in \mathbb{Z}$. В данном случае

$$A_0 = \{\omega : X(\omega) = 0\} = \{11000, 10100, 10010, 10001\}, \ A_3 = \{\omega : X(\omega) = 3\} = \{00011\},$$

$$A_1 = \{\omega : X(\omega) = 1\} = \{01100, 01010, 01001\}, \ A_2 = \{\omega : X(\omega) = 2\} = \{00110, 00101\}.$$

Отсюда $\sigma(X) = \sigma\{A_0, A_1, A_2, A_3\}$.

б) X — число черных шаров среди извлеченных. В этом случае $\{\omega : X = 2\} = \Omega$, т.к. X неслучайная величина и порождает тривиальный класс событий.

в) $X = X_1 + X_2$, где X_1 — число белых шаров, предшествующих первому черному шару, X_2 — число черных шаров, предшествующих белому шару. Из определения следует, что для каждого $\omega \in \Omega$, либо $X_1(\omega) = 0$, либо $X_2(\omega) = 0$. Имеем

$$A_1 = \{\omega : X(\omega) = 1\} = \{01100, 01010, 01001, 10100, 10010, 10001\},$$

$$A_2 = \{\omega : X(\omega) = 2\} = \{00110, 00101, 11000\}, \ A_3 = \{\omega : X(\omega) = 3\} = \{00011\}.$$

Отсюда $\sigma(X) = \sigma\{A_1, A_2, A_3\}$.

□

Задача 100. Найти математическое ожидание $\mathbb{E}Z$, где Z — k-ая по величине из координат n точек, взятых наудачу на отрезке $[0; 1]$ ($k \leqslant n$).

Решение. Общие обозначения и факты: пусть X — непрерывная случайная величина с функциями распределения и плотности $F_X(x)$ и $f_X(x)$, $\{X_i\}_{i=1}^n$ — простая выборка случайной величины X (результаты n ее независимых измерений), $\{X_{(i)}\}_{i=1}^n$ — вариационный ряд, полученный из выборки путем ранжирования ее элементов: $X_{(1)} < \ldots < X_{(n)}$. В нашем случае $Z = X_{(k)}$ — k-ая порядковая статистика, функция распределения которой

$$F_k^{(\cdot)}(x) = \mathbb{P}\{X_{(k)} < x\} = \sum_{i=k}^{n} C_n^i [F_X(x)]^i [1 - F_X(x)]^{n-i}$$

равна вероятности того, что не менее k элементов выборки $\{X_{(i)}\}_{i=1}^n$ имеют значения, меньшие x. Для плотности распределения $X_{(k)}$ получим

$$f_k^{(\cdot)}(x) = \frac{dF_k^{(\cdot)}(x)}{dx} = nC_{n-1}^{k-1}[F_X(x)]^{k-1}[1 - F_X(x)]^{n-k} f_X(x)$$

(легко видеть, что здесь после дифференцирования все члены полученной суммы, кроме первого, сокращаются). По условию задачи $X \in \mathrm{Gl}(0,1)$ (равномерное распределение на отрезке $[0,1]$) и

$$f_k^{(\cdot)}(x) = \begin{cases} nC_{n-1}^{k-1}x^{k-1}[1-x]^{n-k}, & \text{если } 0 \leqslant x \leqslant 1, \\ 0, & \text{в остальных случаях }. \end{cases}$$

В результате для $\mathbb{E}Z$ получаем

$$\mathbb{E}Z = \int\limits_0^1 x\, f_k^{(\cdot)}(x)\, dx = \frac{k}{n+1}.$$

$\square$

Задача 101. Случайный вектор $W = (W_1, W_2, \ldots, W_n)'$ имеет плотность распределения

$$f_W(w) = \begin{cases} n!, & \text{если } 0 \leqslant w_1 \leqslant w_2 \leqslant \ldots \leqslant w_n \leqslant 1, \\ 0, & \text{в остальных случаях }. \end{cases}$$

Найти распределение вектора $U = (U_1, U_2, \ldots, U_n)'$, если

$$U_1 = W_1, \quad U_2 = W_2 - W_1, \ldots, U_n = W_n - W_{n-1}.$$

Решение. В общем случае:

$$U = \varphi(W) \;\Rightarrow\; f_U(u) = f_W\big(\varphi^{-1}(u)\big)\Big|J\Big(\frac{\partial W}{\partial U}\Big)\Big|, \quad J(\cdot) = \det\Big(\frac{\partial \varphi_i^{-1}(u)}{\partial u_j}\Big).$$

В нашем случае $w_1 = u_1$, $w_2 = u_1 + u_2$, …, $w_n = u_1 + \ldots + u_n$ и

$$J(\cdot) = \det\Big(\frac{\partial \varphi_i^{-1}(u)}{\partial u_j}\Big) = 1. \quad \text{Отсюда}$$

$$f_U(u) = \begin{cases} n!, & \text{если } \forall u_i \geqslant 0,\ \sum\limits_{i=1}^n u_i \leqslant 1; \\ 0, & \text{в прочих случаях }. \end{cases}$$

$\square$

Задача 102. Пусть случайная величина X_n принимает значения $\sqrt{n}$, 0 и $-\sqrt{n}$ с вероятностями $\frac{1}{2n}$, $1-\frac{1}{n}$, $\frac{1}{2n}$ соответственно. Выполняется ли для последовательности независимых случайных величин X_1, X_2, … закон больших чисел?

Решение. $\mathbb{E}X_i = 0$; $\operatorname{Var} X_i = 1$.

$$\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n} X_i\Big] = 0; \quad \operatorname{Var}\Big[\frac{1}{n}\sum_{i=1}^{n} X_i\Big] = \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var} X_i = \frac{1}{n}.$$

Из неравенства Чебышева:

$$\mathbb{P}\Big(\Big|\frac{1}{n}\sum_{i=1}^{n} X_i\Big| \geqslant \varepsilon\Big) \leqslant \frac{\operatorname{Var}\Big[\frac{1}{n}\sum\limits_{i=1}^{n} X_i\Big]}{\varepsilon^2} = \frac{1}{n\varepsilon^2} \xrightarrow{n\to\infty} 0$$

следует выполнение ЗБЧ.

Либо сразу заметить, что ЗБЧ выполняется в силу достаточного условия — равномерной ограниченности дисперсий $\operatorname{Var} X_i$. □

Задача 103. В поселке N жителей, каждый из которых в среднем n раз в месяц ездит в город, выбирая дни поездки независимо от остальных. Поезд из поселка в город идет один раз в сутки. Какова должна быть вместимость поезда для того, чтобы он переполнился с вероятностью, не превышающей заданного числа β?

Решение. Обозначим X — число пассажиров (в день), p — вероятность поездки жителя (в данный день), a — вместимость поезда. $X \in \mathrm{Bi}(N, p)$, $p = n/30$ (30 — число дней в м-це).

Условие для a: $\mathbb{P}\{X > a\} < \beta$. Нормальное приближение:

$$\frac{X - \mathbb{E}X}{\sqrt{\operatorname{Var} X}} \xrightarrow{d} T \in \mathcal{N}(0,1), \quad \mathbb{E}X = \frac{Nn}{30}, \ \operatorname{Var} X = \frac{Nn}{30}\Big(1 - \frac{n}{30}\Big).$$

Условие на a принимает приближенный вид $\mathbb{P}\{X > a\} < \beta \Rightarrow$

$$\mathbb{P}\Big\{\frac{X - \mathbb{E}X}{\sqrt{\operatorname{Var} X}} > \frac{a - \mathbb{E}X}{\sqrt{\operatorname{Var} X}}\Big\} < \beta \Rightarrow 1 - \Phi^*\Big(\frac{a - \mathbb{E}X}{\sqrt{\operatorname{Var} X}}\Big) < \beta \Rightarrow$$

$$a > \mathbb{E}X + \sqrt{\operatorname{Var} X}\,\arg(1-\beta).$$

□

Задача 104 (парадокс транзитивности)**.** Будем говорить, что случайная величина X больше по вероятности случайной величины Y, если $\mathbb{P}(X > Y) > \mathbb{P}(X \leq Y)$. Пусть известно, что для случайных величин X, Y, Z, W выполнена следующая цепочка равенств:

$$\mathbb{P}(X > Y) = \mathbb{P}(Y > Z) = \mathbb{P}(Z > W) = \alpha > \frac{1}{2}.$$

Верно ли, что X больше по вероятности W и почему?

Решение. Покажем, что не всегда это верно. Рассмотрим дискретные случайные величины X, Y, Z, W, где X и W одинаково распределены,

$$\mathbb{P}(X = 18) = \mathbb{P}(X = 10) = \mathbb{P}(X = 9) = \mathbb{P}(X = 8) = \mathbb{P}(X = 7) = \mathbb{P}(X = 5) = \frac{1}{6},$$

$$\mathbb{P}(Y = 17) = \mathbb{P}(Y = 16) = \mathbb{P}(Y = 15) = \mathbb{P}(Y = 4) = \mathbb{P}(Y = 3) = \mathbb{P}(Y = 2) = \frac{1}{6},$$

$$\mathbb{P}(Z = 14) = \mathbb{P}(Z = 13) = \mathbb{P}(Z = 12) = \mathbb{P}(Z = 11) = \mathbb{P}(Z = 6) = \mathbb{P}(Z = 1) = \frac{1}{6},$$

и пусть случайные величины X, Y, Z, W попарно независимы. Тогда

$$\mathbb{P}(X > Y) = \mathbb{P}(Y > Z) = \mathbb{P}(Z > W) = \frac{21}{36} > \frac{1}{2}.$$

При этом $\mathbb{P}(X > W) = \frac{15}{36} < \frac{21}{36} = \mathbb{P}(X \leq W)$, т.е. величина X по вероятности не больше случайной вечины W. $\square$

Задача 105 (парадокс Стефана Банаха)**.** В двух спичечных коробках имеется по n спичек. На каждом шаге наугад выбирается коробок, и из него удаляется (используется) одна спичка. Найти вероятность того, что в момент, когда один из коробков опустеет, в другом останется k спичек.

Решение. *Первый вариант*

Событие, удовлетворяющее условию задачи — выбран пустой коробок, а в другом коробке имеется k спичек.

Рассмотрим производящую функцию $P(w, z) = \sum\limits_{m,n} p_{m,n} w^m z^n$, где $p_{m,n}$ есть вероятность, начав с m спичек в одной коробке и n — в другой, получить обе пустые коробки, когда впервые выбирается пустая коробка.

При $m = n = 0$ имеем $p_{0,0} = 1$. Получим

$$P(w, z) = 1 + \frac{1}{2}(w + z)P(w, z) \Rightarrow P(w, z) = \Big(1 - \frac{1}{2}(w + z)\Big)^{-1}.$$

Разлагая в ряд по степеням w и z, находим

$$P(w, z) = 1 + \sum_{k=1}^{\infty} \frac{(w + z)^k}{2^k}.$$

Отсюда $p_{m,n} = \frac{C_{m+n}^n}{2^{m+n}}$.

Далее вводим производящую функцию $P_k(w, z) = \sum\limits_{m,n} p_{k,m,n} w^m z^n$, где $p_{k,m,n}$ есть вероятность, начав с m спичек в одной коробке и n — в другой, в момент выбрасывания первой пустой коробки иметь вторую коробку с k спичками. Тогда

$$P_k(w, z) = \frac{1}{2}(w^k + z^k)P(w, z)$$

$$\Rightarrow\ p_{k,m,n} = \frac{C^m_{m+n-k} + C^n_{m+n-k}}{2^{1+m+n-k}}.$$

Для данной задачи искомая вероятность равна

$$p_{k,n,n} = \frac{C^n_{2n-k}}{2^{2n-k}}.$$

Второй вариант

Событие, удовлетворяющее условию задачи — из выбранной коробки взяли последнюю спичку, а в другом коробке имеется k спичек.

Рассмотрим процесс изъятия спичек из коробок как последовательность нулей и единиц (например, нули соответствуют спичкам первой коробки, единицы – второй коробки) длины $2n$, с числом нулей и единиц равным n. Общее число возможных исходов равно $N = C^n_{2n}$.

Исходы, удовлетворяющие условию задачи, устроены так: если сначала пустеет первый коробок, то на $2n-k$-м месте стоит нуль, а в первых $2n-k-1$ позициях имеется $n-1$ нулей в любом порядке. Аналогично когда сначала пустеет второй коробок. Число таких исходов равно $m = 2 \cdot C^{n-1}_{2n-k-1}$. Искомая вероятность равна

$$p = \frac{m}{N} = \frac{2 \cdot C^{n-1}_{2n-k-1}}{C^n_{2n}} = \frac{2 \cdot n!\, n!\, (2n-k-1)!}{(2n)!\, (n-1)!\, (n-k)!} = \frac{n!\, (2n-k-1)!}{(2n-1)!\, (n-k)!} = \frac{C^k_n}{C^k_{2n-1}}.$$

□

Задача 106 (парадокс Монти–Холла). Представьте, что вы стали участником игры, в которой находитесь перед тремя дверями. Ведущий поместил за одной из трех пронумерованных дверей автомобиль, а за двумя другими дверями — по козе (козы тоже пронумерованы) случайным образом — это значит, что все $3! = 6$ вариантов расположения автомобиля и коз за пронумерованными дверями равновероятны). У вас нет никакой информации о том, что за какой дверью находится. Ведущий говорит: «Сначала вы должны выбрать одну из дверей. После этого я открою одну из оставшихся дверей (при этом если вы выберете дверь, за которой находится автомобиль, то я с вероятностью $1/2$ выберу дверь, за которой находится коза номер 1, и с вероятностью $1 - 1/2 = 1/2$ дверь, за которой находится коза номер 2). Затем я предложу вам изменить свой первоначальный выбор и выбрать оставшуюся закрытую дверь вместо той, которую вы выбрали сначала. Вы можете последовать моему совету и выбрать другую дверь, либо подтвердить свой первоначальный выбор. После этого я открою дверь, которую вы выбрали, и вы выиграете то, что находится за этой дверью.» Вы выбираете дверь номер 3. Ведущий открывает дверь номер 1 и показывает, что за ней находится коза. Затем ведущий предлагает вам выбрать дверь номер 2. Увеличатся ли ваши шансы выиграть автомобиль, если вы последуете его совету?

Решение. Обозначим с.в. $C \in \{1,2,3\}$ — дверь, за которой машина, $S \in \{1,2,3\}$ — дверь, выбранная игроком, $H \in \{1,2,3\}$ — дверь, открытая ведущим.

Для всякого $c \in \{1,2,3\}$ имеем $\mathbb{P}(C=c)=1/3$, а также по условию задачи $\mathbb{P}(C\,|\,S)=\mathbb{P}(C)$. Далее,

$$\mathbb{P}(H\,|\,(C,S))=\begin{cases}0, & \text{если } H=S \text{ (ведущий не открывает дверь, выбранную игроком)}\\ 0, & \text{если } H=C \text{ (ведущий не откроет дверь с машиной)}\\ 1/2, & \text{если } S=C \text{ — по условию задачи}\\ 1, & \text{если } H\neq C \text{ и } S\neq C\end{cases}$$

$$\mathbb{P}(C\,|\,(H,S))=\frac{\mathbb{P}(H\,|\,(C,S))\cdot\mathbb{P}(C\,|\,S)}{\mathbb{P}(H\,|\,S)},$$

где $\mathbb{P}(H\,|\,S)=\sum\limits_{C=1}^{3}\mathbb{P}(H\,|\,(C,S))\mathbb{P}(C)=\sum\limits_{C=1}^{3}\mathbb{P}(H\,|\,(C,S))\mathbb{P}(C\,|\,S)$.

Поэтому в условиях задачи при смене двери вероятность выигрыша

$$\mathbb{P}(C=2\,|\,(H=1,\ S=3))=\frac{1\cdot 1/3}{0\cdot 1/3+1\cdot 1/3+1/2\cdot 1/3}=\frac{2}{3}.$$

Соответственно без смены двери вероятность выигрыша

$$\mathbb{P}(C=3\,|\,(H=1,\ S=3))=\frac{1/2\cdot 1/3}{0\cdot 1/3+1\cdot 1/3+1/2\cdot 1/3}=\frac{1}{3}.$$

Значит, если игрок последует за советом ведущего, его шансы увеличатся. □

Задача 107 (переход к полярным координатам, якобиан преобразования)**.** Спортсмен стреляет по круговой мишени. Вертикальная и горизонтальная координаты точки попадания пули (при условии, что центр мишени – начало координат) – независимые случайные величины, каждая с распределением $N(0,1)$. Покажите, что расстояние от точки попадания до центра имеет плотность распределения вероятностей $r\exp\left(-r^2/2\right)$ для $r\geq 0$. Найдите медиану этого распределения.

Решение. Обозначим (X,Y) — случайный вектор, состоящий из горизонтальной и вертикальной координат точки попадания пули. Его функция распределения имеет вид

$$F(x,y)=\mathbb{P}(X\leq x, Y\leq y)=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y}\frac{1}{2\pi}\exp\Big\{-\frac{X^2+Y^2}{2}\Big\}\,dX\,dY.$$

Подынтегральная функция является плотностью случайного вектора (X,Y). Переход к полярным координатам $X=r\cos\varphi$, $Y=r\sin\varphi$, имеющий якобиан преобразования $|J|=r$, приводит к плотности распределения вида $p(r,\varphi)=\frac{1}{2\pi}re^{-r^2/2}$.

Для случайной величины $Z=\sqrt{X^2+Y^2}=r$, равной расстоянию от точки попадания до центра, функция распределения имеет вид

$$\mathbb{P}(Z\leq R)=\int\limits_{r\leq R}\int\limits_{\varphi\in[0,2\pi]}p(r,\varphi)\,dr\,d\varphi=\int\limits_{r\leq R}re^{-r^2/2}\,dr,$$

откуда получим искомую плотность $r\exp\left(-r^2/2\right)$, где $r \geq 0$.

Медиану m этого распределения находим из условия

$$\int\limits_0^m re^{-r^2/2}\,dr = \frac{1}{2} \;\Rightarrow\; 1 - e^{-m^2/2} = \frac{1}{2} \;\Rightarrow\; m = \sqrt{\ln 4} \approx 1.177.$$

□

Задача 108 (распределение Коши)**.** Радиоактивный источник испускает частицы в случайном направлении (при этом все направления равновероятны). Источник находится на расстоянии d от фотопластины, которая представляет собой бесконечную вертикальную плоскость.

А) При условии, что частица попадает в плоскость, покажите, что горизонтальная координата точки попадания (если начало координат выбирается в точке, ближайшей к источнику) имеет плотность распределения:

$$p\left(x\right) = \frac{d}{\pi\left(d^2 + x^2\right)}.$$

Это распределение известно как *распределение Коши.*

Б) Можно ли вычислить среднее (математическое ожидание) этого распределения?

Решение. **А)** Введем случайную величину X — горизонтальная координата точки попадания. Тогда

$$\mathbb{P}(|X| \leq x) = \frac{\arctan x/d}{\pi/2}.$$

Для функции распределения $F(x)$ с.в. X выполнено соотношение $\mathbb{P}(|X| \leq x) = F(x) - F(-x)$, отсюда находим

$$\frac{d}{dx}\mathbb{P}(|X| \leq x) = 2p(x) = \frac{2}{\pi} \cdot \frac{d}{d^2 + x^2},$$

откуда вытекает искомая формула для плотности $p(x)$.

Б) Чтобы было определено (конечное или бесконечное) математическое ожидание $\mathbb{E}X$, требуется существование $\mathbb{E}X_+$, $\mathbb{E}X_-$ и ограниченность хотя бы одного из них, где $X_+ = X \cdot \mathcal{I}_{\{X \geq 0\}}$, $X = X_+ - X_-$. Имеем

$$\mathbb{E}X_+ = \int\limits_0^\infty xp(x)\,dx = \frac{d}{\pi}\int\limits_0^\infty \frac{x\,dx}{d^2 + x^2} = +\infty$$

Аналогично $\mathbb{E}X_- = +\infty$. Поэтому не существует среднего для распределения Коши.

□

Задача 109 (рекорды)**.** Пусть $X_1, X_2, \ldots$ - независимые случайные величины с одной и той же плотностью распределения вероятностей $p(x)$. Будем говорить, что наблюдается рекордное значение в момент времени n>1, если $X_n > \max[X_1, ..., X_{n-1}]$. Докажите следующие утверждения.

А) Вероятность того, что рекорд зафиксирован в момент времени n, равна $1/n$.

Б) Математическое ожидание числа рекордов до момента времени n равно

$$\sum_{1<i\le n} \frac{1}{i} \sim \ln n.$$

В) Пусть Y_n — случайная величина, принимающая значение 1, если в момент времени n зафиксирован рекорд, и значение 0 – в противном случае. Тогда случайные величины $Y_1, Y_2, \ldots$ независимы в совокупности.

Г) Дисперсия числа рекордов до момента времени n равна

$$\sum_{1<i\le n} \frac{i-1}{i^2} \sim \ln n.$$

Д) Если T – момент появления первого рекорда после момента времени 1, то $ET = \infty$.

Решение. **А)** Обозначим $F(x) = \mathbb{P}(X_i < x)$ — функция распределения с.в. X_i, ее плоность равна $p(x) = F'(x)$. Тогда функция распреления с.в. $Z_{n-1} = \max[X_1, ..., X_{n-1}]$ равна

$$\mathbb{P}(Z_{n-1} < x) = \mathbb{P}(X_1 < x) \cdot \ldots \cdot \mathbb{P}(X_{n-1} < x) = F^{n-1}(x).$$

Тогда плотность данной с.в. имеет вид $p_{n-1}(x) = \frac{d}{dx}F^{n-1}(x) = (n-1)p(x)F^{n-1}(x)$. Вероятность появления рекорда равна

$$\mathbb{P}(X_n > Z_{n-1}) = \int\limits_{-\infty}^{\infty} p_{n-1}(y)\,dy \int\limits_{y}^{\infty} p(x)\,dx =$$

$$= \int\limits_{-\infty}^{\infty} (n-1)p(y)F^{n-1}(y)[1-F(y)]\,dy = (n-1)\int\limits_{-\infty}^{\infty} (F^{n-2}(y) - F^{n-1}(y))\,dF(y) =$$

$$= (n-1)\int\limits_{0}^{1} [z^{n-2} - z^{n-1}]\,dz = (n-1)\Big[\frac{1}{n-1} - \frac{1}{n}\Big] = \frac{1}{n}.$$

В) Сначала покажем, что любые две с.в. Y_n и Y_m независимы, $m < n$. Достаточно показать, что $\mathbb{P}(Y_m = 1, Y_n = 1) = \mathbb{P}(Y_m = 1)\mathbb{P}(Y_n = 1)$. Рассмотрим

$$\mathbb{P}(X_m < x, X_m > Z_{m-1}) = \int\limits_{-\infty}^{x} p_{m-1}(y)\,dy \int\limits_{y}^{x} p(s)\,ds =$$

$$= \int\limits_{-\infty}^{x} (m-1)p(y)F^{m-2}(y)[F(x)-F(y)]\,dy = (m-1)\int\limits_{-\infty}^{x}[F(x)F^{m-2}(y)-F^{m-1}(y)]\,dF(y) =$$

$$= (m-1)\int\limits_{0}^{F(x)}[F(x)z^{m-2}-z^{m-1}]\,dz = F(x)\frac{F^{m-1}(x)}{m-1} - \frac{F^m(x)}{m} = \frac{F^m(x)}{m}.$$

Далее рассмотрим

$$\mathbb{P}(Z_{n-1} < x, X_m > Z_{m-1}) = \mathbb{P}(X_{n-1} < x, \ldots, X_m < x, X_m > Z_{m-1}) =$$

$$= \mathbb{P}(X_{n-1} < x)\ldots\mathbb{P}(X_{m+1} < x)\mathbb{P}(X_m < x, X_m > Z_{m-1}) = F^{n-1-m}(x)\frac{F^m(x)}{m} = \frac{F^{n-1}(x)}{m}.$$

Отсюда находим (аналогично предыдущему пункту)

$$\mathbb{P}(Y_m = 1, Y_n = 1) = \mathbb{P}(X_n > Z_{n-1}, X_m > Z_{m-1}) = \frac{1}{m}\frac{1}{n} = \mathbb{P}(Y_n = 1)\mathbb{P}(Y_m = 1).$$

В общем случае, когда имеем с.в. Y_n, Y_{m_1}, …, Y_{m_k}, где $n > m_1 > \ldots > m_k$, аналогично можно получить

$$\mathbb{P}(Z_{n-1} < x, X_{m_1} > Z_{m_1-1}, \ldots, X_{m_k} > Z_{m_k-1}) = \frac{F^{n-1}(x)}{m_1 \ldots m_k}$$

$$\Rightarrow\ \mathbb{P}(Y_n = 1, Y_{m_1} = 1, \ldots, Y_{m_k} = 1) = \mathbb{P}(X_n > Z_{n-1}, X_{m_1} > Z_{m_1-1}, \ldots, X_{m_k} > Z_{m_k-1}) =$$

$$= \frac{1}{n}\frac{1}{m_1}\ldots\frac{1}{m_k} = \mathbb{P}(Y_n = 1)\mathbb{P}(Y_{m_1} = 1)\ldots\mathbb{P}(Y_{m_k} = 1),$$

что означает независимость в совокупности с.в. Y_1, Y_2, …, Y_n.

Б) Число рекордов до момента времени n равно $M_n = Y_2 + \ldots + Y_n$. Его среднее значение

$$\mathbb{E}M_n = \sum_{1<i\le n}\frac{1}{i} = S(n).$$

Для интеграла $\int_1^n \frac{dx}{x} = \ln n$ его верхняя и нижняя суммы равны $S(n-1)$ и $S(n)-1$, откуда

$$\ln(n+1) < S(n) < \ln n + 1\ \Rightarrow\ \frac{S(n)}{\ln n} \xrightarrow{n\to\infty} 1,$$

т.е. $\sum\limits_{1<i\le n}\frac{1}{i} \sim \ln n$.

Г) Учитывая независимость с.в. Y_2, …, Y_n, дисперсия числа рекордов M_n равна

$$\operatorname{Var} M_n = \sum_{1<i\le n}\operatorname{Var} Y_i = \sum_{1<i\le n}\frac{i-1}{i^2} = \sum_{1<i\le n}\left(\frac{1}{i} - \frac{1}{i^2}\right) = \ln n + c_n,$$

где $|c_n| < 1 + \sum\limits_{1<i\le n}\frac{1}{i^2} < 2$. Отсюда

$$\frac{\operatorname{Var} M_n}{\ln n} \xrightarrow{n\to\infty} 1\ \Rightarrow\ \operatorname{Var} M_n \sim \ln n.$$

Д) Если T – момент появления первого рекорда после момента времени 1, то

$$\mathbb{P}(T=i)=\mathbb{P}(Y_2=\ldots=Y_{i-1}=0,Y_i=1)=\Big(1-\frac{1}{2}\Big)\ldots\Big(1-\frac{1}{i-1}\Big)\frac{1}{i}=\frac{1}{i(i-1)}$$

$$\Rightarrow\ \mathbb{E}T=\sum_{i=2}^{\infty}i\cdot\mathbb{P}(T=i)=\sum_{i=2}^{\infty}\frac{1}{i-1}=\lim_{n\to\infty}(\ln n+r_n)=\infty,$$

где $|r_n|<1$.

$\square$

2. Раздел 1. Вероятностные методы в комбинаторике

Вероятностные методы в комбинаторике (идеи П.Эрдеша (оценка чисел Рамсея), приложения предельных теорем теории вероятностей в асимптотической комбинаторике)

Задача 110. На турнир приехало n игроков. Каждая пара игроков, согласно регламенту турнира, должна провести одну встречу (ничьих быть не может). Пусть

$$C_n^k \cdot (1-2^{-k})^{n-k} < 1.$$

Докажите, что тогда игроки могли сыграть так, что для каждого множества из k игроков найдется игрок, который побеждает их всех.

Решение. Введем на множестве всех турниров равномерную меру, т.е. считаем, что все $2^{C_n^2}$ турнира равновероятны. Введем A_K — событие, состоящее в том, что не существует игрока, побеждающего всех игроков из множества K. Докажем, что

$$\mathbb{P}\Big(\bigcup_{K\subset\{1,..,n\},|K|=k} A_K\Big) \leqslant C_n^k \cdot (1-2^{-k})^{n-k}.$$

I. Число различных множеств $K: \ K \subset \{1,\dots,n\}, |K| = k$ равно C_n^k, а все соответствующие события A_K равновероятны.

II. Найдем $\mathbb{P}(A_K)$, например, для $K = \{1,\dots,k\}$. Среди всех $C_n^2 = \frac{n(n-1)}{2}$ встреч рассмотрим $k(n-k)$ встреч, в которых один из игроков из множества K, а другой игрок — не из K. Эти пары имеют вид:

   1) $(k+1,1), (k+1,2), \dots, (k+1,k)$,

   2) $(k+2,1), (k+2,2), \dots, (k+2,k)$,

   $n-k$) $(n,1), (n,2), \dots, (n,k)$.

   Событие A_K заключается в том, что в каждом выписанном наборе из k встреч есть хотя бы одна встреча, в которой выигрывает игрок из множества K. Число таких исходов в этих k встречах равно 2^k-1(любой исход, кроме одного — когда все игроки из множества K побеждены). Тогда число возможных исходов в рассматриваемых $k(n-k)$ встречах равно $(2^k-1)^{n-k}$.

   Остальные $C_n^2 - k(n-k)$ встреч для события A_K могут иметь любой исход. Отсюда находим

$$\mathbb{P}(A_K) = \frac{(2^k-1)^{n-k}\cdot 2^{C_n^2-k(n-k)}}{2^{C_n^2}} = \frac{(2^k-1)^{n-k}}{2^{k(n-k)}} = (1-2^{-k})^{n-k}.$$

Таким образом,

$$\mathbb{P}\Big(\bigcup_{K\subset\{1,..,n\},|K|=k} A_K\Big) \leqslant C_n^k \cdot \mathbb{P}(A_K) = C_n^k \cdot (1-2^{-k})^{n-k}.$$

Из условия задачи вытекает

$$\mathbb{P}\Big(\bigcup_{K\subset\{1,..,n\},|K|=k} A_K\Big) < 1,$$

откуда

$$\mathbb{P}\Big(\overline{\bigcup_{K\subset\{1,..,n\},|K|=k} A_K}\Big) > 0.$$

Это и означает, что игроки могли сыграть так, что для каждого множества из k игроков найдется игрок, который побеждает их всех.

□

Задача 111. Рассмотрим матрицу $n\times n$, составленную из лампочек, каждая из которых любо включена ($a_{ij}=1$), либо выключена ($a_{ij}=-1$). Предположим, что для каждой строки и каждого столбца имеется переключатель, поворот которого ($x_i=-1$ для строки i и $y_j=-1$ для столбца j) переключает все лампочки в соответствующей линии: с «вкл.» на «выкл.» и с «выкл.» на «вкл.». Тогда для любой начальной конфигурации лампочек можно установить такое положение переключателей, что разность между числом включенных и выключенных лампочек будет не меньше $(\sqrt{2/\pi}+o(1))n^{3/2}$.

Решение. Пусть $y_1,\dots,y_n$ — независимые одинаково распределенные с.в., с законом распределения $y_j=\begin{cases}1, & p=1/2\\ -1, & p=1/2\end{cases}$. Введем с.в. $R_i=\sum\limits_{j=1}^{n}a_{ij}y_j$, $i=1,\dots,n$ и $R=\sum\limits_{i=1}^{n}|R_i|$. Т.к. $a_{ij}=1$ либо $a_{ij}=-1$, каждая с.в. $a_{ij}y_j$ распределена так же, как y_j, поэтому R_i распределена так же, как $S_n=\sum\limits_{j=1}^{n}y_j$.

По смыслу R_i равна разности между числом включенных и выключенных лампочек в i-й строке при переключателях $y_1,\dots,y_n$, а R — это разность между числом включенных и выключенных лампочек в матрице при переключателях $y_1,\dots,y_n$ и $x_1,\dots,x_n$, таких что $x_iR_i=|R_i|$.

Найдем распределение для с.в. $|R_i|=|S_n|$. Применим ЦПТ для $S_n=\sum\limits_{j=1}^{n}y_j$. Т.к. $\mathbb{E}y_j=0$, $\operatorname{Var}y_j=1$, имеем

$$S_n=(\xi+o(1))\sqrt{n},\quad \xi\sim N(0,1)$$
$$\Rightarrow\quad |S_n|=(|\xi|+o(1))\sqrt{n}.$$

Т.к. $\mathbb{E}|\xi|=2\int\limits_0^{+\infty}\frac{x\cdot e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx=\sqrt{\frac{2}{\pi}}$, имеем

$$\mathbb{E}|S_n|=(\sqrt{2/\pi}+o(1))\sqrt{n}.$$

Также согласно ЦПТ

$$|S_n|-\mathbb{E}|S_n|=(\eta+o(1))\sqrt{\operatorname{Var}|S_n|},\quad \eta\sim N(0,1)$$

$$\operatorname{Var}|\xi|=1-\frac{2}{\pi}\Rightarrow\quad |S_n|=(\sqrt{2/\pi}+o(1))\sqrt{n}+\eta\cdot\sqrt{n(1-2/\pi)}.$$

Выбирая такие события, для которых $\eta = o(1)$ (при $n \to \infty$), получим

$$\mathbb{P}\big(|S_n| = (\sqrt{2/\pi} + o(1))\sqrt{n}\big) > 0.$$

Учитывая равенство распределений R_i и S_n, для с.в. R аналогично получим

$$\mathbb{P}\big(R = (\sqrt{2/\pi} + o(1))n^{3/2}\big) > 0.$$

Это означает, что при $n \to \infty$ найдутся конфигурации $\{x_1, \ldots, x_n\}$, $\{y_1, \ldots, y_n\}$, на которых

$$R = (\sqrt{2/\pi} + o(1))n^{3/2},$$

что завершает доказательство.

□

Задача 112. Поверхность некоторой шарообразной планеты состоит из океана и суши (множество мелких островков). Суша занимает больше половины площади планеты. Также известно, что суша есть множество, принадлежащее борелевской σ-алгебре на сфере. На планету хочет совершить посадку космический корабль, сконструированный так, что концы всех шести его ножек лежат на поверхности планеты. Посадка окажется успешной, если не меньше четырех ножек из шести окажутся на суши. Возможна ли успешная посадка корабля на планету?

Решение.

1. Зафиксируем шесть точек корабля на сфере в качестве начального положения.
2. Считаем площадь планеты равной единице.
3. Выберем в качестве пространства элементарных событий множество всех движений на сфере относительного данного начального положения, и каждой из шести ножек корабля поставим в соответствие случайную величину ξ_i, $i = 1, \ldots, 6$, так что для заданного элементарного события ω
   - $\xi_i(\omega) = 1$, если при соответствующем движении планеты i-я ножка корабля окажется на суше;
   - $\xi_i(\omega) = 0$, если при соответствующем движении планеты i-я ножка корабля попадет в океан.
4. Тогда математическое ожидание $\mathbb{E}\xi_i$ равно площади суши на планете, и по условию задачи $\mathbb{E}\xi_i > \frac{1}{2}$. Поэтому
$$\mathbb{E}(\xi_1 + \ldots + \xi_6) > 3.$$
5. Поскольку случайная величина $\xi = \xi_1 + \ldots + \xi_6$ принимает только целые значения, найдется элементарное событие ω, для которого $\xi(\omega) \geqslant 4$.

Значит, при движении относительного начального состояния, которому соответствует элементарное событие ω, количество ножек корабля, попадающих на сушу, равно $\xi(\omega) \geqslant 4$, поэтому возможна успешная посадка корабля на планету.

□

Задача 113. Покажите, что можно так раскрасить в два цвета ребра полного графа с n вершинами (т.е. графа (без петель), в котором любые две различные вершины соединены одним ребром), что любой его полный подграф с m вершинами, где $2C_n^m(1/2)^{C_m^2} < 1$, имеет ребра разного цвета.

Решение. 1. Введем на множестве графов из n вершин с ребрами, раскрашенными в два цвета, равномерную меру. Число ребер в полном графе равно C_n^2, поэтому вероятность одной раскраски графа равна

$$p = \frac{1}{2^{C_n^2}}.$$

2. Пронумеруем вершины графа $\{1, 2, \ldots, n\}$. Для всякого подмножества $K \subset \{1, 2, \ldots, n\}$, состоящего из m элементов, введем событие A_K, заключающееся в том, что полный подграф с вершинами из множества K имеет ребра одного цвета. В таком случае заданные C_m^2 ребер графа имеют одинаковый цвет — один из двух, а остальные $C_n^2 - C_m^2$ ребра могут иметь любой цвет. Поэтому число всевозможных таких графов равно

$$N_m = 2 \cdot 2^{C_n^2 - C_m^2} \Rightarrow \mathbb{P}(A_K) = N_m \cdot p = \frac{2}{2^{C_m^2}}.$$

3. Событие, заключающееся в том, что существует полный подграф с m вершинами, имеющий ребра одного цвета, имеет вид

$$A_m = \bigcup_{K:\, \|K\|=m} A_K.$$

Т.к. число всевозможных $K : \|K\| = m$, равно C_n^m, получим

$$\mathbb{P}(A_m) \leqslant \sum_{K:\, \|K\|=m} \mathbb{P}(A_K) = 2C_n^m(1/2)^{C_m^2}.$$

4. Из условия задачи следует $\mathbb{P}(A_m) < 1$, откуда

$$\mathbb{P}(\overline{A_m}) > 0.$$

Событие $\overline{A_m}$ заключается в том, что не существует полный подграф с m вершинами, имеющий ребра одного цвета.

Это и означает, что существует раскраска ребер полного графа из n вершин в два цвета, так что любой его полный подграф с m вершинами, где $2C_n^m(1/2)^{C_m^2} < 1$, имеет ребра разного цвета.

□

Задача 114 (устойчивые системы большой размерности; В.И. Опойцев, 1985)**.** Из курсов функционального анализа и вычислительной математики хорошо известно, что если спектральный радиус матрицы $A = \|a_{ij}\|_{i,j=1}^n$ меньше единицы: $\rho(A) < 1$, то итерационный процесс $\vec{x}^{n+1} = A\vec{x}^n + \vec{b}$ (СОДУ $\dot{\vec{x}} = -\vec{x} + A\vec{x} + \vec{b}$), вне зависимости от точки старта $\vec{x}^0$, сходится к единственному решению уравнения $\vec{x}^* = A\vec{x}^* + \vec{b}$. Скажем, если $\|A\| = \max_i \sum_j |a_{ij}| < 1$, то и $\rho(A) < 1$ (обратное, конечно, не верно). Предположим, что существует такое маленькое $\varepsilon > 0$, что

$$\frac{1}{n}\sum_{i,j}|a_{ij}| < 1 - \varepsilon. \quad (S)$$

Очевидно, что отсюда, тем более, не следует: $\rho(A) < 1$. Тем не менее, введя на множестве матриц, удовлетворяющих условию (S), равномерную меру, покажите, что относительная мера тех матриц (удовлетворяющих условию (S)), для которых спектральный радиус не меньше единицы, стремится к нулю с ростом n (ε — фиксировано и от n не зависит).

Решение. Поскольку условие (S) не изменится, если любой из элементов a_{ij} матрицы заменить на $-a_{ij}$, при доказательстве достаточно рассматривать матрицы с неотрицательными элементами. Поэтому далее полагаем $a_{ij} \geq 0$, $\forall i, j$.

Также отметим, что достаточно доказать утверждение задачи на множестве матриц, удовлетворяющих условию

$$\frac{1}{n}\sum_{i,j} a_{ij} = 1 - \varepsilon. \quad (SE)$$

Действительно, если матрица A удовлетворяет условию (SE) и $\rho(A) < 1$, то kA удовлетворяет условию (S), при всяком $k \in (0,1)$, и $\rho(kA) < k < 1$. Если же матрица A удовлетворяет условию (S) и ее спектральный радиус не меньше единицы, то $\exists \alpha > 1$, так что αA удовлетворяет условию (SE), и для этой матрицы тем более спектральный радиус не меньше единицы.

Значит, если будет доказано, что относительная мера тех матриц (удовлетворяющих условию (SE)), для которых спектральный радиус не меньше единицы, стремится к нулю с ростом n, это тем более будет выполнено для матриц, удовлетворяющих условию (S), поскольку доля таких матриц на множестве S ниже, чем на множестве SE.

Далее положим $a_{ij} \in \mathrm{Exp}\big(n/(1-\varepsilon)\big)$ — независимые одинаково распределенные случайные величины. Поскольку $\mathbb{E}a_{ij} = (1-\varepsilon)/n$, $\operatorname{Var} a_{ij} = (1-\varepsilon)^2/n^2$, имеем

$$\mathbb{E}\frac{1}{n}\sum_{i,j} a_{ij} = 1 - \varepsilon, \quad \operatorname{Var}\frac{1}{n}\sum_{i,j} a_{ij} = (1-\varepsilon)^2/n^2 \xrightarrow{n\to\infty} 0.$$

Из неравенства Чебышева получим $\frac{1}{n}\sum\limits_{i,j} a_{ij} \xrightarrow{P} 1 - \varepsilon$. Кроме того, распределение любого набора $\{a_{ij},\, (i,j) \in K\}$, $K \subset \{1,\ldots,n\} \times \{1,\ldots,n\}$, определяется только мощностью, или мерой, множества K.

Значит, при $n \to \infty$ распределение элементов случайной матрицы $A = \|a_{ij}\|_{i,j=1}^n$ будет сходиться к равномерному распределению на SE.

Поэтому для так выбранных элементов a_{ij} можно рассмотреть $P_n = \mathbb{P}(\|A\| \geq 1) \geq \mathbb{P}(\rho(A) \geq 1)$ — асимптотически верхняя оценка для относительной меры тех матриц из множества S, для которых спектральный радиус не меньше единицы. Тогда

$$P_n = \mathbb{P}\Big(\max_i \sum_j a_{ij} \geq 1\Big) = \mathbb{P}\Big(\big\{\sum_j a_{1j} \geq 1\big\} \cup \big\{\sum_j a_{2j} \geq 1\big\} \cup \ldots \cup \big\{\sum_j a_{nj} \geq 1\big\}\Big) \leq$$

$$\leq \sum_i \mathbb{P}\Big(\sum_j a_{ij} \geq 1\Big) = n\mathbb{P}\Big(\sum_j a_{1j} \geq 1\Big).$$

Случайная величина $X = \sum\limits_j a_{1j} \in \Gamma\big(n, \frac{1-\varepsilon}{n}\big)$, ее центральный момент четвертого порядка равен

$$\mathbb{E}(X - \mathbb{E}X)^4 = \mathbb{E}(X - (1-\varepsilon))^4 = \frac{(1-\varepsilon)^4}{n^4}(3n^2 + 6n) = O\Big(\frac{1}{n^2}\Big).$$

Используя неравенство Чебышева, получим

$$P_n \le n\mathbb{P}\Big(\sum_j a_{1j} \ge 1\Big) = n\mathbb{P}\Big(X \ge 1\Big) \le n\mathbb{P}\Big(|X - (1-\varepsilon)| \ge \varepsilon\Big) =$$

$$= n\mathbb{P}\Big(|X - \mathbb{E}X| \ge \varepsilon\Big) \le \frac{n}{\varepsilon^4}\mathbb{E}(X - \mathbb{E}X)^4 = O\Big(\frac{1}{n}\Big) \xrightarrow{n\to\infty} 0.$$

Отсюда следует требуемое утверждение $P_n \xrightarrow{n\to\infty} 0$. □

Задача 115 (модель случайного графа Эрдёша – Реньи, 1959-1961)**.** Задан случайный граф $G(n,p)$ (на n вершинах, любые две из которых соединяются ребром с вероятностью p, $p \in [0,1]$). Случайная величина $T_n(G)$ - равна числу треугольников, образованных ребрами, в случайном графе $G(n,p)$. Воспользовавшись неравенством Маркова, доказать, что если $p = o\,(1/n)$, то почти наверное треугольников в случайном графе нет, т.е.

$$P\left\{T_n(G) = 0\right\} \underset{n\to\infty}{\longrightarrow} 1.$$

Решение. To be done □

Задача 116. $V = \{1, ..., m\}$, $\mathrm{M} = \{M_1, ..., M_n\}$, $M_k \subseteq V$.

$\chi: \quad V \to \{-1, 1\}$ (можно интерпретировать, как раскраску множества V в два цвета).

$\chi(M_i) = \sum\limits_{a\in M_i} \chi(a)$ ($|\chi(M_i)|$ отвечает за “равномерность” покраски множества M_i в два цвета).

$disc(\mathrm{M}, \chi) = \max\limits_{i=1..n} |\chi(M_i)|$ (от слова discrepancy - уклонение) - мера того, что хотя бы один объект в M раскрашен “неравномерно”.

$disc(\mathrm{M}) = \min\limits_{\chi} disc(\mathrm{M}, \chi)$(“поиск” наилучшей раскраски).

Показать, что для $\forall n\ \forall m\ \forall \mathrm{M} \quad disc(\mathrm{M}) \le \sqrt{2m\ln(2n)}$. Т. е. $\exists\chi:\ disc(\mathrm{M},\chi) \le \sqrt{2m\ln(2n)}$.

Решение. To be done □

Задача 117. Оцените, сколько можно найти подмножеств множества $\{1, 2, ..., n\}$ таких, что симметрическая разность любых двух из них имеет мощность не менее $n/3$?

Решение. To be done □

Задача 118 (вероятностный метод в теории чисел; Харди – Рамануджан – Туран – Эрдёш - Кац, 1920, 1934, 1940)**.** Пусть $\nu(n)$ обозначает количество простых чисел p, делящих n. Тогда для любого λ

$$\lim_{n\to\infty}\frac{1}{n}\left|\left\{k:\ 1\le k\le n,\ \nu(k)\ge \ln\ln n+\lambda\sqrt{\ln\ln n}\right\}\right| = \frac{1}{\sqrt{2\pi}}\int\limits_{\lambda}^{\infty} e^{-t^2/2}dt.$$

Решение. To be done $\square$

3. Раздел 2. Производящие (характеристические) функции

Производящие (характеристические) функции в теории вероятностей и (асимптотической) комбинаторике (пропаганда идеи: часто для нахождения чисел, имеющих комбинаторную или вероятностную природу, выгодно изучить некоторое функциональное уравнение относительно неизвестной функции, которая в себе всю информацию об этих числах): формальные грамматики (теорема Лагранжа), перечислительная комбинаторика (теория Д.Пойа), метод включения и исключения и его обобщения (подход Дж.-К. Рота), метод отыскания значений (и их асимптотик) различных комбинаторных сумм (подход Г.П.Егорычева), сведение вычисления (асимптотики) чисел, имеющих вероятностную природу, к вычислению вычетов (к исследованию асимптотического поведения интегралов в комплексной плоскости, зависящих от параметра, с помощью метода перевала или стационарной фазы), предельные теоремы и законы больших чисел

Задача 119. Пусть ξ — случайная величина, равномерно распределенная на множестве всех пар векторов $(x, y) \in \{0,1\}^n \otimes \{0,1\}^n$, равная $\xi(x,y) = (x,y) = \sum\limits_{k=1}^{n} x_k y_k$. Найдите:

$$P_k = \mathbb{P}(\xi = k),\ \mathbb{E}\,\xi,\ \operatorname{Var}\xi.$$

Решение. Найдем производящую функцию для распределения $P_k = \mathbb{P}(\xi = k)$. Имеем

$$F(z) = \frac{1}{2^n \cdot 2^n} \sum_{x,y} z^{(x,y)} = \frac{1}{4^n} \sum_{x,y} z^{\sum\limits_{i=1}^{n} x_i y_i} = \frac{1}{4^n} \sum_{x_1,y_1} z^{x_1 y_1} \cdot \ldots \cdot \sum_{x_n,y_n} z^{x_n y_n} = \frac{(z+3)^n}{4^n}.$$

С другой стороны, $F(z) = \sum\limits_{k=0}^{n} P_k z^k$. Отсюда

$$P_k = \frac{C_n^k}{4^n} \cdot 3^{n-k}.$$

Далее

$$\mathbb{E}\,\xi = F'(1) = \frac{n}{4},$$

$$\operatorname{Var}\xi = F''(1) + F'(1) - [F'(1)]^2 = \frac{n(n-1)}{16} + \frac{n}{4} - \frac{n^2}{16} = \frac{3n}{16}.$$

□

Задача 120. Пусть ξ — с.в., равномерно распределенная на множестве бинарных матриц (т.е. матриц с элементами типа 0 и 1) порядка $m \times n$ и равная числу нулевых столбцов матрицы. Доказать, что

a) $P_k(m,n) = \mathbb{P}(\xi = k) = C_n^k \cdot \left(2^m - 1\right)^{n-k} \Big/ 2^{m \cdot n}$;

б) $\mathbb{E}\xi = n/\,2^m$;

в) если $2^m - 1 = \alpha \cdot n$, где α не зависит от n, то

$$\lim_{n\to\infty} P_k(m,n) = e^{-\lambda}\,\frac{\lambda^k}{k!},\quad \text{где } \lambda = \alpha^{-1}.$$

Решение. _Первый способ._

а) Число всевозможных бинарных матриц равно $2^{m\cdot n}$, число всевозможных ненулевых столбцов равно $2^m - 1$, поэтому

$$P_k(m,n) = \mathbb{P}(\xi = k) = C_n^k\,\frac{(2^m-1)^{n-k}}{2^{m\cdot n}}.$$

б) Имеем

$$\mathbb{E}\xi = \sum_{k=1}^{n} k\,C_n^k \frac{(2^m-1)^{n-k}}{2^{m\cdot n}} = n\sum_{k=1}^{n} C_{n-1}^{k-1}\frac{(2^m-1)^{n-k}}{2^{m\cdot n}} =$$

$$= \frac{n}{2^m}\sum_{k=1}^{n} C_{n-1}^{k-1}\frac{(2^m-1)^{n-k}}{2^{m(n-k)}\cdot 2^{m(k-1)}} = \frac{n}{2^m}\sum_{k=0}^{n-1} C_{n-1}^{k}\frac{(2^m-1)^{n-k}}{2^{m(n-k)}\cdot 2^{m\cdot k}} = \frac{n}{2^m}\Big(\frac{2^m-1}{2^m}+\frac{1}{2^m}\Big)^{n-1} = \frac{n}{2^m}.$$

в) Подставим в выражение для $P_k(m,n)$ условие $2^m - 1 = \alpha\cdot n$, учитывая $\lambda = \alpha^{-1}$:

$$P_k(m,n) = C_n^k\cdot\frac{(\alpha n)^{n-k}}{(\alpha n+1)^n} = \frac{n!}{k!(n-k)!}\cdot\frac{1}{\alpha^k n^k\cdot(1+\frac{1}{\alpha n})^n} =$$

$$= \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{n^k\cdot k!}\cdot\frac{\lambda^k}{(1+\frac{\lambda}{n})^n} =$$

$$= \Big(1-\frac{1}{n}\Big)\cdot\Big(1-\frac{2}{n}\Big)\cdot\ldots\Big(1-\frac{k-1}{n}\Big)\cdot\frac{\lambda^k}{k!}\cdot\frac{1}{(1+\frac{\lambda}{n})^n} \xrightarrow{n\to\infty} \frac{\lambda^k}{k!}\,e^{-\lambda}.$$

Второй способ.

Рассмотрим следующие случайные величины:

$$\xi_i = \begin{cases} 1, & i\text{-й столбец нулевой},\\ 0, & \text{иначе}. \end{cases}$$

Тогда $\xi_i \in \mathrm{Be}\big(\frac{1}{2^m}\big)$, $\xi = \xi_1 + \ldots + \xi_n$. Производящая функция с.в. ξ_i имеет вид

$$\psi_{\xi_i}(z) = \mathbb{P}(\xi_i = 0) + z\cdot\mathbb{P}(\xi_i = 1) = \frac{1}{2^m}\big(z + (2^m-1)\big)$$

$$\Rightarrow\ \psi_\xi(z) = \big[\psi_{\xi_i}(z)\big]^n = \frac{1}{2^{mn}}\big(z+(2^m-1)\big)^n.$$

а) Величина $\mathbb{P}(\xi = k)$ равна коэффициенту при z^k в многочлене $\psi_\xi(z)$, который по формуле бинома Ньютона равен

$$P_k(m,n) = \mathbb{P}(\xi = k) = \frac{1}{2^{mn}} C_n^k (2^m - 1)^{n-k}.$$

б) Имеем

$$\mathbb{E}\xi = \frac{d}{dz}\psi_\xi(z)\Big|_{z=1} = \frac{n}{2^{mn}}(1 + 2^m - 1)^{n-1} = \frac{n}{2^m}.$$

в) Подставим в характеристическую функцию условие $2^m - 1 = \alpha \cdot n$, учитывая $\lambda = \alpha^{-1}$:

$$\psi_\xi(z) = \frac{1}{(\alpha n + 1)^n}(z + \alpha n)^n = \left(\frac{z + \alpha n}{\alpha n + 1}\right)^n =$$

$$= \left(1 + \frac{z-1}{\alpha n + 1}\right)^n \xrightarrow{n\to\infty} f(z) = \exp\bigl\{\lambda(z-1)\bigr\}.$$

Функция $f(z)$ является характеристической функцией случайной величины $\eta \in \mathrm{Po}(\lambda)$. Поэтому в силу теоремы о предельном переходе для характеристических функций пуассоновское распределение является предельным распределением:

$$\lim_{n\to\infty} P_k(m,n) = \mathbb{P}(\eta = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

□

Задача 121. Сколько раз нужно подбросить монету, чтобы решка выпала два раза подряд?

Решение. Вероятностное пространство состоит из всех последовательностей букв Р и О, оканчивающихся на РР, но не содержащих двух Р подряд ранее. Любой элемент тогда имеет вид:

$$(O + PO) * PP \text{ — регулярное выражение .}$$

Некоммутативная производящая функция имеет вид:

$$S = PP + OPP + POPP + OOPP + OPOPP + POOPP + OOOPP + \ldots = (1 - O - PO)^{-1}PP$$

Заменяем $P \to pz$; $O \to qz$, получим производящую функцию

$$G(z) = \frac{p^2 z^2}{(1 - qz - pqz^2)}$$

Для случайной величины числа бросаний до появления двух последовательных решек математическое ожидание равно $G'(1) = 6$ (монета симметричная $p = q = 1/2$). □

Задача 122 (игра У. Пенни, 1969)**.** Алиса и Билл играют в игру: они бросают монету до тех пор, пока не встретится PPO или POO. Если первой появится последовательность PPO, выигрывает Алиса, если POO – Билл. Будет ли игра честной?

Решение. Производящая функция времени ожидания последовательности PPO:

$$G(z) = \frac{z^3}{z^3 - 8(z-1)}.$$

Такая же производящая функция и для времен ожидания POO. Таки образом, если каждый играет сам по себе (т.е. каждый подкидывает монетку), то ни у кого нет преимущества. Но если рассматривать одну игровую последовательность (кидается одна монетка), то это не верно.

Рассмотрим конфигурации, выигрышные для Алисы (A) и для Билла (B):

$$\text{A=PPO+PPPO+OPPO+PPPPO+OPPPO+POPPO+OOPPO+}\dots$$
$$\text{B=POO+OPPO+PPOO+PPPOO+OPPOO+POPPO+OOPPO+}\dots$$

Рассмотрим конфигурации всех последовательностей, для которых пока ни один из игроков не выиграл:

$$\text{N=1+P+O+PP+PO+OP+OO+PPP+POP+OPP+}\dots$$

Справедливы следующие равенства:

$$\text{1+N(P+O)=N+A+B,}$$

$$\text{NPPO=A,}$$

$$\text{NPOO=B+AO.}$$

Подставляя вместо O и P 1/2, получим систему уравнений, решением которой является A = 2/3, B = 1/3. Алиса будет выигрывать примерно вдвое чаще Билла! □

Задача 123. Теперь трое игроков: Алиса, Билл и Компьютер. Играют пока не выпадет одна из следующих последовательностей: A=PPOP, B=POPP, C=OPPP. Каковы шансы каждого выиграть?

Решение. To be done □

Задача 124. Рассматривается игра У. Пенни из примера 2. Показать, что последовательность $a_1a_2\dots a_l$ всегда уступает последовательности $\bar{a}_2a_1a_2\dots a_{l-1}$, $l \geq 3$.

Решение. To be done □

Задача 125 (об оценке хвостов)**.** Пусть $\Psi(z) = Mz^X$ — производящая функция случайной величины (ПФСВ) X. Докажите, что

$$\mathrm{P}(X \le r) \le x^{-r}\Psi(x), \text{ для } 0 < x \le 1;$$

$$\mathrm{P}(X \ge r) \le x^{-r}\Psi(x), \text{ для } x \ge 1.$$

Решение. Если $0 < x \le 1$, получим

$$\Psi(x) = \mathbb{E}x^X \ge \mathbb{E}\big(x^X \cdot \mathcal{I}_{\{X \le r\}}\big) \ge \mathbb{E}\big(x^r \cdot \mathcal{I}_{\{X \le r\}}\big) = x^r\, \mathbb{E}\mathcal{I}_{\{X \le r\}} = x^r\, \mathbb{P}(X \le r),$$

откуда получим первое неравенство $\mathbb{P}(X \le r) \le x^{-r}\Psi(x)$.

Если $x \ge 1$, тогда

$$\Psi(x) = \mathbb{E}x^X \ge \mathbb{E}\big(x^X \cdot \mathcal{I}_{\{X \ge r\}}\big) \ge \mathbb{E}\big(x^r \cdot \mathcal{I}_{\{X \ge r\}}\big) = x^r\, \mathbb{E}\mathcal{I}_{\{X \ge r\}} = x^r\, \mathbb{P}(X \ge r),$$

откуда получим второе неравенство $\mathbb{P}(X \ge r) \le x^{-r}\Psi(x)$. □

Задача 126 (о червяке Шредингера)**.** В вершине пятиугольника $ABCDE$ находится яблоко, а на расстоянии двух ребер, в вершине C, находится червяк. Каждый день червяк переползает в одну из двух соседних вершин с равной вероятностью. Так, через один день червяк окажется в вершине B или D с вероятностью $1/2$. По прошествии двух дней червяк может снова оказаться в C, поскольку он не запоминает своих предыдущих положений. Достигнув вершины A, червячок останавливается пообедать.

1. Чему равны математическое ожидание и дисперсия числа дней прошедших до обеда?
2. Какую оценку дает *неравенство Чебышёва* ($\mathrm{P}\left(|X - MX| \ge a\right) \le DX/a^2$) для вероятности p того, что это число дней будет 100 или больше?
3. Что позволяют сказать о величине p оценки из задачи “об оценки хвостов”.

Решение. To be done □

Задача 127. Пять человек стоят в вершинах пятиугольника $ABCDE$ и бросают друг другу диски Фрисби. У них имеется два диска, которые в начальный момент находятся в соседних вершинах. В очередной момент времени диски бросают либо налево, либо направо с одинаковой вероятностью. Процесс продолжается до тех пор, пока обе тарелки не окажутся в одной вершине.

1. Найдите математическое ожидание и дисперсию числа пар бросков.
2. Найдите “замкнутое” выражение через числа Фибоначчи для вероятности того, что игра продлиться более 100 шагов.

Решение. To be done □

Задача 128. Обобщите задачу 4 на случай m-угольника и найдите математическое ожидание и дисперсию числа пар бросков до столкновения дисков. Докажите, что если m нечетно, то ПФСВ для числа бросаний представимо в следующем виде:[1]

$$G_m(z) = \prod_{k=1}^{(m-1)/2} \frac{p_k z}{1 - q_k z},$$

где

$$p_k = \sin^2 \frac{(2k-1)\pi}{2m}, \quad q_k = \cos^2 \frac{(2k-1)\pi}{2m}.$$

Решение. To be done □

Задача 129 (загадочный случайный суп)**.** Студент, решивший отобедать в столовой, может обнаружить в своей тарелке с супом случайное число N инородных частиц Λ со средним μ и конечной дисперсией. С вероятностью p выбранная частица является мухой, иначе это таракан; типы разных частиц независимы. Пусть F – количество мух и S – количество тараканов.

А) Покажите, что производящая функция случайной величины F равна

$$\psi_F(s) = \psi_N(ps + 1 - p).$$

Б) Предположим, что случайная величина N имеет пуассоновское (poisson) распределение с параметром μ (записывают $\mathrm{N} \in \mathrm{Po}\,(\mu)$). Покажите, что F имеет пуассоновское распределение с параметром $p\mu$, а случайные величины F и S независимы. Покажите, что

$$\psi_N(s) = \psi_N^2\Big(\frac{1}{2}(1+s)\Big).$$

Убедитесь, что случайная величина N имеет пуассоновское распределение.

Решение. **А)** С.в. F можно представить в виде $F = \xi_1 + \ldots + \xi_N$, где $\xi_i \in \mathrm{Be}(p)$. Тогда

$$\psi_F(s) = \mathbb{E}s^F = \mathbb{E}s^{\xi_1+\ldots+\xi_N} = \mathbb{E}\big(\mathbb{E}s^{\xi_i}\big)^N = \psi_N\big(\mathbb{E}s^{\xi_i}\big).$$

Поскольку $\mathbb{E}s^{\xi_i} = p\,s^1 + (1-p)s^0 = ps + 1 - p$, получим искомую формулу для $\psi_F(s)$.

[1] Воспользуйтесь подстановкой $z = 1/\cos^2\theta$.

Б) С.в. $N \in \mathrm{Po}(\mu)$ имеет производящую функцию $\psi_N(s) = \exp\big\{\mu(s-1)\big\}$. Отсюда

$$\psi_N^2\Big(\frac{1}{2}(1+s)\Big) = \Big(\exp\Big\{\mu\Big(\frac{1}{2}(1+s)-1\Big)\Big\}\Big)^2 = \exp\big\{\mu\big((1+s)-2\big)\big\} = \exp\big\{\mu(s-1)\big\} = \psi_N(s).$$

Далее,

$$\psi_F(s) = \psi_N(ps+1-p) = \exp\big\{\mu(ps+1-p-1)\big\} = \exp\big\{p\mu(s-1)\big\},$$

откуда $F \in \mathrm{Po}(p\mu)$.

Поскольку аналогично для с.в. S можно показать, что $S \in \mathrm{Po}\big((1-p)\mu\big)$, из вида характеристической функции для распределений $\mathrm{Po}(\mu)$, $\mathrm{Po}(p\mu)$, $\mathrm{Po}\big((1-p)\mu\big)$ получим

$$\psi_N(s) = \psi_F(s)\cdot\psi_S(s), \quad \forall s > 0.$$

Т.к. $N = F + S$, отсюда следует независимость случайных величин F и S.

$\square$

Задача 130 (производящие функции, вычеты)**.** Обозначим через E^n – множество бинарных последовательностей длины n, или множество вершин единичного n-мерного куба, а через E^n_k – k-ый слой куба E^n, то есть подмножество точек E^n, имеющих ровно k единичных координат. Пусть $X = (\vec{x},\vec{y})$ - случайная величина, где $\vec{x} \in E^n_p$, $\vec{y} \in E^n_q$ - независимые и равномерно распределенные на E^n_p и E^n_q соответственно векторы. Обозначим через $a_{p,q}(k) = P\left\{X = k\right\}$. Доказать следующие утверждения:

1) $\sum\limits_{k=0}^{n} a_{p,q}(k)z^k = \frac{1}{2\pi i}C_n^p \oint\limits_{|u|=\rho} \frac{(1+zu)^p(1+u)^{n-p}}{u^{q+1}}du.$

2) $a_{p,q}(k) = \frac{C_p^k C_{n-p}^{q-k}}{C_n^q}.$

3) $EX = \frac{pq}{n}.$

4) $DX = \frac{pq}{n(n-1)}\left(n + \frac{pq}{n} - (p+q)\right).$

Решение. **2)** Рассматриваем $k \le \min(p,q)$. Число всевозможных векторов $\vec{x} \in E^n_p$ равно C_n^p, число векторов $\vec{y} \in E^n_q$ равно C_n^q. Если считать позиции единиц в векторе $\vec{x}$ зафиксированными, то из них можно выбрать C_p^k способами те позиции, в которых находятся единицы в векторе $\vec{y}$. Тогда в векторе $\vec{y}$ на остальных $n-p$ позициях находится $q-k$ единиц, что дает C_{n-p}^{q-k} вариантов. В итоге получим

$$a_{p,q}(k) = \frac{C_n^p\, C_p^k\, C_{n-p}^{q-k}}{C_n^p\, C_n^q} = \frac{C_p^k C_{n-p}^{q-k}}{C_n^q}.$$

3) Запишем $X = (\vec{x},\vec{y}) = x_1y_1 + \ldots + x_ny_n$. Отсюда

$$\mathbb{E}X = n\,\mathbb{E}x_1y_1 = n\,\mathbb{E}x_1\mathbb{E}y_1 = n\,\frac{p}{n}\frac{q}{n} = \frac{pq}{n},$$

т.к. $\mathbb{E}x_1 = \mathbb{P}(x_1 = 1) = \frac{C_{n-1}^{p-1}}{C_n^p} = \frac{p}{n}$. Аналогично $\mathbb{E}y_1 = \frac{q}{n}$.

4) Рассмотрим

$$\mathbb{E}X^2 = \mathbb{E}(x_1y_1+\ldots+x_ny_n)^2 = n\,\mathbb{E}x_1y_1+n(n-1)\mathbb{E}x_1y_1x_2y_2 = \frac{pq}{n}+n(n-1)\mathbb{E}x_1x_2\,\mathbb{E}y_1y_2.$$

Вычислим $\mathbb{E}x_1x_2 = \frac{C_{n-2}^{p-2}}{C_n^p} = \frac{p(p-1)}{n(n-1)}$. Аналогично $\mathbb{E}y_1y_2 = \frac{q(q-1)}{n(n-1)}$. Отсюда

$$\mathbb{E}X^2 = \frac{pq}{n} + \frac{p(p-1)q(q-1)}{n(n-1)}$$

$$\Rightarrow\ \operatorname{Var} X = \mathbb{E}X^2-(\mathbb{E}X)^2 = \frac{pq}{n}+\frac{p(p-1)q(q-1)}{n(n-1)}-\frac{p^2q^2}{n^2} = \frac{pq}{n(n-1)}\left(n+\frac{pq}{n}-(p+q)\right).$$

□

4. Раздел 4. Метод Монте-Карло

Задача 131. На плоскости дано ограниченное измеримое по Лебегу множество S. Требуется найти площадь (меру Лебега) этого множества с заданной точностью ε.

Поскольку по условию множество ограничено, то вокруг него можно описать квадрат со стороной a. Выберем декартову систему координат в одной из вершин квадрата с осями, параллельными сторонам квадрата. Рассмотрим n независимых с.в. $\{X_k\}_{k=1}^n$, имеющих одинаковое равномерное распределение в этом квадрате, т.е. $X_k \in R([0,a]^2)$. Введем с.в.

$$Y_k = I(X_k \in S) = \begin{cases} 1, & X_k \in S \\ 0, & X_k \notin S \end{cases}.$$

Тогда $\{Y_k\}_{k=1}^n$ — независимые одинаково распределенные с.в.. Ясно, что $Y_k \in \mathrm{Be}(p(S))$. Следовательно, по у.з.б.ч.

$$\frac{Y_1 + \ldots + Y_n}{n} \xrightarrow{\text{п.н.}} \mathbb{E}(Y_1) = p(S) = \frac{\mu(S)}{a^2} \quad \text{при } n \to \infty.$$

Оцените сверху следующую вероятность

$$\mathbb{P}\Big(\Big|\frac{Y_1 + \ldots + Y_n}{n} - \frac{\mu(S)}{a^2}\Big| > \delta\Big).$$

Решение. Обозначим $S_n = Y_1 + \ldots + Y_n$. Имеем

$$\mathbb{P}\Big(\Big|\frac{Y_1 + \ldots + Y_n}{n} - \frac{\mu(S)}{a^2}\Big| > \delta\Big) = \mathbb{P}\Big(\Big|\frac{S_n - \mathbb{E}S_n}{n}\Big| > \delta\Big) =$$

$$= \mathbb{P}\Big(\Big|\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}}\Big| > \delta\sqrt{\tfrac{n}{p(S)(1-p(S))}}\Big).$$

Обозначим $\delta\sqrt{\tfrac{n}{p(S)(1-p(S))}} = \delta_1$. Согласно ц.п.т. при больших n приближенно выполнено $\dfrac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} \in N(0,1)$. Поэтому

$$\mathbb{P}\Big(\Big|\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}}\Big| > \delta_1\Big) \approx 2\big[1 - \Phi\big(\delta_1\big)\big].$$

Наибольшее возможное значение величины $1 - \Phi\Big(\delta\sqrt{\tfrac{n}{p(S)(1-p(S))}}\Big)$ достигается при наибольшем возможном $p(S) \cdot (1 - p(S)) \quad \Rightarrow$ при $p(S) = 0.5$. Отсюда

$$2\big[1 - \Phi\big(\delta_1\big)\big] \leqslant 2\big[1 - \Phi\big(2\delta\sqrt{n}\big)\big].$$

Оценим ошибку данного приближения. Имеем

$$\mathbb{P}\Big(\Big|\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}}\Big| > \delta_1\Big) = 1 - \Big[\mathbb{P}\Big(\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} < \delta_1\Big) - \mathbb{P}\Big(\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} < -\delta_1\Big)\Big].$$

Согласно неравенству Берри-Эссена для всякого значения x выполнено

$$\Big|\mathbb{P}\Big(\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} < x\Big) - \Phi(x)\Big| \leqslant \frac{C_0\mu^3}{\sigma^3\sqrt{n}}.$$

Здесь $C_0 < 0.7056$, $\sigma^2 = \operatorname{Var} X_k = p(1-p) = 0.25$,

$$\mu^3 = \mathbb{E}|X_k - p|^3 = p(1-p)^3 + (1-p)p^3 = 0.125.$$

Отсюда

$$\Big|\mathbb{P}\Big(\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} < x\Big) - \Phi(x)\Big| \leqslant \frac{C_0}{\sqrt{n}}$$

$$\Rightarrow \mathbb{P}\Big(\Big|\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}}\Big| > \delta_1\Big) \leqslant 1 - \big[\Phi(\delta_1) - \Phi(-\delta_1)\big] + \frac{2C_0}{\sqrt{n}} =$$

$$= 2\big[1 - \Phi(\delta_1)\big] + \frac{2C_0}{\sqrt{n}} \leqslant 2\big[1 - \Phi(2\delta\sqrt{n})\big] + \frac{2C_0}{\sqrt{n}}.$$

Таким образом, исходная вероятность оценивается как

$$\mathbb{P}\Big(\Big|\frac{S_n - \mathbb{E}S_n}{n}\Big| > \delta\Big) \leqslant 2\big[1 - \Phi(2\delta\sqrt{n})\big] + \frac{2C_0}{\sqrt{n}}.$$

□

Задача 132. На плоскости проведены параллельные прямые на единичном расстоянии друг от друга, и на плоскость наугад бросается иголка длиной $L < 1$. Угол между прямыми и иголкой и расстояние от середины иглы до ближайшей прямой являются независимыми с.в., равномерно распределенными на соответственно $(0, 2\pi)$ и $(-1/2, 1/2)$. С помощью серии таких опытов вычислить число π с заданной точностью $\delta = 1\%$ и с вероятностью ошибки не больше $\varepsilon = 5\%$.

Решение. 1. Угол между прямыми и иголкой и расстояние от середины иглы до ближайшей прямой обозначим через φ и r. Условие попадания иголки на прямую можно записать в виде

$$\frac{L}{2}|\cos\varphi| > |r|.$$

Из геометрического смысла достаточно рассматривать угол между прямыми и иголкой в промежутке $(0, \pi/2)$. Поэтому введем с.в. ψ, равномерно распределенную на $(0, \pi/2)$, удовлетворяющую условию $\cos\psi = |\cos\varphi|$. Также для удобства введем с.в. $R = 2|r|$, равномерно распределенную на $(0, 1)$.

Тогда условие попадания иголки на прямую примет вид $\cos\psi > \frac{R}{L}$. Вероятность данного события равна

$$\mathbb{P}\big(\cos\psi > \tfrac{R}{L}\big) = \mathbb{P}\big(\psi \in [0, \arccos\tfrac{R}{L}]\big) =$$

$$= \int\limits_0^L dr \int\limits_0^{\arccos\frac{r}{L}} \frac{d\psi}{\pi/2} = \frac{2}{\pi}\int_0^L \arccos\frac{r}{L}\,dr = \frac{2L}{\pi}\int\limits_0^1 \arccos x\,dx = \frac{2L}{\pi}.$$

2. Серии рассматриваемых в задаче опытов описываются с помощью с.в. $X_k \in \mathrm{Be}(p)$, $p = \frac{2L}{\pi}$. С помощью ц.п.т. можно оценивать величину $\frac{2L}{\pi}$. Чтобы оценить число π с относительной точностью δ, достаточно оценить $\frac{2L}{\pi}$ с абсолютной точностью $\frac{2L\delta}{\pi^2}$. Действительно, если $\tilde{\pi} \in [\pi(1-\delta), \pi(1+\delta)]$, то

$$\Big|2L(\frac{1}{\tilde{\pi}} - \frac{1}{\pi})\Big| \leqslant \frac{2L\delta}{\pi(1-\delta)} \approx \frac{2L\delta}{\pi}.$$

Аналогично предыдущей задаче получим

$$\mathbb{P}\Big(\Big|\frac{X_1 + \ldots + X_n}{n} - \frac{2L}{\pi}\Big| > \frac{2L\delta}{\pi}\Big) \leqslant 2\Big[1 - \Phi\Big(2\sqrt{n}\frac{2L\delta}{\pi}\Big)\Big] = 2\Big[1 - \Phi\Big(\frac{4L\delta\sqrt{n}}{\pi}\Big)\Big].$$

Чтобы вероятность ошибки была не больше $\varepsilon = 5\%$, достаточно условия

$$2\Big[1 - \Phi\Big(\frac{4L\delta\sqrt{n}}{\pi}\Big)\Big] = \varepsilon$$

$$\Rightarrow \quad n = \Big(\frac{\pi}{4\delta \cdot L} \cdot \gamma_{1-\varepsilon/2}\Big)^2,$$

где γ_α — квантиль нормального распределения: $\Phi(\gamma_\alpha) = \alpha$.

Таким образом, с помощью серии из n опытов находим относительную долю успехов — когда игла пересекает одну из прямых. Полученная величина является оценкой для $\frac{2L}{\pi}$, из которой получается искомая оценка для числа π.

□

Задача 133. Требуется вычислить с заданной точностью ε и с заданной доверительной вероятностью γ абсолютно сходящийся интеграл

$$J = \int\limits_{[0,1]^m} f(x)\,dx.$$

Считается, что $\forall x \in [0,1]^m : \ |f(x)| \leqslant 1$.

Решение. Введем случайный m-вектор $X \in R([0,1]^m)$ и с.в. $\xi = f(X)$. Тогда

$$\mathbb{E}\xi = \int\limits_{[0,1]^m} f(x)\,dx = J.$$

Поэтому получаем оценку интеграла $\overline{J}_n = \frac{1}{n}\sum\limits_{k=1}^{n} f(x^k)$, где $x^k, \ \ k = 1, \ldots, n$ — повторная выборка значений случайного вектора X (т.е. все $x^k, \ k = 1, \ldots, n$ — независимы и одинаково распределены: так же как и вектор X). В задаче требуется оценить сверху число n $(n > m)$, начиная с которого $\mathbb{P}(|J - \overline{J}_n| \leqslant \varepsilon) \geqslant \gamma$.

Обозначим $S_n = \sum\limits_{k=1}^{n} f(x^k) = n\overline{J}_n,\ \ \sigma^2 = \operatorname{Var}\xi \leqslant \mathbb{E}\xi^2 \leqslant 1$. Имеем

$$\mathbb{P}(|J - \overline{J}_n| \leqslant \varepsilon) = \mathbb{P}\Big(\Big|\frac{S_n - nJ}{\sigma\sqrt{n}}\Big| \leqslant \frac{\sqrt{n}}{\sigma}\varepsilon\Big).$$

Согласно ц.п.т. можно считать при достаточно больших n, что с.в. $\frac{S_n - nJ}{\sigma\sqrt{n}}$ имеет распределение $N(0,1)$, поэтому

$$\mathbb{P}\Big(\Big|\frac{S_n - nJ}{\sigma\sqrt{n}}\Big| \leqslant \frac{\sqrt{n}}{\sigma}\varepsilon\Big) \approx 2\Phi\big(\tfrac{\sqrt{n}}{\sigma}\varepsilon\big) - 1.$$

Поскольку $\sigma \leqslant 1$, чтобы выполнялось $\mathbb{P}(|J - \overline{J}_n| \leqslant \varepsilon) \geqslant \gamma$, достаточно условия

$$2\Phi\big(\sqrt{n}\varepsilon\big) - 1 \geqslant \gamma$$

$$\Rightarrow\ n \geqslant \Big[\frac{1}{\varepsilon}\Phi^{-1}\big(\tfrac{\gamma+1}{2}\big)\Big]^2.$$

$\square$

5. Раздел 5. Вероятностный анализ алгоритмов

Вероятностный анализ алгоритмов (сложность в среднем, сложность для почти всех входов), вероятностные алгоритмы и их анализ (проверка тождеств с помощью метода Монте-Карло, вероятностное округление), дерандомизация, вероятностные вычисления (вероятностная машина Тьюринга, полиномиальный вероятностный алгоритм проверки простоты числа)

Задача 134. Рассмотрим «NP-полную» задачу

$$\sum_{j=1}^{n} x_j \to \max$$

$$x_j \in \{0,1\}, \quad j = 1, \ldots, n;$$
$$\sum_{j=1}^{n} a_{ij} x_j \leqslant 1, \quad i = 1, \ldots, m \tag{5.1}$$
$$(a_{ij} \in \{0,1\}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n).$$

Булев вектор x длины n будем называть допустимым, если он удовлетворяет системе (5.1). Обозначим через $T(j)$ множество всех допустимых булевых векторов для системы (5.1) с $n-j$ нулевыми последними компонентами и через e_j — вектор длины n с единичной j–ой компонентой и с остальными нулевыми компонентами. Рассмотрим алгоритм: 1) строим множество допустимых решений $T(j)$ на основе множества $T(j-1)$, пытаясь добавить вектор e_j ко всем булевым векторам $T(j-1)$; 2) среди $|T(n)|$ допустимых булевых векторов ищем «наилучший».

а) Покажите, что сложность описанного алгоритма составляет $O(|T(n)|mn)$. При каких $a_{ij} \in \{0,1\}$ алгоритм будет работать экспоненциально долго?

б) Оцените сложность в среднем (математическое ожидание времени работы алгоритма), т.е. $O(\mathbb{E}(|T(n)|)mn)$, если с.в. $\{a_{ij}\}_{i,j=1}^{n}$ — независимые и одинаково распределенные по закону Бернулли $\mathrm{Be}(p)$ $(mp^2 \geqslant \ln n)$.

Решение. а) Очевидно, $T(0) = \{(0, \ldots, 0) \in \{0,1\}^n\}$, $T(j) \supseteq T(j-1)$.

На каждом шаге $k = 1, \ldots, n$ уже известны значения выражений $\sum\limits_{j=1}^{k-1} a_{ij} x_j$, $i = 1, \ldots, m$ для каждого $x = (x_1, \ldots, x_{k-1}, 0, \ldots, 0) \in T(k-1)$. Чтобы построить множество $T(k)$, нужно добавить к каждому $x \in T(k-1)$ вектор e_k и проверить m условий вида (5.1). Это равносильно тому, чтобы к каждому уже вычисленному выражению $\sum\limits_{j=1}^{k-1} a_{ij} x_j$, прибавить a_{ik}, $i = 1, \ldots, m$, и проверить условие (5.1). Каждая из этих операций имеет некоторую фиксированную сложность C.

Таким образом, на k-ом шаге требуется выполнить m операций сложения сложности C для каждого вектора из множества $T(k-1)$, и общее число операций оценивается как

$$S = \sum_{k=1}^{n} Cm \cdot |T(k-1)| \leqslant Cmn \cdot |T(n)|,$$

т.е. $S = O(|T(n)|mn)$.

Алгоритм будет работать экспоненциально долго, если выполнена экспоненциальная оценка снизу для $|T(n)|$. Например, если число m ограничений (5.1) фиксировано, и все векторы $a_i = \{a_{i1}, \ldots, a_{in}\}$, $i = 1, \ldots, m$, имеют ограниченный вес: $|a_i| = \sum\limits_{j=1}^{n} a_{ij} \leqslant a$, $i = 1, \ldots, m$, то каждое условие (5.1), $i = 1, \ldots, m$, является ограничением на не более чем a компонент вектора x (остальные компоненты в условии (5.1) не участвуют), поэтому в совокупности условия (5.1) являются ограничениями на не более чем $a \cdot m$ компонент допустимого вектора x. Остальные же $(n - a \cdot m)$ компонент вектора в них не участвуют и могут принимать любое из значений 0 или 1. Поэтому мощность множества $T(n)$ можно оценить как

$$|T(n)| \geqslant 2^{n-a\cdot m} = O(2^n).$$

В таком случае $S = O(mn\,2^n)$.

б) Пусть $k > 0$. Положим: $x^k = x_{j_1,\ldots,j_k}$ — вектор с k единицами (на позициях $\{j_1, \ldots, j_k\}$) и $n-k$ нулями; p_{ki} — вероятность выполнения i-го неравенства системы (5.1) для $x_{j_1,\ldots,j_k}$; P_k — вероятность того, что x^k — допустимое решение. Поскольку все с.в. $\{a_{ij}\}_{i,j=1}^{n}$ независимы и одинаково распределены, вероятности p_{ki} и P_k не зависят от набора $\{j_1, \ldots, j_k\}$. Поэтому, например, при $\{j_1, \ldots, j_k\} = \{1, \ldots, k\}$ выполнено i-е неравенство системы (5.1), если $\sum\limits_{j=1}^{k} a_{ij} \leqslant 1$. Т.к. по определению с.в. $\sum\limits_{j=1}^{k} a_{ij}$ имеет биномиальное распределение с параметрами p и k, то

$$p_{ki} = \mathbb{P}\Big(\sum_{j=1}^{k} a_{ij} \leqslant 1\Big) = (1-p)^k + kp(1-p)^{k-1} = (1-p)^{k-1} \cdot (1 + (k-1)p) \leqslant$$

$$\leqslant (1-p)^{k-1} \cdot (1+p)^{k-1} = (1-p^2)^{k-1} \leqslant e^{-p^2(k-1)}.$$

Тогда $P_k = (p_{ki})^m \leqslant e^{-mp^2(k-1)}$. По определению

$$\mathbb{E}(|T(n)|) = \sum_{k=0}^{n} C_n^k P_k.$$

Для биномиальных коэффициентов при $k \geqslant 2$ выполнена оценка

$$C_n^k = \frac{n(n-1)\ldots(n-k+1)}{k!} < n^k.$$

Поэтому

$$\mathbb{E}(|T(n)|) = \sum_{k=0}^{n} C_n^k P_k < 1 + n + \sum_{k=2}^{n} n^k e^{-mp^2(k-1)} = 1 + n + n\sum_{k=2}^{n} e^{(k-1)(\ln n - mp^2)}.$$

Т.к. по условию $mp^2 \geqslant \ln n$, для $\mathbb{E}(|T(n)|)$ выполнена оценка сверху

$$\mathbb{E}(|T(n)|) < 1 + n + n^2$$

и сложность алгоритма в среднем составляет $O(\mathbb{E}(|T(n)|)mn) = O(n^3 m)$.

□

Задача 135 (быстрая сортировка)**.** Оценить сложность в среднем (временную и пространственную) алгоритма быстрой сортировки.

Решение. To be done □

Задача 136. Найти среднее число присваиваний m:= при выполнении алгоритма поиска наименьшего элемента массива:

```
m:=x[1];
for i=2 to n do
if x[i]<m then m:=x[i]
end;
```

Предполагается, что все возможные взаимные порядки элементов в исходном массиве длины n являются равновероятными; в качестве размера входа берется n.

Решение. To be done □

Задача 137 (сортировка вставками)**.** На k-ом цикле (k=1,..,n-1) первые k элементов уже отсортированы (по возрастанию). Требуется найти позицию k+1 элемента относительно предыдущих элементов. Для этого можно сравнивать x[k+1] элемент с отсортированными k элементами, пока не найдем элемент ≤x[k+1], тем самым найдется место для x[k+1] в отсортированном массиве длины k. Для того чтобы всегда нашелся такой элемент, можно дополнить исходный массив элементом x[0] =−∞. Число выполненных сравнений (на k –ом цикле) равно числу элементов исходного массива с индексами меньшими k+1, но большими x[k+1] плюс еще одно сравнение. В нижеследующих терминах общее число сравнений выполненных алгоритмом рано $I(n) + n - 1$. Найдите среднее время работы алгоритма (среднее число сравнений), а также дисперсию числа сравнений, пользуясь следующим указанием.

Указание.

Определение. Говорят, что элементы (σ_i, σ_j) образуют инверсию в перестановке $(\sigma_1, \sigma_2, \dots, \sigma_n)$, если $i < j$, но $\sigma_i > \sigma_j$.

Сопоставим перестановке $(\sigma_1, \sigma_2, \dots, \sigma_n)$ последовательность $(b_1, b_2, \dots, b_n)$, где b_k - число элементов перестановки, предшествующих k, но больших k. Заметим, что $0 \le b_k \le n-k$, $k = 1, ..., n$. Покажите, что это соответствие взаимно однозначно. Число инверсий в перестановке длины n равно $I(n) = \sum\limits_{k=1}^{n} b_k$. Последовательности $(b_1, b_2, \dots, b_n)$ сопоставим моном $x_{b_1} x_{b_2} \cdots x_{b_n}$. Тогда производящая функция

$$F(x_0, x_1, \dots, x_n) = \sum_{\sigma \in S_n} x_{b_1} x_{b_2} \cdots x_{b_n}$$

– есть сумма всех таких мономов, соответствующих каждой из n! возможных перестановок.

Покажите, что

$$F(x_0, x_1, \ldots, x_n) = x_0 (x_0 + x_1)(x_0 + x_1 + x_2) \cdots (x_0 + x_1 + \cdots + x_{n-1}).$$

Покажите, что обычная производящая функция числа инверсий в перестановке длины n есть

$$F(z) = z^0 \left(z^0 + z^1\right) \left(z^0 + z^1 + z^2\right) \cdots \left(z^0 + z^1 + \cdots + z^{n-1}\right).$$

Производящая функция случайной величины, равной числу инверсий в перестановке тогда имеет вид:

$$\Phi(z) = \sum_{k=0}^{n(n-1)/2} P\{I(n) = k\} z^k = \frac{1}{n!} F(z) = \prod_{k=1}^{n} b_k(z),$$

$$b_k(z) = \frac{z^0 + z^1 + \ldots + z^{n-k}}{n-k+1}.$$

Далее для решения задачи нужно воспользоваться связью между производящей функцией случайной величины и моментами этой случайной величины.

Решение. To be done □

Задача 138 (быстрая сортировка)**.** Покажите, что временная сложность рандомизированного алгоритма быстрой сортировки, где разбивающий элемент выбирается случайно, допускает оценку $O(n \log n)$.

Решение. To be done □

Задача 139. А) Пусть имеется генератор случайных чисел, в результате обращения к которому появляется 0 или 1 с одинаковой вероятностью равной 1/2 (аналог подбрасывания симметричной монеты). Пусть задано вещественное число $0 \leq p \leq 1$. С помощью имеющегося генератора определить генератор randp, в результате обращения к которому появляется 0 или 1 с вероятностями p и $1 - p$ соответственно (незначительные отклонения допустимы). Оцените сложность в среднем алгоритма получения одного случайного числа с помощью randp (затраты определяются числом обращений к изначально имеющемуся генератору). **Б)** Пусть имеется генератор случайных чисел randp (описанный выше). Известно, что $p \neq 0, \quad p \neq 1$. Как с помощью него сконструировать генератор, в результате обращения которому появляется 0 или 1 с одинаковой вероятностью 1/2. **В)** Чему равно математическое ожидание числа обращений к изначально имеющемуся генератору случайных чисел при построении последовательности пар до появления 0,1 или 1,0? Найти сложность в среднем алгоритма получения k "равновероятных" нулей и единиц с помощью сконструированного генератора (затраты определяются количеством обращений к изначально имеющемуся генератору). Можно ли указать значения p, для которых эта сложность имеет минимальное и, соответственно, максимальное значение?

Решение. To be done $\square$

Задача 140 (сублинейный приближенный вероятностный алгоритм для матричных игр; Григориадис – Хачиян, 1995)**.** Рассматривается симметричная антагонистическая игра двух лиц X и Y. Смешанные стратегии X и Y будем обозначать соответственно $\vec{x}$ и $\vec{y}$. При этом x_k - вероятность того, что игрок X выберет стратегию с номером k, аналогично определяется y_k. Таким образом, $\vec{x},\vec{y} \in S = \left\{\vec{x} \in \mathrm{R}^n : \ \ \vec{e}^T\vec{x} = 1,\ \vec{x} \geq \vec{0}\right\}$, где $\vec{e} = (1,...,1)^T$. Выигрыш игрока X: $V_X(\vec{x},\vec{y}) = \vec{y}^T A\vec{x}$, а выигрыш игрока Y: $V_Y(\vec{x},\vec{y}) = -\vec{y}^T A\vec{x}$ (игра антагонистическая). Каждый игрок стремится максимизировать свой выигрыш, при заданном ходе оппонента. Равновесием Нэша (в смешанных стратегиях) называется такая пара стратегий $(\vec{x}^*,\ \vec{y}^*)$, что

$$\vec{x}^* \in \operatorname{Arg}\max_{\vec{x}\in S} \vec{y}^{*T} A\vec{x}, \quad \vec{y}^* \in \operatorname{Arg}\min_{\vec{y}\in S} \vec{y}^T A\vec{x}^*.$$

Ценой игры называют $\max\limits_{\vec{x}\in S}\min\limits_{\vec{y}\in S} \vec{y}^T A\vec{x} = \min\limits_{\vec{y}\in S}\max\limits_{\vec{x}\in S} \vec{y}^T A\vec{x} = \vec{y}^{*T} A\vec{x}^*$. Поскольку, по условию, игра также симметричная, то $A = -A^T$ - матрица$n \times n$. С помощью стандартной редукции можно свести к этому случаю общий случай произвольной матричной игры. В рассматриваемом же случае цена игры (выигрыш игроков в положении равновесия Нэша) есть 0, а множества оптимальных стратегий игроков совпадают. Требуется найти с точностью $\varepsilon > 0$ положение равновесия Нэша (оптимальную стратегию), т.е. требуется найти такой вектор $\vec{x}$, что $A\vec{x} \leq \varepsilon\vec{e}$, $\vec{x} \in S$. Покажите, считая элементы матрицы A равномерно ограниченными, скажем, единицей, что приводимый ниже алгоритм находит с вероятностью не меньшей 1/2 (вместо 1/2 можно взять любое положительное число меньшее единицы) такой $\vec{x}$ за время $\mathrm{O}\left(\varepsilon^{-2} n \log^2 n\right)$, т.е. в определенном смысле даже не вся матрица (из n^2 элементов) просматривается. Отметим также, что в классе детерминированных алгоритмов, время работы растет с ростом n не медленнее чем $\sim n^2$ (эта нижняя оценка получается из информационных соображений). Другими словами, никакой детерминированный алгоритм не может также асимптотически быстро находить приближенно равновесие Нэша. Точнее говоря, описанный ниже вероятностный алгоритм дает почти квадратичное ускорение по сравнению с детерминированными.

Алгоритм

1. **Инициализация:** $\vec{x} = \vec{U} = \vec{0}$, $\vec{p} = \vec{e}/n$, $t = 0$.
2. **Повторить:**
3. **Счетчик итераций:** $t := t + 1$.
4. **Датчик случайных чисел:** выбираем $k \in \{1,...,n\}$ с вероятностью p_k.
5. **Модификация $\vec{X}$:** $X_k := X_k + 1$.
6. **Модификация $\vec{U}$:** $U_i := U_i + a_{ik}$, $i = 1,...,n$.
7. **Модификация $\vec{p}$:** $p_i := p_i \exp(\varepsilon a_{ik}/2) \Bigg/ \left(\sum\limits_{j=1}^{n} p_j \exp(\varepsilon a_{jk}/2)\right)$, $i = 1,...,n$.

8. **Критерий останова:** если $\vec{U}\Big/t \le \varepsilon\vec{e}$, то останавливаемся и печатаем $\vec{x} = \vec{X}\Big/t$.

Указание. Покажите, что с вероятностью не меньшей, чем 1/2 алгоритм остановится через $t^* = 4\varepsilon^{-2}\ln n$ итераций. Для этого введите $P_i\left(t\right) = \exp\left(\varepsilon U_i\left(t\right)/2\right)$ и $\Phi\left(t\right) = \sum\limits_{i=1}^{n} P_i\left(t\right)$. Покажите, что

$M\left[\Phi\left(t+1\right)|\,\vec{P}\left(t\right)\right] = \Phi\left(t\right)\sum\limits_{i,k=1}^{n} p_i\left(t\right)p_k\left(t\right)\exp\left(\varepsilon a_{ik}/2\right)$ и $\exp\left(\varepsilon a_{ik}/2\right) \le 1 + \varepsilon a_{ik}/2 + \varepsilon^2/6$.

Используя это и кососимметричность матрицы A, покажите

$$M\left[\Phi\left(t+1\right)\right] \le M\left[\Phi\left(t\right)\right]\left(1 + \varepsilon^2/6\right).$$

Следовательно, $M\left[\Phi\left(t\right)\right] \le n\exp\left(t\varepsilon^2/6\right)$ и $M\left[\Phi\left(t^*\right)\right] \le n^{5/3}$. Отсюда по неравенству Маркова имеем, что $(n \ge 8)$

$$P\left(\Phi\left(t^*\right) \le n^2\right) \ge P\left(\Phi\left(t^*\right) \le 2n^{5/3}\right) \ge 1/2.$$

Тогда $P\left(\varepsilon U_i\left(t^*\right)/2 \le 2\ln n,\ i = 1, ..., n\right) \ge 1/2$. Откуда уже следует, что $P\left(\vec{x}\left(t^*\right) \le \varepsilon\vec{e}\right) \ge 1/2$.

Решение. To be done $\square$

6. Раздел 14. Явление концентрации меры

Явление концентрации меры (А. Пуанкаре, П. Леви, В. Мильман, М. Громов, М. Талагран) или геометрическое (изопериметрическое) толкование предельных теорем и законов больших чисел теории вероятностей (объяснение концентрации значений «хороших» функций (таких, например, как хроматическое число графа или перманент матрицы), определенных на случайных комбинаторных объектах (графах или матрицах) большой размерности, в окрестности медианы функции (асимптотически, по размерности объекта, равной математическому ожиданию); а также объяснение «фазовых переходов» в случайных комбинаторных объектах большой размерности), нелинейный закон больших чисел (некоторые результаты В.И. Опойцева), элементы стохастического агрегирования, экспандеры, понятие спектра, предельные формы (А.М. Вершик, Я.Г. Синай)

Задача 141. а) Пусть имеются абсолютно непрерывные (имеющие плотность) независимые случайные величины $X_1, \ldots, X_n$ и $Y_n = G_n(X_1, \ldots, X_n)$ — также случайная величина. Докажите следующее неравенство, описывающее нелинейный закон больших чисел:

$$\operatorname{Var} Y_n \leqslant \int_{\mathbb{R}^n} \sum_{i=1}^{n} \Big(\frac{\partial G_n(x_1, \ldots, x_n)}{\partial x_i} \Big)^2 f_i^*(x_i) \prod_{j \neq i} f_j(x_j)\, dx_1 \ldots dx_n,$$

где сопряженные плотности $f_i^*(x_i)$ существуют и определяются следующим образом

$$f_i^*(x_i) = \mu_i(\infty) \int\limits_{-\infty}^{x_i} f_i(t)\, dt - \mu_i(x_i), \quad \mu_i(x) = \int\limits_{-\infty}^{x} t f_i(t)\, dt.$$

б) Пусть независимые одинаково распределенные случайные величины $X_1, \ldots, X_n$ имеют равномерное распределение на отрезке $[0,1]$ (часто пишут $X_i \in R[0,1]$), а

$$\max_{x \in [0,1]^n} \| \operatorname{grad} G_n(x_1, \ldots, x_n) \| \xrightarrow{n \to \infty} 0.$$

Тогда $Y_n \xrightarrow{P} \mathbb{E} Y_n$ при $n \to \infty$.

в) Пусть независимые одинаково распределенные случайные величины $X_1, \ldots, X_n$ имеют равномерное распределение на отрезке $[0,1]$, а $Y_n = G_n(X_1, \ldots, X_n) = \max\limits_{i=1,\ldots,n} X_i$.

Тогда $Y_n \xrightarrow{P} \mathbb{E} Y_n$ при $n \to \infty$.

Решение. а) Сначала докажем в одномерном случае. Имеем

$$\operatorname{Var} Y = \int\limits_{-\infty}^{\infty} [G(x) - \mathbb{E} G(X)]^2 f(x)\, dx = \frac{1}{2} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} [G(x) - G(y)]^2 f(x) f(y)\, dx\, dy =$$

$$= \int\limits_{-\infty}^{\infty} \int\limits_{y}^{\infty} \Big[\int\limits_{y}^{x} G'(t)\, dt \Big]^2 f(x) f(y)\, dx\, dy.$$

По неравенству Коши-Буняковского получим

$$\Big[\int\limits_y^x G'(t)\,dt\Big]^2 \leqslant (x-y)\int\limits_y^x [G'(t)]^2\,dt,$$

откуда

$$\operatorname{Var} Y \leqslant \int\limits_{-\infty}^{\infty}\int\limits_{y}^{\infty}(x-y)\int\limits_y^x [G'(t)]^2\,dt f(x)f(y)\,dx\,dy.$$

Поменяем местами порядок интегрирования, получим

$$\operatorname{Var} Y \leqslant \int\limits_{-\infty}^{\infty}[G'(t)]^2\int\limits_{t}^{\infty}\int\limits_{-\infty}^{t}(x-y)f(x)f(y)\,dx\,dy\,dt = \int\limits_{-\infty}^{\infty}[G'(t)]^2 f^*(t)\,dt,$$

поскольку

$$\int\limits_{t}^{\infty}\int\limits_{-\infty}^{t}(x-y)f(x)f(y)\,dx\,dy = \int\limits_{t}^{\infty} xf(x)\,dx \int\limits_{-\infty}^{t} f(y)\,dy - \int\limits_{t}^{\infty} f(x)\,dx\int\limits_{-\infty}^{t} yf(y)\,dy =$$

$$= (\mu(\infty)-\mu(t))\cdot\int\limits_{-\infty}^{t} f(y)\,dy - \mu(t)\int\limits_{t}^{\infty} f(y)\,dy = \mu(\infty)\int\limits_{-\infty}^{t} f(y)\,dy - \mu(t) = f^*(t).$$

Поэтому в одномерном случае неравенство доказано.

Далее действуем по индукции, пусть неравенство справедливо в размерности $n-1$. Введем

$$m_{n-1}(x_n) = \int_{\mathbb{R}^{n-1}} G(x)f_1(x_1)\ldots f_{n-1}(x_{n-1})\,dx_1\ldots dx_{n-1},$$

$$D_{n-1}(x_n) = \int_{\mathbb{R}^{n-1}} [G(x)-m_{n-1}(x_n)]^2 f_1(x_1)\ldots f_{n-1}(x_{n-1})\,dx_1\ldots dx_{n-1}.$$

Имеем

$$\operatorname{Var} Y_n = \int_{\mathbb{R}^n}[G(x)-m_{n-1}(x_n)+m_{n-1}(x_n)-\mathbb{E}G(X)]^2 f_1(x_1)\ldots f_n(x_n)\,dx_1\ldots dx_n =$$

$$= \int_{\mathbb{R}} D_{n-1}(x_n)f_n(x_n)\,dx_n + \int_{\mathbb{R}}[m_{n-1}(x_n)-\mathbb{E}G(X)]^2 f_n(x_n)\,dx_n.$$

В одномерном случае уже доказано, что

$$\int_{\mathbb{R}}[m_{n-1}(x_n)-\mathbb{E}G(X)]^2 f_n(x_n)\,dx_n \leqslant \int_{\mathbb{R}}\Big[\frac{dm_{n-1}(x_n)}{dx_n}\Big]^2 f_n^*(x_n)\,dx_n.$$

При этом

$$\Big[\frac{dm_{n-1}(x_n)}{dx_n}\Big]^2 = \Big[\frac{d}{dx_n}\int_{\mathbb{R}^{n-1}} G(x)f_1(x_1)\ldots f_{n-1}(x_{n-1})\,dx_1\ldots dx_{n-1}]^2 \leqslant$$

$$\leqslant \int_{\mathbb{R}^{n-1}} \left(\frac{dG}{dx_n}\right)^2 f_1(x_1)\dots f_{n-1}(x_{n-1})\,dx_1\dots dx_{n-1}.$$

Применяя полученное неравенство и индуктивный переход для величины

$$\int_{\mathbb{R}} D_{n-1}(x_n) f_n(x_n)\,dx_n,$$

получим искомое неравенство в размерности n.

б) Применим неравенство из пункта а): здесь $f_i(t) = \mathbb{I}_{\{t\in[0,1]\}}$, $f_i^*(x) \leqslant 1/2$, поэтому условие

$$\max_{x\in[0,1]^n} \|\mathrm{grad} G_n(x_1,\dots,x_n)\| \xrightarrow{n\to\infty} 0$$

дает $\mathrm{Var}\, Y_n \xrightarrow{n\to\infty} 0$. Тогда из неравенства Чебышева

$$\mathbb{P}(|Y_n - \mathbb{E}Y_n| > \varepsilon) \leqslant \frac{\mathrm{Var}\, Y_n}{\varepsilon^2}, \quad \forall \varepsilon > 0$$

следует, что $Y_n \xrightarrow{P} \mathbb{E}Y_n$ при $n \to \infty$.

в) Функция распределения с.в. Y_n:

$$F(x) = \mathbb{P}(Y_n \leqslant x) = \mathbb{P}(X_1 \leqslant x,\dots,X_n \leqslant x) = \mathbb{P}(X_1 \leqslant x)\cdot\ldots\cdot\mathbb{P}(X_n \leqslant x) = x^n \cdot \mathbb{I}_{\{x\in[0,1]\}}.$$

Отсюда

$$\mathbb{E}Y_n = \int\limits_0^1 x\,dF(x) = \frac{n}{n+1} \xrightarrow{n\to\infty} 1.$$

Имеем

$$\forall \varepsilon > 0,\ \mathbb{P}(|Y_n - \mathbb{E}Y_n| < \varepsilon) = \int\limits_{n/(n+1)-\varepsilon}^{n/(n+1)+\varepsilon} dF(x) =$$

$$= \left[\text{ при больших } n\text{: } n/\,(n+1) + \varepsilon > 1\ \right] = \int\limits_{n/(n+1)-\varepsilon}^{1} dF(x) \geqslant 1 - (1-\varepsilon)^n \xrightarrow{n\to\infty} 1.$$

Это по определению означает, что $Y_n \xrightarrow{P} \mathbb{E}Y_n$ при $n \to \infty$.

$\square$

Задача 142. Пусть X_n — случайный вектор с равномерным распределением на единичной сфере в $\mathbb{R}^n$. Равномерное распределение характеризуется тем, что оно инвариантно относительно группы ортогональных преобразований. Пусть Y_n обозначает первую координату X_n. Докажите, что $\sqrt{n}\,Y_n \xrightarrow{d} N(0,1)$ при $n \to \infty$. Заметим, что в статистической физике с помощью утверждения этой задачи получался закон распределения Максвелла скоростей частиц одномерного идеального газа.

Решение. Пусть $\xi_1, \ldots, \xi_n$ — независимые в совокупности с.в., имеющие одинаковое распределение $N(0,1)$. Рассмотрим случайный вектор $Z_n = (\xi_1, \xi_2, \ldots, \xi_n)$. Тогда $Z_n \in N(0, E_n)$, E_n — единичная матрица размера n. Для всякого ортогонального преобразования $O_n : \mathbb{R}^n \to \mathbb{R}^n$ пространства $\mathbb{R}^n$ случайный вектор

$$O_n Z_n \in N(0, O_n E_n O_n^T) = N(0, E_n), \qquad \text{т.к. } O_n O_n^T = O_n^T O_n = E_n.$$

Получим, что распределение Z_n инвариантно относительно группы ортогональных преобразований. Равномерное распределение X_n отличается тем, что $\|X_n\|_{\mathbb{R}^n} = 1$, поэтому распределения случайных векторов

$$X_n \quad \text{и} \quad \frac{Z_n}{\|Z_n\|_{\mathbb{R}^n}} \quad \text{совпадают.}$$

Поэтому имеет место равенство по распределению с.в.

$$Y_n = \frac{\xi_1}{\sqrt{\xi_1^2 + \ldots + \xi_n^2}}$$

$$\Rightarrow \quad \sqrt{n} Y_n = \frac{\xi_1}{\sqrt{(\xi_1^2 + \ldots + \xi_n^2)/n}}.$$

По теореме Колмогорова для у.з.б.ч. $\frac{\xi_1^2+\ldots+\xi_n^2}{n} \xrightarrow{\text{п.н.}} 1 = \mathbb{E}\xi_k^2$.

Отсюда при $n \to \infty$

$$\sqrt{n}\, Y_n \xrightarrow{d} N(0,1).$$

□

Задача 143 (принцип концентрации площади сферы; А. Пуанкаре, 1911)**.** Покажите, что если в многомерном шаре задано равномерное распределение вероятностей и согласно этому распределению вероятностей сгенерировано два случайных вектора, то с вероятностью близкой к единице концы этих векторов будут лежать почти на границе шара и эти два случайных вектора будут почти ортогональны.

Решение. Объем n-мерного шара радиуса r равен $V_n(r) = \frac{\pi^{n/2} r^n}{\Gamma(n/2+1)}$. Мера границы шара — граничного слоя толщины Δr, равна

$$\Delta V_n(r) = V_n(r) - V_n(r - \Delta r) \Rightarrow \frac{\Delta V_n(r)}{V_n(r)} = 1 - \left(\frac{r - \Delta r}{r}\right)^n \xrightarrow{n\to\infty} 1,$$

т.е. для сколь угодно малой толщины Δr при достаточно больших n почти вся мера шара сосредоточена на граничном слое толщины Δr.

Далее, для доказательства того, что сгенерированные два случайных вектора являются почти ортогональными, не ограничивая общности, полагаем, что они имеют единичную длину и один из них равен $\xi_1 = (1, 0, \ldots, 0)$.

Первый способ.

Тогда другой из векторов x имеет равномерное распределение на единичной сфере S^{n-1} из $\mathbb{R}^n$.

Пусть $l_n = \mathbb{E}\|x-\xi_1\|$ — среднее расстояние между точками x и ξ_1, $b_n^2 = \operatorname{Var}\|x-\xi_1\|$ — среднеквадратическое отклонение этого расстояния.

Поскольку $\|x-\xi_1\|^2 = 2(1-\langle x,\xi_1\rangle)$, для доказательства почти ортогональности векторов x и ξ_1 достаточно показать, что $l_n \xrightarrow{n\to\infty} \sqrt{2}$ и $b_n^2 \xrightarrow{n\to\infty} 0$. Действительно, тогда из неравенства Чебышева получим $\forall\varepsilon>0$,

$$\mathbb{P}(|\|x-\xi_1\| - l_n| > \varepsilon) \le \frac{b_n^2}{\varepsilon^2} \xrightarrow{n\to\infty} 0,$$

т.е. $\|x-\xi_1\| \xrightarrow{P} \sqrt{2}$ и $\langle x,\xi_1\rangle \xrightarrow{P} 0$.

Запишем

$$l_n = \mathbb{E}\|x-\xi_1\| = \frac{1}{\sigma_{n-1}}\int\limits_{S^{n-1}} \|x-\xi_1\|\,d\mu_{n-1} = \sqrt{2}\frac{1}{\sigma_{n-1}}\int\limits_{S^{n-1}} \sqrt{1-\langle x,\xi_1\rangle}\,d\mu_{n-1}.$$

В силу неравенства Коши-Буняковского

$$\frac{1}{\sigma_{n-1}}\int\limits_{S^{n-1}} \sqrt{1-\langle x,\xi_1\rangle}\,d\mu_{n-1} \le \frac{1}{\sigma_{n-1}}\sqrt{\int\limits_{S^{n-1}} (1-\langle x,\xi_1\rangle)\,d\mu_{n-1}}\sqrt{\int\limits_{S^{n-1}} d\mu_{n-1}} = 1. \qquad (6.1)$$

Определим для всякого $\varepsilon\in(0,1)$ множество $S_\varepsilon^{n-1} = \{x\in S^{n-1}\,|\,\langle x,\xi_1\rangle > \varepsilon\}$. В силу неравенства Чебышева мера множества S_ε^{n-1} оценивается как

$$\mu_{n-1}S_\varepsilon^{n-1} \le \frac{1}{\varepsilon}\int_{S_{n-1}} |\langle x,\xi_1\rangle|\,d\mu_{n-1} \le$$

$$\le \frac{1}{\varepsilon}\sqrt{\int\limits_{S^{n-1}} \langle x,\xi_1\rangle^2\,d\mu_{n-1}}\sqrt{\int\limits_{S^{n-1}} d\mu_{n-1}} = \frac{1}{\varepsilon}\sqrt{\int\limits_{S^{n-1}} \frac{1}{n}\|x\|^2\,d\mu_{n-1}}\sqrt{\int\limits_{S^{n-1}} d\mu_{n-1}} = \frac{\sigma_{n-1}}{\varepsilon\sqrt{n}}.$$

Итак,

$$\frac{\mu_{n-1}S_\varepsilon^{n-1}}{\sigma_{n-1}} \le \frac{1}{\varepsilon\sqrt{n}}.$$

Отсюда $\forall\varepsilon>0\ \exists n_\varepsilon = \left[\frac{1}{\varepsilon^4}\right]+1$, так что $\forall n>n_\varepsilon$ имеем оценку

$$\frac{1}{\sigma_{n-1}}\int\limits_{S^{n-1}} \sqrt{1-\langle x,\xi_1\rangle}\,d\mu_{n-1} \ge \frac{1}{\sigma_{n-1}}\int\limits_{S^{n-1}\setminus S_\varepsilon^{n-1}} \sqrt{1-\varepsilon}\,d\mu_{n-1} =$$

$$= \frac{\sigma_{n-1}-\mu_{n-1}S_\varepsilon^{n-1}}{\sigma_{n-1}}\sqrt{1-\varepsilon} \ge$$

$$\ge \sqrt{1-\varepsilon}\Big(1-\frac{1}{\varepsilon\sqrt{n}}\Big) \ge \sqrt{1-\varepsilon}(1-\varepsilon) = (1-\varepsilon)^{3/2}.$$

Вкупе с формулой (6.1) это значит, что

$$\lim_{n\to\infty}\frac{1}{\sigma_{n-1}}\int\limits_{S^{n-1}} \sqrt{1-\langle x,\xi_1\rangle}\,d\mu_{n-1} = 1,$$

откуда $\lim\limits_{n\to\infty} l_n = \sqrt{2}$.

Далее,

$$\mathbb{E}\|x-\xi_1\|^2 = \frac{1}{\sigma_{n-1}} \int\limits_{S^{n-1}} \|x-\xi_1\|^2\, d\mu_{n-1} = \frac{2}{\sigma_{n-1}} \int\limits_{S^{n-1}} (1-\langle x,\xi_1\rangle)\, d\mu_{n-1} = 2$$

$\Rightarrow\ b_n^2 = \mathbb{E}\|x-\xi_1\|^2 - l_n^2 \xrightarrow{n\to\infty} 0.$

Второй способ.

Достаточно доказать, что для всякого сколь угодно малого $\delta > 0$ проекция второго вектора на ось x_1 с вероятностью, близкой к единице, лежит в промежутке $[-\delta,\delta]$ при $n\to\infty$. Это равносильно тому, что доля от площади всей сферы $S^{n-1}(r)$, которую занимает сферический слой $S_\delta^{n-1}(r)$, проектирующийся в отрезок $[-\delta,\delta]$ оси x_1, может быть сделана сколь угодно близкой к 1 при $n\to\infty$.

Переходя к n-мерным сферическим координатам и обратно, находим меру сферического слоя $S_\delta^{n-1}(r)$:

$$\mu_{n-1}S_\delta^{n-1}(r) = \int\limits_{\varphi_1^{(2)}}^{\varphi_1^{(1)}} \int\limits_0^{\pi} \dots \int\limits_0^{\pi} r^{n-1}\sin^{n-2}\varphi_1 \sin^{n-3}\varphi_2 \dots \sin\varphi_{n-2}\, d\varphi_1\dots d\varphi_{n-1} =$$

$$= Cr^{n-1}\int\limits_{\varphi_1^{(2)}}^{\varphi_1^{(1)}} \sin^{n-2}\varphi_1\, d\varphi_1 = Cr^{n-1}\int\limits_{-\delta}^{\delta}\Big(1-(x/r)^2\Big)^{(n-3)/2} dx.$$

Вероятность попадания в данный слой $S_\delta^{n-1}(r)$ равна

$$\mathbb{P}[-\delta,\delta] = \frac{\int_{-\delta}^{\delta}\Big(1-(x/r)^2\Big)^{(n-3)/2} dx}{\int_{-r}^{r}\Big(1-(x/r)^2\Big)^{(n-3)/2} dx}.$$

Данное отношение на зависит от r, поэтому можно считать $r=1$. Для нахождения асимптотики имеющихся интегралов при $n\to\infty$ применим классические результаты относительно асимптотики интеграла Лапласа $F(\lambda) = \int_a^b f(x)e^{\lambda S(x)}\,dx$ при $\lambda\to+\infty$. Если обе функции f и S определны и регулярны на промежутке $I=[a,b]$ и функция S имеет единственный глобальный максимум на I, который достигается в точке $x_0\in I$, $f(x_0)\neq 0$, то асимптотика интеграла такая же, как в окрестности точки x_0 (принцип локализации). В зависимости от расположения точки x_0 и свойств функции $S(x)$ возможны следующие тейлоровские разложения при $\lambda\to+\infty$:

$$F(\lambda) = \frac{f(x_0)}{-S'(x_0)}e^{\lambda S(x_0)}\lambda^{-1}\big(1+O(\lambda^{-1})\big),$$

если $x_0=a$ и $S'(x_0)\neq 0$ (т.е. $S'(x_0)<0$);

$$F(\lambda) = \sqrt{\frac{\pi}{-2S''(x_0)}}f(x_0)e^{\lambda S(x_0)}\lambda^{-1/2}\big(1+O(\lambda^{-1/2})\big),$$

если $x_0 = a$, $S'(x_0) = 0$, $S''(x_0) \neq 0$ (т.е. $S''(x_0) < 0$);

$$F(\lambda) = \sqrt{\frac{2\pi}{-S''(x_0)}} f(x_0) e^{\lambda S(x_0)} \lambda^{-1/2} \big(1 + O(\lambda^{-1/2})\big),$$

если $a < x_0 < b$, $S'(x_0) = 0$, $S''(x_0) \neq 0$ (т.е. $S''(x_0) < 0$).

Третий случай применим для интеграла

$$\int_{-1}^{1} \Big(1 - x^2\Big)^{(n-3)/2} dx = \int_{-1}^{1} \exp\Big(\frac{n-3}{2} \ln(1 - x^2)\Big)\, dx \sim \sqrt{\frac{2\pi}{n}},$$

а первая формула — для интеграла

$$\int_{\delta}^{1} \Big(1 - x^2\Big)^{(n-3)/2} dx = \int_{\delta}^{1} \exp\Big(\frac{n-3}{2} \ln(1 - x^2)\Big)\, dx \sim \frac{1}{n\delta}(1 - \delta^2)^{(n-1)/2}.$$

Отсюда находим

$$\mathbb{P}[-\delta, \delta] = \frac{\int\limits_{-\delta}^{\delta} \Big(1 - x^2\Big)^{(n-3)/2} dx}{\int\limits_{-1}^{1} \Big(1 - x^2\Big)^{(n-3)/2} dx} \sim 1 - \frac{\sqrt{2}}{\delta\sqrt{\pi n}}(1 - \delta^2)^{(n-1)/2} \xrightarrow{n\to\infty} 1.$$

□

Задача 144 (изопериметрическое неравенство и принцип концентрации меры; П. Леви, 1919)**.** Число μ_f называют медианой функции f, если

$\mu\left(\vec{x} \in S_1^n : \;\; f\left(\vec{x}\right) \geq \mu_f\right) \geq 1/2$ и $\mu\left(\vec{x} \in S_1^n : \;\; f\left(\vec{x}\right) \leq \mu_f\right) \geq 1/2$,

где $\mu\left(d\vec{x}\right)$ - равномерное мера на единичной сфере S_1^n в R^n. Пусть A - измеримое (борелевское) множество на сфере S_1^n. Через A_δ - будем обозначать δ-окрестность множества A на сфере S_1^n. Предположим теперь, что в некотором царстве, расположенном на S_1^n, царь предложил царице Диодоне построить огород с заданной длиной забора. Царица хочет, чтобы её огород при заданном периметре имел наибольшую площадь. Таким образом, царице надо решить изопериметрическую задачу (такие задачи обычно рассматриваются в курсах вариационного исчисления). Решение этой задачи хорошо известно – "круглый огород". Для нас же полезно, рассмотрение двойственной задачи, имеющей такое же решение: при заданной площади огорода спроектировать его так, чтобы он имел наименьшую длину забора, его ограждающего. Используя решение этой задачи, покажите, что если $\mu\left(A\right) \geq 1/2$, то

$$\mu\left(A_\delta\right) \geq 1 - \sqrt{\pi/2} \exp\left(-\delta^2 n/2\right).$$

Пусть теперь на S_1^n задана функция с модулем непрерывности

$$\omega_f\left(\delta\right) = \sup\left\{\left|f\left(\vec{x}\right) - f\left(\vec{y}\right)\right| : \;\; \rho\left(\vec{x}, \vec{y}\right) \leq \delta, \; \vec{x}, \vec{y} \in S_1^n\right\}.$$

Тогда
$$\mu\left(\vec{x} \in S_1^n:\quad \left|f\left(\vec{x}\right) - \mu_f\right| \geq \omega_f\left(\delta\right)\right) \leq \sqrt{\pi/2}\exp\left(-\delta^2 n/2\right).$$
Можно показать, что при весьма естественных условиях медиана асимптотически близка к среднему значению (математическому ожиданию). Аналогичное неравенство можно получить (М. Талагран, 1994), например, для модели случайных графов (Эрдёша - Реньи). И исследовать плотную концентрацию около среднего значения различные функций на случайных графов: число независимости, хроматическое число и т.п..

Решение. To be done □

7. Раздел 15. Геометрические вероятности

Геометрические вероятности (задача об игле Бюффона и необходимые понятия интегральной геометрии, средняя площадь проекции и гауссова кривизна (формула Гаусса-Бонне))

Задача 145. Стержень длины l разрезан в двух случайно выбранных точках. Определить вероятность того, что из полученных частей стержня (отрезков) можно построить:

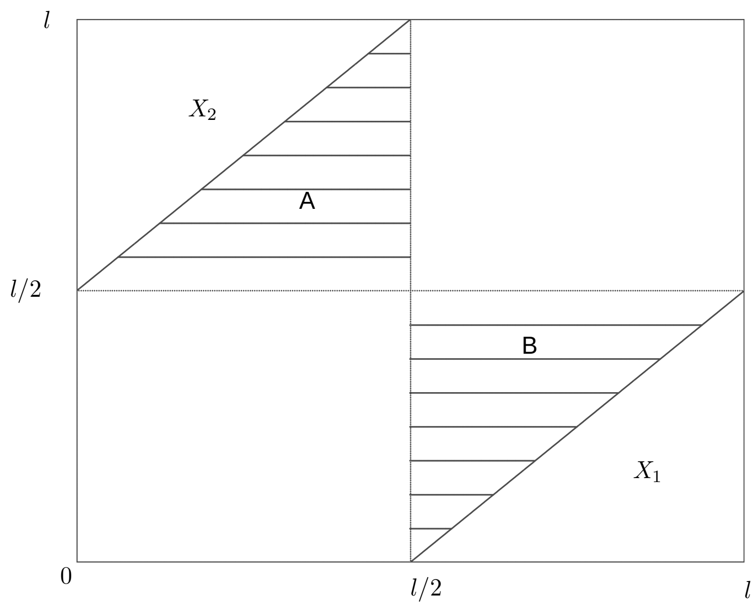
а) треугольник;

б) прямоугольный треугольник.

Решение. Обозначим: X_1 — координата точки первого разреза, X_2 — координата точки второго разреза.

а) Условие возможности построения треугольника: $(X_1, X_2) \in A \cup B$ (см. рисунок ниже), $p = \mathbb{P}\{(X_1, X_2) \in A \cup B\} = 1/4$ (равномерное распределение);

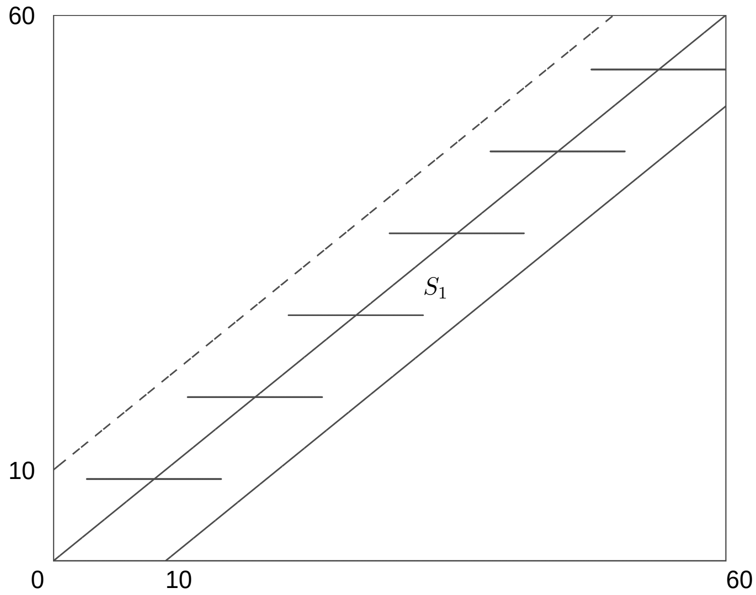
б) условие возможности построения прямоугольного треугольника имеет вид $(X_1, X_2) \in C$, где C — множество точек, связанных функционально теоремой Пифагора и, ввиду непрерывности распределений случайных величин X_1 и X_2, имеющее нулевую меру. Поэтому вероятность построения прямоугольного треугольника равна здесь нулю.



□

Задача 146. Двое условились о встрече между 10 и 11 часами утра, причем договорились ждать друг друга не более 10 минут. Считая, что момент прихода на встречу каждым выбирается «наудачу» в пределах указанного часа, найти вероятность того, что встреча состоится.

Решение. Моменты появления участников встречи образуют двумерный вектор (T_1, T_2), равномерно распределенный на квадрате $S = [0,60] \times [0,60]$ с площадью $|S| = 60^2$. Встреча состоится, если $|T_1 - T_2| < 10$, что соответствует попаданию точки (T_1, T_2) в заштрихованную область S_1 с площадью $|S_1| = 1100$ (см. рис ниже). Вероятность встречи равна $\mathbb{P}(A) = \dfrac{11}{36}$.



$\square$

Задача 147 (изогнутая игла Бюффона)**.** Любопытный студент швейного техникума решил повторить опыты Бюффона по бросанию иглы (студент хочет оценить число π). Для этого он подготовил горизонтально расположенный лист бумаги, разлинованный параллельными прямыми так, что расстояние между соседними прямыми равно 1. Однако в распоряжении студента оказалось только погнутая иголка. Иголка имеет форму кочерги, но студент не имеет точного представления о том, как именно погнута иголка. Ему известно лишь то, что длина иголки, до того как она погнулась, была равна 2. Студент бросил погнутую иголку 1000000 раз и посчитал суммарное число пересечений, учитывая кратность. Помогите студенту оценить число π:

а) с помощью неравенства Чебышёва;

б) с помощью ц.п.т. (центральной предельной теоремы) и оценок скорости сходимости в ц.п.т., например, с помощью неравенства Берри–Эссена или более точных аппроксимаций.

Решение. Очевидно, любые два прямолинейных участка иголки равной длины $L \leq 1$ могут пересечь прямую с одинаковой вероятностью $p < 1$ (не более одного раза). Отсюда следует, что математическое ожидание числа попаданий иглы на прямую линейно зависит от длины иглы.

Сначала рассмотрим окружность диаметра 1, т.е. длины π. Такая окружность с вероятностью 1 пересекает дважды одну из прямых. Тогда, исходя из линейности математического ожидания числа попаданий иглы на прямую относительно длины иглы, для иглы длиной $L < 1$ имеем $\mathbb{E}\xi_L = 2L/\pi$. Но в случае иглы длиной $L < 1$ число попаданий на прямую есть бернуллиевскя случайная величина, поэтому $\xi_L \in \mathrm{Be}(2L/\pi)$. Разбив исходную иголку длины 2 в форме кочерги на три линейных участка с длинами $L_i \le 1$, $i = 1$, 2, 3, находим число попаданий иглы на прямую $\xi = \xi_{L_1} + \xi_{L_2} + \xi_{L_3}$, где $\xi_{L_i} \in \mathrm{Be}(2L_i/\pi)$, $L_1 + L_2 + L_3 = 2$. Отсюда

$$\mathbb{E}\xi = \frac{4}{\pi},$$

$$\operatorname{Var}\xi = \sum_{i,j=1}^{3} \operatorname{cov}(\xi_{L_i}, \xi_{L_j}) \le \Big(\sqrt{\operatorname{Var}\xi_{L_1}} + \sqrt{\operatorname{Var}\xi_{L_2}} + \sqrt{\operatorname{Var}\xi_{L_3}}\Big)^2.$$

Т.к. $\operatorname{Var}\xi_{L_i} = 2L_i/\pi\big(1 - 2L_i/\pi\big)$, $\max\limits_{L_i}\big\{2L_i/\pi\big(1 - 2L_i/\pi\big)\big\} = 1/4$, получим

$$\operatorname{Var}\xi \le \left(\frac{3}{2}\right)^2 = \frac{9}{4}.$$

1) Итак, имеется $N = 1000000$ реализаций ξ_i, $\mathbb{E}\xi_i = 4/\pi$, $\operatorname{Var}\xi_i \le 9/4$. Для среднего числа попаданий $X_N = S_N/N$, где $S_N = \xi_1 + \ldots + \xi_N$, запишем неравенство Чебышева

$$\mathbb{P}\Big(\Big|X_N - \frac{4}{\pi}\Big| < \varepsilon\Big) \ge 1 - \frac{\operatorname{Var} X_N}{\varepsilon^2} \ge 1 - \frac{9}{4N\varepsilon^2}$$

Если нужно оценить число π с доверительной вероятностью $\delta = 0.99$, достаточно потребовать

$$1 - \frac{9}{4N\varepsilon^2} = \delta = 0.99 \Rightarrow \varepsilon = \sqrt{\frac{225}{N}} = 0.015.$$

Отсюда с вероятностью $\delta = 0.99$ выполнено

$$\Big|X_N - \frac{4}{\pi}\Big| < \varepsilon \Rightarrow \pi \in \Big(\tfrac{4}{X_N + 0.015}, \ \tfrac{4}{X_N - 0.015}\Big).$$

2) Близость величины $\frac{S_N - \mathbb{E}S_N}{\sqrt{\operatorname{Var} S_N}}$ к стандартной нормально распределенной (согласно ц.п.т.) в смысле близости их функций распределения определяется из неравенства Берри-Эссена

$$\sup_x \Big|\mathbb{P}\Big(\frac{S_N - \mathbb{E}S_N}{\sqrt{\operatorname{Var} S_N}} < x\Big) - \Phi(x)\Big| \le \frac{C_0\mu^3}{\sigma^3\sqrt{N}},$$

где $C_0 < 0.7056$, $\mu^3 = \mathbb{E}|\xi_i - \mathbb{E}\xi_i|^3$, $\sigma^2 = \operatorname{Var}\xi_i$, $\Phi(x) = \int_{-\infty}^{x} \frac{e^{-t^2/2}}{\sqrt{2\pi}}\, dt$.

Обозначив событие $A = \Big\{\frac{S_N - \mathbb{E}S_N}{\sqrt{\operatorname{Var} S_N}} \in (-3, 3)\Big\}$, получим

$$|\mathbb{P}(A) - [\Phi(3) - \Phi(-3)]| \le \frac{2C_0\mu^3}{\sigma^3\sqrt{N}}.$$

Имеем $\operatorname{Var} S_N = N \cdot \operatorname{Var}\xi_i$, где $\operatorname{Var}\xi_i \le \frac{9}{4}$. Далее, остается оценить

$$\frac{\mu^3}{\sigma^3} = \frac{p_1(4/\pi)^3 + p_2(4/\pi - 1)^3 + p_3(2 - 4/\pi)^3}{\big(p_1(4/\pi)^2 + p_2(4/\pi - 1)^2 + p_3(2 - 4/\pi)^2\big)^{3/2}} = F(p_1, p_2, p_3),$$

где $p_1 = \mathbb{P}(\xi_i = 0)$, $p_2 = \mathbb{P}(\xi_i = 1)$, $p_3 = \mathbb{P}(\xi_i = 2)$, $p_1 + p_2 + p_3 = 1$ (т.к. изогнутая игла длины 2 может пересекать не более двух прямых).

Численные результаты показывают, что можно считать $\max\limits_{p_1+p_2+p_3=1} F(p_1, p_2, p_3) \approx 2$. Отсюда

$$|\mathbb{P}(A) - [\Phi(3) - \Phi(-3)]| \leq \frac{2C_0\mu^3}{\sigma^3\sqrt{N}} \leq \frac{2.8}{\sqrt{N}} = 2.8 \cdot 10^{-3}.$$

Учитывая $[\Phi(3) - \Phi(-3)] \approx 0.9973$, получим

$$\mathbb{P}(A) \geq [\Phi(3) - \Phi(-3)] - 2.8 \cdot 10^{-3} \approx 0.9945.$$

Значит, с вероятностью $\delta = 0.9945$ выполнено событие A:

$$\left|\frac{S_N - \mathbb{E}S_N}{\sqrt{\operatorname{Var} S_N}}\right| < 3 \Rightarrow \left|X_N - \frac{4}{\pi}\right| \leq \frac{3\sqrt{\operatorname{Var}\xi_i}}{\sqrt{N}} < 4.5 \cdot 10^{-3}$$

$$\Rightarrow \pi \in \Big(\frac{4}{X_N + 0.0045}, \frac{4}{X_N - 0.0045}\Big).$$

□

Задача 148 (средняя площадь поверхности и интегральная геометрия)**.** Покажите, что средняя площадь ортогональной проекции куба с ребром единица на случайную плоскость равна 3/2.

Решение. Очевидно, средняя ортогональная проекция любой плоской области определяется только площадью этой области. Аппроксимируя границу всякого измеримого выпуклого тела многогранниками и переходя к пределу, получим, что средняя площадь S_1 ортогональной проекции всякого измеримого тела линейно зависит от площади его границы V_1.

Поэтому рассмотрим такое тело, у которого легко вычисляется средняя площадь ортогональной проекции. Таким телом является единичный шар B, поскольку ортогональная проекция B на любую плоскость является единичным кругом, чья площадь равна $S_1^B = \pi$. Учитывая, что площадь границы данного шара $V_1^B = 4\pi$, площадь границы единичного куба $V_1^C = 6$, получим

$$\frac{S_1^C}{S_1^B} = \frac{V_1^C}{V_1^B} \Rightarrow S_1^C = S_1^B \frac{V_1^C}{V_1^B} = \pi\frac{6}{4\pi} = \frac{3}{2}.$$

□

Замечание. Обозначим через S_k k-мерный объем ортогональной проекции рассматриваемой области в $\mathbb{R}^n$ на случайную k-мерную плоскость, или, что то же самое, среднее значение (усредненное по всем k-мерным плоскостям, предполагаемым равновероятными) площади ортогональной k-мерным проекции области. Оказывается, что S_k также равны средним значениям (усредненным по поверхности рассматриваемой области) симметрических функций от главных кривизн поверхности, и участвуют в (удивительной) формуле для объема h-окрестности этой области:

$$V(h) = V_0 + V_1 h + V_2 h^2 + ... + V_n h^n,$$

где V_0 – объем области; V_1 – $(n-1)$-мерный объем границы области, пропорциональный среднему значению от числа 1; число V_k пропорционально S_k и выражается через средние значения от произведений k главных кривизн. В случае $n = 3$, из главных кривизн k_1 и k_2 в каждой точке можно составить *среднюю кривизну* $k_1 + k_2$ и *гауссову кривизну* $K = k_1 k_2$. В этом случае объем h-окрестности получается $V(h) = V_0 + V_1 h + V_2 h^2 + V_3 h^3$, где V_2 пропорционально интегралу от средней кривизны по всей поверхности, а V_3 – от гауссовской:

$$V_3 = \frac{4}{3}\pi \iint K dS.$$

Например, для сферы радиуса R

$$V(h) = \frac{4}{3}\pi \cdot (R+h)^3 = \frac{4}{3}\pi R^3 + h \cdot \left(4\pi R^2\right) + h^2 (4\pi R) + \frac{4}{3}\pi h^3.$$

Здесь

$$k_1 = k_2 = 1/R, \quad k_1 + k_2 = 2/R, \quad k_1 k_2 = 1/R^2, \quad \iint (k_1 + k_2)\, dS = 8\pi R,$$

$\iint (k_1 k_2)\, dS = 4\pi$ (формула Гаусса-Бонне).

Коэффициент V_3 не зависит от деталей области, а зависит только от *эйлеровой характеристики* поверхности рассматриваемой области. Это обстоятельство привело Г. Вейля к созданию теории характеристических классов и чисел, обобщающих формулу Гаусса-Бонне.

8. Раздел 17. Понятие количества информации

Понятие количества информации (теорема Шеннона-Макмиллана)

Задача 149. Пусть буква X — дискретная с.в., принимающая значения из алфавита $(x_1, \dots, x_m)$ с вероятностями $(p_1, \dots, p_m)$. Имеется случайный текст из n букв X (предполагается, что буквы в тексте независимы друг от друга). Общее количество таких текстов $2^{n\log m}$. Поэтому можно закодировать все эти слова, используя $n\log m$ бит. Однако, используя то обстоятельство, что $(p_1, \dots, p_m)$ — в общем случае неравномерное распределение, предложите лучший способ кодирования, основанный на усиленном законе больших чисел.

Решение. Пусть $\Omega = \{\omega:\ \omega = (X_1, X_2, \dots, X_n),\ X_i \in 1, 2, \dots, m\}$ — пространство элементарных исходов. Вероятность появления слова $\omega = (X_1, X_2, \dots, X_n)$ равна $p(\omega) = p_{X_1}\cdot\ldots\cdot p_{X_n}$. По теореме Колмогорова об у.з.б.ч.

$$-\frac{1}{n}\log p(\omega) = -\frac{1}{n}\sum_{i=1}^{n}\log p_{X_i} \xrightarrow{\text{п.н.}} -\mathbb{E}p(\omega) = -\sum_{i=1}^{m} p_i \log p_i = H(p) \qquad \text{при } n \to \infty.$$

В частности, $\frac{S_n}{n} \xrightarrow{P} H(p)$, где $S_n = -\log p(\omega) = -\sum\limits_{i=1}^{n}\log p_{X_i}$. Это можно записать в виде

$$\mathbb{P}\Big(\Big|\frac{S_n - nH(p)}{n}\Big| > \delta\Big) \xrightarrow{n\to\infty} 0. \tag{8.1}$$

Будем называть текст ω δ-типичным, если

$$2^{-n(H(p)+\delta)} < p(\omega) < 2^{-n(H(p)-\delta)}.$$

1) Существует не более $2^{n(H(p)+\delta)}$ типичных текстов, поскольку

$$p(\omega) > 2^{-n(H(p)+\delta)}, \quad \sum_{\omega\in\Omega} p(\omega) = 1.$$

2) Для $n > n(\varepsilon, \delta)$ существует, по крайней мере, $(1-\varepsilon)2^{n(H(p)-\delta)}$ типичных текстов, т.к.

$$p(\omega) < 2^{-n(H(p)-\delta)},$$

$$2^{-n(H(p)+\delta)} < p(\omega) < 2^{-n(H(p)-\delta)} \ \Leftrightarrow\ |-\log p(\omega) - nH(p)| < n\delta,$$

$$(8.1) \Rightarrow \quad \text{при } n > n(\varepsilon, \delta) \quad \mathbb{P}\Big(\Big|\frac{S_n - nH(p)}{n}\Big| < \delta\Big) = \mathbb{P}(|-\log p(\omega) - nH(p)| < n\delta) > 1-\varepsilon.$$

3) Множество нетипичных текстов имеет вероятность $\leqslant \varepsilon$, исходя из (8.1).

Из п.1) следует, что можно осуществить эффективное кодирование данных, используя все двоичные последовательности длины $n(H(p)+\delta)$, чтобы закодировать все δ-типичные тексты и отбросить нетипичные. Из п.3) следует, что вероятность ошибки при таком кодировании будет не больше ε. Из п.2) следует, что любой код, использующий двоичные последовательности длины $n(H(p)-\delta)$, имеет асимптотически не исчезающую вероятность ошибки, стремящуюся к единице при $n \to \infty$.

Функцию $H(p)$, которую можно проинтерпретировать как меру количества информации (в битах на передаваемый символ) в случайном тексте, называют энтропией. $nH(p)$ характеризует меру неопределённости случайного текста. При этом для

$$p^0 = (p_1^0, \dots, p_m^0) = (1/m, \dots, 1/m)$$

энтропия максимальна: $H(p^0) = \log m$, и эффективное кодирование невозможно.

Действительно, из неравенства Йенсена для выпуклой функции $g(x) = x \cdot \ln x$, $g(0) = 0$, беря в качестве узлов $p_i, i = 1, .., m$, в качестве весов — величины $\frac{1}{m}$, получим

$$H(p^0) = -\ln\frac{1}{m} = -m \cdot \left(\sum_{i=1}^{m} \frac{p_i}{m}\right) \cdot \ln\left(\sum_{i=1}^{m} \frac{p_i}{m}\right) \geqslant$$

$$\geqslant -m \cdot \sum_{i=1}^{m} \frac{1}{m} p_i \ln p_i = H(p).$$

$\square$

Задача 150 (о шляпах Тода Эберта (1998)). А) Трех игроков отводят в комнату, где на них надевают (случайно и независимо) белые и черные шляпы. Каждый видит цвет других шляп и должен написать на бумажке одно из трех слов: «белый», «черный», «пас» (не советуясь с другими и не показывая им свою бумажку). Команда выигрывает, если хотя бы один из игроков назвал правильный цвет своей шляпы и ни один не назвал неправильного. Как им сговориться, чтобы увеличить шансы?

Б) Решите эту же задачу, если игроков $n = 2^m - 1$ ($m \in \mathbb{N}$).

Решение. 1) Сначала докажем, что существует 2^{2^n-n-1} кодов Хемминга, состоящих из $2^n - 1$ элементов. Пронумеруем разряды в двоичной системе от 00...1 до 11...1 (из n разрядов).

Разряды с номерами 2^k, $k = \overline{0, .., n-1}$ назовем вспомогательными, все остальные разряды (их назовем основные) разделим на n пересекающихся множеств (групп). Разряд с номером l назовем принадлежащим к группе p, если в двоичной записи этого разряда на p-ом месте с конца стоит 1. Таким образом, каждый разряд принадлежит как минимум двум группам.

Теперь во все не вспомогательные (основные) разряды (их $2^n - n - 1$) произвольным образом расставим 0 и 1. Теперь в каждый разряд с номером 2^k поставим таким образом 0 или 1, чтобы в нем и в $(k+1)$-й группе в сумме было четное число единиц.

Докажем, что в результате любые два из полученных кодов будут отличаться как минимум на 3 символа.

Пусть $(\alpha_1, \dots, \alpha_{2^n-n-1})$ и $(\beta_1, \dots, \beta_{2^n-n-1})$ — две произвольные последовательности 0 и 1 в основных разрядах. Если они отличаются хотя бы на 3 символа, то предположение доказано.

а) Если они отличаются на два символа: $\alpha_i \neq \beta_i$, $\alpha_j \neq \beta_j$, то, не ограничивая общности, i-й разряд принадлежит к группе p, к которой на принадлежит j-й разряд. Тогда на 2^{p-1}-м месте у кодов стоят разные цифры, т.е. коды отличаются хотя бы на 3 символа.

б) Если они отличаются на один символ: $\alpha_i \neq \beta_i$, то i-й разряд принадлежит хотя бы двум группам p и q, и на 2^{p-1}-м и 2^{q-1}-м местах у кодов стоят разные цифры, т.е. коды отличаются хотя бы на 3 символа.

Очевидно, имеется 2^{2^n-n-1} кодов. Обозначим через H множество таких кодов. Пусть L — множество кодов, отличающихся от какого-то кода из H ровно на один символ. Тогда каждый из таких кодов не может отличаться от какого-то еще кода из H на один символ, поскольку доказано выше, что коды из H отличаются как минимум на три символа. Отсюда находим $|L| = |H| \cdot (2^n - 1) = 2^{2^n-n-1} \cdot (2^n - 1)$.

2) Обозначим один цвет шляпы — 0, другой — 1. Дадим i-му игроку следующие инструкции: если $\exists u$ (0 или 1), так что

$$\alpha_1 \ldots \alpha_{i-1} u \, \alpha_{i+1} \ldots \alpha_{2^n-1} \in H,$$

$$\alpha_1 \ldots \alpha_{i-1} \overline{u} \, \alpha_{i+1} \ldots \alpha_{2^n-1} \in L,$$

тогда i-й игрок говорит, что цвет его шляпы — $\overline{u}$. Иначе — пасует. Здесь α_j — цвет шляпы j-го игрока.

Найдем вероятность выигрыша. Все исходные 2^{2^n-1} расположения шляп на игроках равновероятны. Докажем, что если расположение шляп на игроках $m \in L$, то они выиграют. Действительно, найдется единственный номер i такой, что

$$m = m_1 \ldots m_i \ldots m_{2^n-1}, \quad m' = m_1 \ldots \overline{m_i} \ldots m_{2^n-1} \in H,$$

тогда игрок под номером i угадает цвет своей шляпы, а все остальные скажут «пас».

Если же $m \in H$, то все игроки дадут неверный ответ, и команда проиграет.

Значит, веротяность того, что команда не выиграет, равна $\dfrac{|H|}{|H| + |L|} = \dfrac{1}{2^n}$, т.е.

$$\mathbb{P}(\text{«выигрыша»}) = 1 - \frac{1}{2^n}.$$

3) Заметим, что при $n = 2$ есть $2^2 - 1 = 3$ игрока и вероятность выигрыша $1 - \frac{1}{2^2} = \frac{3}{4}$, т.е. существует стратегия, при которой вероятность выигрыша $\frac{3}{4}$.

Докажем, что стратегий лучше не бывает.

Единственная информация, которой владеет i-й игрок — это цвета шляп двух других. Поэтому стратегия для i-го игрока должна зависеть только от этих двух цветов. В каждом случае имеется три варианта ответа для игрока: 0, 1 или «пас», т.е. всего 3^{12} различных стратегий. Поскольку есть 8 вариантов расположения шляп на игроках, более выгодная стратегия должна обеспечивать выигрыш в 7 вариантах. Тогда один из игроков должен угадать свой цвет в 3 ситуациях. Значит, имеются для него ответы $\alpha_{i_1 j_1}$, $\alpha_{i_2 j_2}$, не являющиеся пасами. Но тогда в ситуациях $\overline{\alpha_{i_1 j_1}} i_1 j_1$ и $\overline{\alpha_{i_2 j_2}} i_2 j_2$ он ошибется, что противоречит предположению о 7 выигрышных ситуациях.

Таким образом, максимальная вероятность выигрыша равна $\frac{3}{4}$.

4) Аналогичным образом для всякого n докажем, что приведенная стратегия дает наибольшую вероятность выигрыша. i-й игрок обладает информацией о цветах шляп остальных игроков, и в зависимости от этого должен сказать 0, 1 или «пас». Имеется конечное число различных стратегий игроков, и если существует стратегия с вероятностью проигрыша меньше $\frac{1}{2^n}$, то ситуаций, в которых какой-то игрок угадывает свой цвет, будет не меньше $(2^{2^n-1}-2^{2^n-n-1}+1)$. По признаку Дирихле найдется игрок (пусть это первый игрок), который должен угадать в количестве ситуаций не менее
$$\left[\frac{2^{2^n-1}-2^{2^n-n-1}+1}{2^n-1}\right]+1=2^{2^n-n-1}+1.$$
Пусть это комбинации вида
$$xl_2^1\dots l_{2^n-1}^1,$$
$$\dots,$$
$$xl_2^{2^{2^n-n-1}+1}\dots l_{2^n-1}^{2^{2^n-n-1}+1},$$
в которых он угадывает свой цвет $\alpha^1_{l_2^i\dots l_{2^n-1}^i}$, тогда на комбинациях
$$\overline{\alpha^1_{l_2^1\dots l_{2^n-1}^1}}l_2^1\dots l_{2^n-1}^1,$$
$$\dots,$$
$$\overline{\alpha^1_{l_2^{2^{2^n-n-1}+1}\dots l_{2^n-1}^{2^{2^n-n-1}+1}}}l_2^{2^{2^n-n-1}+1}\dots l_{2^n-1}^{2^{2^n-n-1}+1}$$
первый игрок ошибется.

Отсюда число выигрышных комбинаций первого игрока не превышает $2^{2^n-1}-(2^{2^n-n-1}+1)$, что противоречит предположению о $(2^{2^n-1}-2^{2^n-n-1}+1)$ выигрышных ситуациях.

□

Задача 151. Имеются две урны, содержащие по 20 шаров – 10 белых, 5 черных и 5 красных в первой и 8 белых, 8 черных и 4 красных во второй. Из каждой урны вытаскивают по одному шару. Исход какого из этих двух опытов следует считать более неопределенным?

Решение. Набор вероятностей $\{p_i\}$ для первого опыта имеет вид $\left\{\frac{1}{2},\frac{1}{4},\frac{1}{4}\right\}$, для второго опыта — $\left\{\frac{2}{5},\frac{2}{5},\frac{1}{5}\right\}$. Чтобы определить, какой из опытов более неопределенный, вычислим меру неопределенности $H=\sum\limits_i p_i\ln\frac{1}{p_i}$. Имеем
$$H_1=\frac{1}{2}\ln 2+2\cdot\frac{1}{4}\ln 4=\frac{3}{2}\ln 2,$$
$$H_2=\frac{1}{5}\ln 5+2\cdot\frac{2}{5}\ln\frac{5}{2}=\ln 5-\frac{4}{5}\ln 2.$$
Вычисления дают $H_1<H_2$. Значит, исход второго из опытов более неопределенный. □

Задача 152 (о джентльменах и шляпах)**.** Три джентльмена заходят в комнату. Имеется три красных шляпы и две белых. Внезапно выключается свет, и в темноте каждый из джентльменов одевает одну из шляп, цвета которой он не видит. После включения света происходит следующий диалог:

Первый: я не знаю, какая на мне шляпа.

Второй: я тоже не знаю, какая на мне шляпа.

Третий: а я теперь знаю, какая на мне шляпа.

Какая шляпа на третьем джентльмене?

Решение. Обозначим цвета шляп джентльменов через c_i, $i = 1, 2, 3$, имеющих значения из множества $\{к, б\}$.

Первый джентльмен может понять, какая на нем шляпа, только в случае $c_2 = c_3 = б$ (тогда $c_1 = к$). Поэтому $\{c_2, c_3\}$ может иметь вид $\{к, б\}$, $\{к, к\}$ или $\{б, к\}$. Второй джентльмен может знать свой цвет лишь в случае $c_3 = б$ (тогда $c_2 = к$). Т.к. он своего цвета не знает, заключаем $c_3 = к$.

Таким образом, на третьем джентльмене красная шляпа. □

Задача 153. Пусть для некоторого пункта (скажем, для г. Долгопрудный) вероятность того, что 15 июня будет дождь, равна 0.4, а вероятность того, что дождя не будет, равна 0.6. Пусть далее для этого же пункта вероятность дождя 15 октября равна 0.8, а вероятность отсутствия осадков равна 0.2. Предположим, что определенный метод прогноза погоды 15 июня оказывается верным в 3/5 всех тех случаев, когда предсказывается дождь и в 4/5 тех случаев, в которых прогнозируется отсутствие осадков. Применительно к погоде на 15 октября этот метод оказывается правильным в 9/10 тех случаев, когда предсказывается дождь, и в половине случаев, когда предсказывается его отсутствие. В какой из указанных двух дней прогноз дает нам больше информации о реальной погоде?

Решение. Определим с.в. $X \in \mathrm{Be}(0.4)$ — индикатор события того, что 15 июня будет дождь. Аналогично вводим для 15 октября с.в. $Y \in \mathrm{Be}(0.8)$.

Результаты прогноза погоды определенным методом на 15 июня и 15 октября обозначим через Z_1 и Z_2, которые также являются бернуллиевскими с.в. Из условия задачи получим

$$\mathbb{P}(Z_1 = 1|X = 1) = \frac{3}{5}, \quad \mathbb{P}(Z_1 = 0|X = 0) = \frac{4}{5},$$

$$\mathbb{P}(Z_2 = 1|X = 1) = \frac{9}{10}, \quad \mathbb{P}(Z_2 = 0|X = 0) = \frac{1}{2}.$$

Введем с.в. $T_1 = \mathcal{I}_{\{Z_1 = X\}}$, $T_2 = \mathcal{I}_{\{Z_2 = Y\}}$, являющиеся индикаторами того, что прогноз погоды на 15 июня и 15 октября соответственно верен. Находим

$$\mathbb{P}(T_1 = 1) = \mathbb{P}(Z_1 = 1, X = 1) + \mathbb{P}(Z_1 = 0, X = 0) =$$

$$= \mathbb{P}(Z_1 = 1|X = 1)\mathbb{P}(X = 1) + \mathbb{P}(Z_1 = 0|X = 0)\mathbb{P}(X = 0) = 0.72,$$

$$\mathbb{P}(T_2 = 1) = \mathbb{P}(Z_2 = 1, Y = 1) + \mathbb{P}(Z_2 = 0, Y = 0) =$$

$$= \mathbb{P}(Z_2 = 1|Y = 1)\mathbb{P}(Y = 1) + \mathbb{P}(Z_2 = 0|Y = 0)\mathbb{P}(Y = 0) = 0.82.$$

Энтропии с.в. T_1 и T_2 являются мерой неопределенности того, что прогноз для соответствующей даты правильный. Поэтому тот прогноз дает больше информации, у которого энтропия соответствующей с.в. T_i меньше. Вычисления показывают

$$H_1 = 0.72 \ln \frac{1}{0.72} + 0.28 \ln \frac{1}{0.28} > H_2 = 0.82 \ln \frac{1}{0.82} + 0.18 \ln \frac{1}{0.18}.$$

Значит, прогноз на 15 октября дает больше информации о реальной погоде. $\square$

9. Раздел 18. Безгранично делимые распределения

Безгранично делимые распределения (процесс Леви, представление Леви-Хинчина)

Задача 154. Пусть с.в. X имеет экспоненциальное распределение с параметром λ ($X \in \mathrm{Exp}(\lambda)$), т.е. $\mathbb{P}(X > x) = e^{-\lambda x}$, $x \geqslant 0$. Покажите, что имеет место «отсутствие последействия»:

$$\mathbb{P}(X > x + y \,|\, X > x) = \mathbb{P}(X > y).$$

Решение. По определению условной вероятности при $x, y \geqslant 0$

$$\mathbb{P}(X > x + y \,|\, X > x) = \frac{\mathbb{P}(X > x + y,\, X > y)}{\mathbb{P}(X > x)} = \frac{\mathbb{P}(X > x + y)}{\mathbb{P}(X > x)} =$$

$$= \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} = e^{-\lambda y} = \mathbb{P}(X > y).$$

$\square$

Задача 155. Представьте, что вы владелец киоска по продаже мороженного (одного вида, скажем, пломбира) и хотите оценить, сколько мороженного $K(T)$ вам удастся продать за рабочий день T. Имеет место формула

$$K(T) = \max\Big\{n : \sum_{k=1}^{n} X_k < T\Big\},$$

где $X_1, X_2, X_3, \ldots$ — независимые одинаково распределенные по закону $\mathrm{Exp}(\lambda)$ с.в. (X_k интерпретируется как время между $k-1$ и k сделкой (продажей)). Покажите, что

1) вероятность $\mathbb{P}(K(T+t) - K(T) = k)$, где $t \geqslant 0$ и $k = 0, 1, 2, \ldots$, не зависит от $T \geqslant 0$;

2) $\forall n \geqslant 1,\ 0 \leqslant t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n \quad \to \quad \big\{K(t_k) - K(t_{k-1})\big\}_{k=1}^{n}$ — независимые с.в.;

3) $\mathbb{P}(K(t) > 1) = o(t)$, $t > 0$;

4) $\forall n \geqslant 1,\ 0 \leqslant t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n,\ 0 \leqslant k_1 \leqslant k_2 \leqslant \ldots \leqslant k_n,\ k_1, \ldots, k_n \in \mathbb{N} \cup \{0\}$

$$\mathbb{P}(K(t_1) = k_1, \ldots, K(t_n) = k_n) = e^{-\lambda t_1}\frac{(\lambda t_1)^{k_1}}{k_1!} \cdot e^{-\lambda(t_2 - t_1)}\frac{(\lambda(t_2 - t_1))^{k_2 - k_1}}{(k_2 - k_1)!} \cdot \ldots \cdot e^{-\lambda(t_n - t_{n-1})}\frac{(\lambda(t_n - t_{n-1}))^{k_n - k_{n-1}}}{(k_n - k_{n-1})!},$$

в частности, $K(T) \in \mathrm{Po}(\lambda T)$.

Решение. 1) По определению с вероятностью 1 выполнены соотношения

$$\sum_{k=1}^{K(T)} X_k < T, \qquad \sum_{k=1}^{K(T)+1} X_k \geqslant T.$$

Введем $\Delta x = T - \sum\limits_{k=1}^{K(T)} X_k$. Распределение с.в. $\sum\limits_{k=1}^{K(T)+1} X_k$ имеет вид

$$\mathbb{P}\Big(\sum_{k=1}^{K(T)+1} X_k \geqslant T + x\Big) = \mathbb{P}(X_{K(T)+1} \geqslant x + \Delta x \,|\, X_{K(T)+1} \geqslant \Delta x) = \mathbb{P}(X_{K(T)+1} \geqslant x),$$

исходя из доказанного ранее свойства «отсутствия последействия» для $X_k \in \mathrm{Exp}(\lambda)$. Поэтому одинаково распределены с.в.

$$\sum_{k=1}^{K(T)+s} X_k - T \quad \text{и} \quad \sum_{k=1}^{s} X_k,$$

$$\text{а также } \max\Big\{n : \sum_{k=1}^{K(T)+n} X_k < T + t\Big\} \quad \text{и} \quad \max\Big\{n : \sum_{k=1}^{n} X_k < t\Big\} = K(t).$$

Т.к. $\max\Big\{n : \sum\limits_{k=1}^{K(T)+n} X_k < T + t\Big\} = K(T+t) - K(T)$, одинаково распределены с.в. $K(T+t) - K(T)$ и $K(t)$.

Следовательно, вероятности $\mathbb{P}(K(T+t) - K(T) = k)$ равны величинам $\mathbb{P}(K(t) = k)$, не зависящим от T.

2) Предварительно найдем распределение с.в. $K(t)$. По индукции доказывается, что сумма $S_n = \sum\limits_{k=1}^{n} X_k$ имеет функцию распределения

$$F_{S_n}(t) = \mathbb{P}(S_n < t) = 1 - e^{-\lambda t}\left(1 + \frac{\lambda \cdot t}{1!} + \ldots + \frac{(\lambda \cdot t)^{n-1}}{(n-1)!}\right).$$

Тогда

$$\mathbb{P}(K(t) = k) = \mathbb{P}(S_k < t, S_{k+1} \geqslant t) = F_{S_k}(t) - F_{S_{k+1}}(t) = e^{-\lambda t}\frac{(\lambda \cdot t)^k}{k!}.$$

Значит, $K(t) \in \mathrm{Po}(\lambda t)$.

Характеристическая функция с.в. $\xi \in \mathrm{Po}(\lambda)$ имеет вид

$$\varphi(t) = \mathbb{E}e^{it\xi} = e^{\lambda(e^{it}-1)}.$$

Из этой формулы имеем

$$\varphi_{K(t_n)}(t) = \varphi_{K(t_1)}(t) \cdot \varphi_{K(t_2)-K(t_1)}(t) \cdot \ldots \cdot \varphi_{K(t_n)-K(t_{n-1})}(t).$$

Из свойств характеристических функций следует, что с.в. $\big\{K(t_k) - K(t_{k-1})\big\}_{k=1}^{n}$ независимы.

3)

$$\mathbb{P}(K(t)>1)=1-\mathbb{P}(K(t)\leqslant 1)=1-e^{-\lambda t}(1+\lambda t)=1-(1-\lambda t+o(t))\cdot(1+\lambda t)=o(t).$$

4) Из независимости с.в. $\big\{K(t_k)-K(t_{k-1})\big\}_{k=1}^{n}$ имеем

$$\mathbb{P}(K(t_1)=k_1,\ldots,K(t_n)=k_n)=$$
$$=\mathbb{P}(K(t_1)=k_1,K(t_2)-K(t_1)=k_2-k_1\ldots,K(t_n)-K(t_{n-1})=k_n-k_{n-1})=$$
$$=e^{-\lambda t_1}\frac{(\lambda t_1)^{k_1}}{k_1!}\cdot e^{-\lambda(t_2-t_1)}\frac{(\lambda(t_2-t_1))^{k_2-k_1}}{(k_2-k_1)!}\cdot\ldots\cdot e^{-\lambda(t_n-t_{n-1})}\frac{(\lambda(t_n-t_{n-1}))^{k_n-k_{n-1}}}{(k_n-k_{n-1})!}.$$

□

Задача 156. В течение рабочего дня фирма осуществляет $K(T)\in \mathrm{Po}(\lambda T)$ сделок ($K(T)$ — с.в., имеющая распределение Пуассона с параметром $\lambda T=100$[сделок/час] $*$ 10[часов]). Каждая сделка приносит доход $V_n\in R[a,b]$ (V_n — с.в., имеющая равномерное распределение на отрезке $[a,b]=[\$10,\$100]$, n — номер сделки). Считая, что K, V_1, V_2, … — независимые в совокупности с.в., найдите математическое ожидание и дисперсию выручки за день $Q(T)=\sum\limits_{k=1}^{K(T)}V_k$.

Решение. Используем формулы для условных вероятностей:

$$\mathbb{E}X=\mathbb{E}(X|Y),$$
$$\mathrm{Var}(\mathbb{E}(X|Y))+\mathbb{E}(\mathrm{Var}(X|Y))=\mathbb{E}(\mathbb{E}(X|Y))^2-(\mathbb{E}(\mathbb{E}(X|Y)))^2+\mathbb{E}\big(\mathbb{E}(X^2|Y)-(\mathbb{E}(X|Y))^2\big)=$$
$$=\mathbb{E}(\mathbb{E}(X|Y))^2-(\mathbb{E}X)^2+\mathbb{E}X^2-\mathbb{E}(\mathbb{E}(X|Y))^2=\mathbb{E}X^2-(\mathbb{E}X)^2=\mathrm{Var}\,X,$$

т.е.

$$\mathrm{Var}\,X=\mathrm{Var}(\mathbb{E}(X|Y))+\mathbb{E}(\mathrm{Var}(X|Y)).$$

Имеем

$$\mathbb{E}Q(T)=\mathbb{E}(\mathbb{E}(Q(T)|K(T)))=\mathbb{E}(K(T)\mathbb{E}(V_k))=\mathbb{E}(K(T))\mathbb{E}(V_k).$$

Т.к. $\mathbb{E}(V_k)=\frac{a+b}{2}$, $\mathbb{E}(K(T))=\lambda T$, математическое ожидание выручки за день равно

$$\mathbb{E}Q(T)=\lambda T\cdot\frac{a+b}{2}=55000\$.$$

Соответственно для дисперсии

$$\mathrm{Var}\,Q(T)=\mathrm{Var}(\mathbb{E}(Q(T)|K(T)))+\mathbb{E}(\mathrm{Var}(Q(T)|K(T)))=\mathrm{Var}(K(T)\mathbb{E}(V_k))+\mathbb{E}(K(T)\,\mathrm{Var}(V_k))=$$
$$=\mathrm{Var}(K(T))(\mathbb{E}(V_k))^2+\mathbb{E}(K(T))\,\mathrm{Var}(V_k).$$

Т.к. $\mathrm{Var}(V_k)=\frac{(b-a)^2}{12}$, $\mathrm{Var}(K(T))=\lambda T$, дисперсия выручки за день равна

$$\mathrm{Var}\,Q(T)=\lambda T\cdot\frac{(a+b)^2}{4}+\lambda T\cdot\frac{(b-a)^2}{12}=\lambda T\cdot\frac{a^2+ab+b^2}{3}=3700000\$^2.$$

□

Задача 157. В течение трех лет фирма из предыдущей задачи работала $N = 1000$ дней (длина рабочего дня и параметры спроса не менялись). Оцените распределение с.в.

$$Q^N = \sum_{k=1}^{N} Q_k(T),$$

где $Q_k(T)$ — выручка за k–ый день. Верно ли, что с.в. Q^N и $Q_k(NT)$ одинаково распределены?

Решение. 1. Согласно ц.п.т. получим, что распределение величины $\frac{Q^N - \mathbb{E}Q^N}{\sqrt{\operatorname{Var} Q^N}}$ близко к $N(0,1)$, поэтому распределение Q^N близко к $N(\mathbb{E}Q^N, \operatorname{Var} Q^N)$. Поскольку

$$\mathbb{E}Q^N = N \cdot \mathbb{E}(Q(T)) = 55 \cdot 10^6\$,$$
$$\operatorname{Var} Q^N = N \cdot \operatorname{Var}(Q(T)) = 37 \cdot 10^8\$^2,$$

распределение Q^N близко к нормальному $N(55 \cdot 10^6\$, 37 \cdot 10^8\$^2)$.

2. Близость величины $\frac{Q^N - \mathbb{E}Q^N}{\sqrt{\operatorname{Var} Q^N}}$ к нормальному $N(0,1)$ в смысле близости их функций распределения определяется из неравенства Берри-Эссена

$$\sup_x \left| \mathbb{P}\left(\frac{Q^N - \mathbb{E}Q^N}{\sqrt{\operatorname{Var} Q^N}} < x \right) - \Phi(x) \right| \leqslant \frac{C_0 \mu^3}{\sigma^3 \sqrt{N}},$$

где $C_0 < 0.7056$, $\mu^3 = \mathbb{E}\left| Q_k(T) - \mathbb{E}Q_k(T) \right|^3$, $\sigma^2 = \operatorname{Var} Q_k(T)$, $\Phi(x) = \int\limits_{-\infty}^{x} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\, dt$.

3. Для сравнения распределений с.в. Q^N и $Q_k(NT)$ найдем их характеристические функции.

$$\varphi_{Q^N}(t) = \varphi^N_{Q_k(T)}(t) = \exp\{N\lambda T(\varphi_{V_k}(t) - 1)\},$$
$$\varphi_{Q_k(NT)}(t) = \exp\{N\lambda T(\varphi_{V_k}(t) - 1)\} = \varphi_{Q^N}(t).$$

Из равенства характеристических функций рассматриваемых с.в. по теореме единственности следует равенство их распределений.

□

Задача 158 (пуассоновский поток событий)**.** Рассмотрим интервал $[-N, N]$ и бросим на него независимо и случайно (точнее равномерно) $M = [\rho N]$ точек, где $\rho > 0$ - некоторая константа, называемая плотностью. Легко вычислить биномиальную вероятность $P_{N,M}(k, I)$ того, что в конечный интервал $I \subset [-N, N]$ попадет ровно k точек. Покажите, что $P_{N,M}(k, I)$ стремится при $N \to \infty$ к пуассоновскому выражению:

$$P(k, I) \stackrel{def}{=} \lim_{N \to \infty} P_{N,M(N)}(k, I) = \frac{(\rho\, |I|)^k}{k!} e^{-\rho |I|}, \quad k = 0, 1, \ldots$$

Покажите также, что если $I_1, I_2 \subset [-N, N]$ и $I_1 \cap I_2 = \emptyset$, то

$$P(k_1, I_1; k_2, I_2) \stackrel{def}{=} \lim_{N \to \infty} P_{N,M(N)}(k_1, I_1; k_2, I_2) = P(k, I_1)\, P(k, I_2).$$

Решение. To be done □

Задача 159 (процесс Леви)**.** Покажите, что если $X(t)$ — процесс Леви, то:

$$\exists!(b, c, \nu(dx)) : \; c \geqslant 0, \; \int_{-\infty}^{\infty} \max(1, x^2)\, \nu(dx) < \infty, \; \nu(dx) \geqslant 0 :$$

$$\forall t \geqslant 0 \;\to\; \varphi_{X(t)}(\mu) = \mathbb{E}e^{i\mu X(t)} = \exp\Big\{t\Big[i\mu b - c\mu^2 \,/\, 2 + \int_{-\infty}^{\infty} \big(e^{i\mu x} - 1 - i\mu x I(|x| \leqslant 1)\big)\, \nu(dx)\Big]\Big\}. \tag{9.1}$$

Решение. Из представления

$$X(1) = [X(t_n) - X(t_{n-1})] + \ldots + [X(t_2) - X(t_1)] + X(t_1),$$

где $t_k = \frac{k}{n}$, и определения процесса Леви следует, что $X(1)$ — безгранично делимая с.в. Выпишем для $X(1)$ представление Леви–Хинчина:

$$\varphi_{X(1)}(\mu) = \mathbb{E}e^{i\mu X(1)} = \exp\Big\{i\mu b - c\mu^2 \,/\, 2 + \int_{-\infty}^{\infty} \big(e^{i\mu x} - 1 - i\mu x I(|x| \leqslant 1)\big)\, \nu(dx)\Big\}.$$

Из свойств характеристических функций для суммы независимых с.в. вытекает справедливость представления (9.1) для любых рациональных t.

Поскольку процесс $X(t)$ стохастически непрерывен, имеет место сходимость по распределению

$$X(t_n) \xrightarrow{n\to\infty} X(t), \qquad \text{где } t_n \xrightarrow{n\to\infty} t, \; t_n \in \mathbb{Q}.$$

По теореме непрерывности для характеристических функций отсюда следует

$$\forall \mu : \; \varphi_{X(t_n)}(\mu) \xrightarrow{n\to\infty} \varphi_{X(t)}(\mu),$$

поэтому $\forall t \geqslant 0$ справедливо представление (9.1).

□

Задача 160. **1) Биномиальная однопериодная модель Кокса – Росса - Рубинштейна.** Пусть на "идеализированном" фондовом рынке имеется всего две ценные бумаги, и торговля осуществляется всего в два момента времени. Пусть цена первой бумаги S (будем называть её акцией (stock)) известна в первый момент. Цена второй бумаги C (будем называть ее call – опционом европейского типа[2]) не известна в первый момент. Пусть с ненулевой вероятностью $p > 0$ к моменту времени 2 цена акции

[2]Опцион характеризуется датой исполнения (в нашем случае - момент времени 2) и платежами в момент исполнения (C_u и C_d). Причем эти платежи - заранее известные функции от цены акции в этот момент (введение опционов было мотивированно, желанием "хеджироваться", страховаться от нежелательных изменений цен акций). Основная задача заключается в установлении "справедливой" цены опциона C в момент времени 1 (см. А.Н. Ширяев, Вероятность-2, М.: МЦНМО, 2004, глава 7, § 11; А.Н. Ширяев, Основы финансовой стохастической математики, М.: ФАЗИС, 2004, т. 1, т. 2).

вырастет в $u > 1$ (up) раз и с вероятностью $1 - p$ цена акции "вырастет" в $d < 1$ (down) раз, т.е. падает. Пусть также известны возможные цены опциона во второй момент: C_u - если акция выросла в цене и C_d - если акция упала в цене. Для простоты будем считать, что банк работает с нулевым процентом, т.е. класть деньги в банк, в расчете на проценты, бессмысленно. Говорят, что ***рынок безарбитражный***, если не существует таких k_S, k_C, что[3]

$X(1) = k_S S + k_C C = 0, \quad P(X(2) \ge 0) = P(k_S S(2) + k_C C(2) \ge 0) = 1$, причем $P(X(2) > 0) > 0$.

Докажите, что рассматриваемый рынок безарбитражный тогда и только тогда, когда[4]

$C = \tilde{p} C_u + (1 - \tilde{p}) C_d$, где $\tilde{p} = \frac{1-d}{u-d}$.

2) **Биномиальная n-периодная модель Кокса – Росса – Рубинштейна.** Предложите обобщение рынка и соответствующих понятий п. 1 на n - периодный рынок. С возможностью класть деньги в банк под процент $r - 1$ $(d < r < u)$ - за один период (под такой же процент брать деньги из банка). Опцион исполняется в заключительный $n + 1$-ый момент. Платежи по опциону в этот момент известны - описываются известной функцией $\bar{C}(S)$ (например, для указанного в п. 1 опциона[5] $\bar{C}(S) = \max\ \{0,\ S - X\}$), т.е. $C_k(n+1) = \bar{C}(S_k(n+1))$, где k - состояние в котором находится рынок в момент времени $n + 1$. Считайте, что $S_k(n+1) = Su^k d^{n-k}$, т.е. k - характеризует то, сколько раз акция поднималась в цене. Также как и в п. 1, требуется определить "справедливую" цену опциона.

   Указание. Обоснуйте формулу Кокса – Росса – Рубинштейна:

   $C = \frac{1}{r^n} \sum\limits_{k=0}^{n} \binom{n}{k} \tilde{p}^k (1 - \tilde{p})^{n-k} \bar{C}\left(Su^k d^{n-k}\right)$, где $\tilde{p} = \frac{r-d}{u-d}$. (i)

3) **Континуальная биномиальная модель Блэка – Шоулса.** Уместим на отрезке времени $[0,\ t]$ $\quad n+1$ - моментов (промежутки между которыми одинаковы), в которые осуществляется торговля согласно п. 2. Введем два параметра: a - **снос**, $\sigma^2 \ge 0$ - **волатильность** (дисперсия). Положим,

---

[3] Не имея в начальный момент 1 капитала $X(1) = 0$, но, проделав некоторую махинацию (продав одних ценных бумаг (в зависимости от специфики рынка, иногда разрешается "вставать в короткую позицию" – продавать ценные бумаги, не имея их в наличие; приобретая при этом долг) и купив на вырученные деньги других бумаг), можно в момент времени 2 гарантированно ничего не проиграть, и при этом с ненулевой вероятностью выиграть (не уточняя сколько – поскольку, "прокручивая" по имеющемуся арбитражу (пропорционально увеличивая коэффициенты k_S, k_C) сколь угодно большую сумму, можно получить с ненулевой вероятностью сколь угодно большой выигрыш).

[4] $\tilde{p}$ - называется мартингальной вероятностью (смысл такого определения будет раскрыт позже) и задает ***мартингальную меру***. Если существует единственная мартингальная мера, то рынок называется ***полным***. На полном рынке неизвестная цена опциона C в начальный момент определяется однозначно и может интерпретироваться как "справедливая цена".

[5] X - называется ценной исполнения опциона и считается известной. Собственно, вид функции $\bar{C}(S) = \max\ \{0,\ S - X\}$ проясняет смысл опциона. Опцион дает право купить (у того, кто продал нам опцион) в момент исполнения опциона акцию по цене X. Если акция стоит дороже в этот момент, то, конечно, мы этим правом воспользуемся и получим прибыль (продавец опциона обязан продать нам акцию). Если же цена акции меньше цены исполнения опциона, то нам уже не выгодно покупать акцию по более дорогой цене, чем рыночная, и мы не исполняем опцион, т.е. ничего не делаем (ведь опцион дает нам право, ни к чему не обязывая).

$\mu = a + \frac{\sigma^2}{2}, \quad r = \exp\left(\mu \frac{t}{n}\right), \quad u = \exp\left(\sigma\sqrt{\frac{t}{n}}\right), \quad d = \exp\left(-\sigma\sqrt{\frac{t}{n}}\right), \quad \tilde{p} = \frac{r-d}{u-d} \approx \frac{1}{2}\left(1 + \frac{a}{\sigma}\sqrt{\frac{t}{n}}\right)$
(ii)

Переходя к пределу при $n \to \infty$ в формуле (i) (с $\bar{C}(S) = \max\ \{0,\ S - X\}$), согласно (ii), получите формулу Блэка – Шоулса для справедливой цены опциона в "континуальной биномиальной модели". Почему вводится именно два параметра (а не один, три и т.д.)? Почему

$$r - 1 \sim \frac{\mu t}{n}, \quad u - 1 \approx \sigma\sqrt{\frac{t}{n}}, \quad d - 1 \approx -\sigma\sqrt{\frac{t}{n}}?$$

Возможны ли какие-нибудь другие осмысленные варианты соотношений типа (ii), при которых будет существовать предел при $n \to \infty$ в формуле (i) (для простоты вычислений считайте, что $\bar{C}(S) := S$)?

4) **Броуновское движение (процесс Башелье) и винеровский процесс.** Исходя из формулы (i), имеем, что "рынок" при определении "справедливой" цены опциона считает, что случайный процесс $S(m)$ (цена акции в момент времени m) - эволюционирует согласно биномиальной модели с неизменными параметрами[6] d, u, $\tilde{p}$. Построим случайный процесс $S(t)$ (в непрерывном времени), исходя из процесса $S(m)$, заданного в дискретном времени предельным переходом, аналогичным п. 3. Назовем, полученный процесс $S(t)$ - ***геометрическим броуновским движением*** (или случайным процессом Башелье - Самуэльсона) с параметрами[7] a, $\sigma^2 \geq 0$, а случайный процесс $B(t) = \ln\frac{S(t)}{S(0)}$ - ***броуновским движением*** с параметрами a, $\sigma^2 \geq 0$. Если $a = 0$, $\sigma^2 = 1$, то такое броуновское движение имеет специальное название – ***винеровский процесс*** $W(t)$. Покажите, что броуновское движение является процессом Леви. Найдите триплет $(b,\ c,\ \nu(dx))$.

Решение. To be done □

Задача 161 (пуассоновский процесс)**.** В приложениях (в теории надежности, теории массового обслуживания) широко используются ***процессы восстановления***(***потоки Пальма***). Основным (и наиболее удобным для анализа) представителем таких процессов является ***пуассоновский процесс***, который можно определить следующим образом:[8]

[6]Кстати говоря, один из альтернативных способов введения мартингальных вероятностей $\tilde{p}$ основывается на, так называемых, "риск нейтральных" или "мартингальных" соображениях. Заключающихся в том, что $\tilde{p}$ выбирается исходя из равенства $M_{\tilde{p}}\left[S u^{\sum_{k=1}^{n} x_k} d^{n - \sum_{k=1}^{n} x_k}\right] = S r^n$, где i.i.d.$x_k \in Be(\tilde{p})$ или исходя из того, что процесс приведенной (продисконтированной) стоимости акции $\tilde{S}(m) = S(m) \big/ r^m$ - должен быть ***мартингалом*** относительно мартингальной меры $\tilde{p}$ (отсюда и название), т.е. $M_{\tilde{p}}\left(\tilde{S}(m+1) \middle| \left(\tilde{S}(1), ..., \tilde{S}(m)\right)\right) = \tilde{S}(m)$ (см. А.В. Летчиков, Лекции по финансовой математике, Москва - Ижевск, РХД, 2004, глава 3).

[7]Эти параметры имеют следующий смысл: $a = \frac{1}{t} M\left[\ln\frac{S(t)}{S(0)}\right]$, $\sigma^2 = \frac{1}{t} D\left[\ln\frac{S(t)}{S(0)}\right]$.

[8]$K(t)$ - число отказов приборов к моменту времени $t \geq 0$ (отказавший прибор сразу же заменяется исправным). Все приборы идентичны, т.е. имеют одинаковое распределение времени безотказной работы. Кроме того, приборы работают независимо друг от друга.

$K(t) = \max\left\{k:\ \sum_{i=1}^{k} T_i < t\right\}$, где[9] i.i.d. с.в. $T_i \in \mathrm{Exp}(\lambda)$, т.е.[10] $P(T_i > t) = e^{-\lambda t}$, $\lambda > 0$. Покажите, что $K(t)$ является процессом Леви. Найдите триплет $(b,\ c,\ \nu(dx))$.

Решение. 1) По определению с вероятностью 1 выполнены соотношения

$$\sum_{k=1}^{K(T)} X_k < T, \qquad \sum_{k=1}^{K(T)+1} X_k \geqslant T.$$

Введем $\Delta x = T - \sum_{k=1}^{K(T)} X_k$. Распределение с.в. $\sum_{k=1}^{K(T)+1} X_k$ имеет вид

$$\mathbb{P}\Big(\sum_{k=1}^{K(T)+1} X_k \geqslant T + x\Big) = \mathbb{P}(X_{K(T)+1} \geqslant x + \Delta x \,|\, X_{K(T)+1} \geqslant \Delta x) = \mathbb{P}(X_{K(T)+1} \geqslant x),$$

исходя из доказанного ранее свойства «отсутствия последействия» для $X_k \in \mathrm{Exp}(\lambda)$. Поэтому одинаково распределены с.в.

$$\sum_{k=1}^{K(T)+s} X_k - T \quad \text{и} \quad \sum_{k=1}^{s} X_k,$$

$$\text{а также } \max\Big\{n:\ \sum_{k=1}^{K(T)+n} X_k < T + t\Big\} \quad \text{и} \quad \max\Big\{n:\ \sum_{k=1}^{n} X_k < t\Big\} = K(t).$$

Т.к. $\max\Big\{n:\ \sum_{k=1}^{K(T)+n} X_k < T + t\Big\} = K(T+t) - K(T)$, одинаково распределены с.в. $K(T+t) - K(T)$ и $K(t)$.

Следовательно, вероятности $\mathbb{P}(K(T+t) - K(T) = k)$ равны величинам $\mathbb{P}(K(t) = k)$, не зависящим от T.

2) Предварительно найдем распределение с.в. $K(t)$. По индукции доказывается, что сумма $S_n = \sum_{k=1}^{n} X_k$ имеет функцию распределения

$$F_{S_n}(t) = \mathbb{P}(S_n < t) = 1 - e^{-\lambda t}\left(1 + \frac{\lambda \cdot t}{1!} + \ldots + \frac{(\lambda \cdot t)^{n-1}}{(n-1)!}\right).$$

Тогда

$$\mathbb{P}(K(t) = k) = \mathbb{P}(S_k < t, S_{k+1} \geqslant t) = F_{S_k}(t) - F_{S_{k+1}}(t) = e^{-\lambda t}\frac{(\lambda \cdot t)^k}{k!}.$$

[9]Параметр $\lambda > 0$ принято называть интенсивностью пуассоновского процесса.

[10]Напомним важную особенность показательного распределения "отсутствие последействия": $P(T_i > t + \tau | T_i > t) = P(T_i > \tau)$. Заметим также, что общие процессы восстановления задаются аналогичной формулой с той лишь разницей, что i.i.d. $T_i \geq 0$ п.н. уже не обязательно распределены по показательному закону.

Значит, $K(t) \in \mathrm{Po}(\lambda t)$.

Характеристическая функция с.в. $\xi \in \mathrm{Po}(\lambda)$ имеет вид

$$\varphi(t) = \mathbb{E}e^{it\xi} = e^{\lambda(e^{it}-1)}.$$

Из этой формулы имеем

$$\varphi_{K(t_n)}(t) = \varphi_{K(t_1)}(t) \cdot \varphi_{K(t_2)-K(t_1)}(t) \cdot \ldots \cdot \varphi_{K(t_n)-K(t_{n-1})}(t).$$

Из свойств характеристических функций следует, что с.в. $\left\{K(t_k) - K(t_{k-1})\right\}_{k=1}^{n}$ независимы.

3) Таким образом, $K(t)$ является процессом Леви, т.к. $K(0) = 0$. Характеристическая функция

$$\varphi_{K(1)}(\mu) = \exp\left\{\lambda(e^{i\mu} - 1)\right\},$$

при этом из представления Леви-Хинчина

$$\varphi_{K(1)}(\mu) = \exp\left\{i\mu b - c\mu^2 \,/\, 2 + \int\limits_{-\infty}^{\infty} \left(e^{i\mu x} - 1 - i\mu x I(|x| \leqslant 1)\right) \nu(dx)\right\}.$$

Тогда триплет $b = \mathbb{E}K(1) = \lambda$, $c = 0$, $\nu(dx) = \lambda I_{\{1\}}(dx)$ удовлетворяет данному равенству, т.к.

$$\exp\left\{i\mu\lambda + \int\limits_{-\infty}^{\infty} \left(e^{i\mu x} - 1 - i\mu x I(|x| \leqslant 1)\right) \lambda I_{\{1\}}(dx)\right\} = \exp\left\{i\mu\lambda + \lambda(e^{i\mu} - 1 - i\mu)\right\} = \exp\left\{\lambda(e^{i\mu} - 1)\right\}.$$

Итак, $(b, c, \nu(dx)) = (\lambda, 0, \lambda I_{\{1\}}(dx))$.

□

Задача 162 (сложный пуассоновский процесс)**.** В микроэкономике, страховании и в ряде других приложений часто возникает ***сложный (взвешенный) пуассоновский процес***, который можно определить следующим образом:[11] $Q\,(t) = \sum\limits_{i=1}^{K(t)} V_i$, где i.i.d. V_i:

$$dF_{V_i}\,(x) = \nu\,(dx)/\nu\,(\mathrm{R})$$

не зависят от пуассоновского процесса $K\,(t)$, интенсивность, которого равна $\lambda = \nu\,(\mathrm{R}) < \infty$. Покажите, что $Q\,(t)$ является процессом Леви. Найдите триплет $(b,\ c,\ \nu\,(dx))$. С помощью теоремы 8.10 (предельная теорема с управляющим параметром) из книги “*Натан А.А., Горбачев О.Г., Гуз С.А.* Теория вероятностей. М.: МЗ Пресс, 2007” исследуйте поведение $Q\,(t)$ при $t \to \infty$.

[11] $Q\,(t)$ может интерпретироваться, как прибыль фирмы к моменту времени t. При этом предполагается, что к этому моменту времени фирма осуществляет $K\,(t)$ сделок (т.е. число сделок, вообще говоря, случайно), причем каждая сделка, в независимости от остальных сделок (и их количества), приносит случайный доход V_i.

Решение. 1. Т.к. $K(0)=0$, получим $Q(0)=0$. Далее, исходя из того, что $K(t)$ является процессом Леви, имеет место равенство по распределению:

$$Q(t+s)-Q(s)=\sum_{i=K(s)+1}^{K(t+s)} V_i=\sum_{i=1}^{K(t+s)-K(s)} V_i=\sum_{i=1}^{K(t)} V_i=Q(t).$$

Также, $\forall t_{n-1}<t_n$, равны по распределению

$$Q(t_n)-Q(t_{n-1})=\sum_{i=K(t_{n-1})+1}^{K(t_n)} V_i=\sum_{i=1}^{K(t_n)-K(t_{n-1})} V_i,$$

поэтому $\forall t_1<\ldots<t_{n-1}<t_n$ независимы с.в. $Q(t_1)$, $Q(t_2)-Q(t_1)$, …, $Q(t_n)-Q(t_{n-1})$, что следует из независимости с.в. V_i, $K(t_1)$, $K(t_2)-K(t_1)$, …, $K(t_n)-K(t_{n-1})$.

2. Таким образом, $Q(t)$ является процессом Леви. Характеристическая функция

$$\varphi_{Q(1)}(\mu)=\exp\big\{\lambda(\varphi_{V_i}(\mu)-1)\big\}=\exp\Big\{\nu(\mathbb{R})\Big(\int\limits_{-\infty}^{+\infty} e^{i\mu x}\nu(dx)/\nu(\mathbb{R})-1\Big)\Big\}=$$

$$=\exp\Big\{\int\limits_{-\infty}^{+\infty}(e^{i\mu x}-1)\nu(dx)\Big\}.$$

При этом из представления Леви-Хинчина

$$\varphi_{Q(1)}(\mu)=\exp\Big\{i\mu b-c\mu^2\,/\,2+\int\limits_{-\infty}^{\infty}\big(e^{i\mu x}-1-i\mu x I(|x|\leqslant 1)\big)\,\nu(dx)\Big\}.$$

Отсюда получим

$$i\mu b-c\mu^2\,/\,2=i\mu\int\limits_{-\infty}^{\infty} xI(|x|\leqslant 1)\,\nu(dx).$$

Данное равенство имеет место при $c=0$, $b=\int\limits_{-1}^{1} x\,\nu(dx)=\lambda\mathbb{E}\big(V_i I_{\{V_i\leqslant 1\}}\big)$.

Итак, $(b,c,\nu(dx))=\Big(\int\limits_{-1}^{1} x\,\nu(dx),0,\nu(dx)\Big)=\Big(\lambda\mathbb{E}\big(V_i I_{\{V_i\leqslant 1\}}\big),0,\lambda\cdot dF_{V_i}(x)\Big)$.

$\square$

Задача 163. Пусть n единичных масс равномерно распределены на $[-n,n]$. На единичную массу в начале координат действует гравитационная сила

$$f_n=\sum_{k=1}^{n}\frac{\operatorname{sign}(X_k)}{X_k^2}.$$

В силу равномерного распределения X_k

$$\mathbb{E}\left[\exp\left\{it\frac{\operatorname{sign}(X_k)}{X_k^2}\right\}\right] = \int\limits_{-n}^{n} \exp\left\{it\frac{\operatorname{sign}(x)}{x^2}\right\}\frac{dx}{2n} = \frac{1}{n}\int\limits_{0}^{n}\cos\left(\frac{t}{x^2}\right)dx.$$

Покажите, что $\mathbb{E}\left[\exp\left\{itf_n\right\}\right] \underset{n\to\infty}{\longrightarrow} \exp\left(-c\sqrt{|t|}\right)$. Что отсюда следует? Заметим, что соответствующее предельное распределение в элементарных функциях не выражается.

Решение. Для доказательства $\mathbb{E}\left[\exp\left\{itf_n\right\}\right] \underset{n\to\infty}{\longrightarrow} \exp\left(-c\sqrt{|t|}\right)$ достаточно показать при $t \geq 0$, что

$$\frac{1}{n}\int\limits_{0}^{n}\cos\left(\frac{t}{x^2}\right)dx = 1 - \frac{c\sqrt{t}+\alpha_n}{n},$$

где $\alpha_n \xrightarrow{n\to\infty} 0$.

Действительно, в таком случае

$$\mathbb{E}\big[\exp\{itf_n\}\big] = \left(\frac{1}{n}\int_0^n \cos\!\left(\frac{t}{x^2}\right) dx\right)^n = \left(1 - \frac{c\sqrt{|t|}+\alpha_n}{n}\right)^n \xrightarrow{n\to\infty} \exp\!\big(-c\sqrt{|t|}\big).$$

Итак, получим

$$n - \big(c\sqrt{t}+\alpha_n\big) = \int_0^n \cos\frac{t}{x^2}\,dx$$

$$\Rightarrow \int_0^n \left(1 - \cos\frac{t}{x^2}\right) dx = c\sqrt{t} + \alpha_n.$$

Последнее соотношение равносильно тому, что $\int_0^\infty \left(1 - \cos\frac{t}{x^2}\right) dx = c\sqrt{t}$. Для его доказательства рассмотрим интеграл с параметром

$$F(t) = \int_0^\infty \left(1 - \cos\frac{t}{x^2}\right) dx.$$

Имеем $F(0) = 0$, $\exists F'(t)$, причем

$$F'(t) = \int_0^\infty \frac{1}{x^2}\cdot\sin\frac{t}{x^2}\,dx = -\int_0^\infty \frac{x}{2t}\cdot\sin\frac{t}{x^2}\,d\!\left(\frac{t}{x^2}\right).$$

По формуле интегрирования по частям находим

$$F'(t) = \int_0^\infty \frac{1}{2t}\left(1 - \cos\frac{t}{x^2}\right) dx = \frac{F(t)}{2t}.$$

Отсюда получим

$$\frac{dF}{F} = \frac{dt}{2t} \;\Rightarrow\; F(t) = c\sqrt{t}.$$

Утверждение о предельном переходе доказано.

Функция $\exp\big(-c\sqrt{|t|}\big)$ непрерывна в нуле, поэтому является характеристической функцией некоторой с.в. X (как предел характеристических функций). По теореме о предельном переходе для характеристических функций получим сходимость по распределению

$$f_n \xrightarrow{d} X \quad (n\to\infty)$$

□

10. Раздел *

Теория чисел и теория вероятностей

Задача 164 (вероятностное доказательство формулы Эйлера)**.** Пусть X – целочисленная случайная величина с распределением

$P\left(X=n\right)=\frac{1}{\varsigma(s)n^s}$, где $\varsigma(s)=\sum\limits_{n\geq 1} n^{-s}, \quad s>1.$

Пусть $1<p_1<p_2<p_3<\ldots$ - простые числа, и пусть A_k - событие = {X делится на p_k}.

А) Найдите $P\{A_k\}$ и покажите, что события $A_1, A_2, \ldots$независимы.

Б) Получите, что

$\prod\limits_{k=1}^{\infty}(1-p_k^{-s})=\frac{1}{\varsigma(s)}$ (формула Эйлера).

Решение. To be done □

11. Раздел **

Перестановки и их свойства

Задача 165 ("сто заключенных")**.** В коридоре находятся 100 человек, у каждого свой номер (от 1 до 100). Их по одному заводят в комнату, в которой находится комод со 100 выдвижными ящиками. В ящики случайным образом разложены карточки с номерами (от 1 до 100). Каждому разрешается заглянуть в не более чем 50 ящиков. Цель каждого – определить, в каком ящике находится его номер. Общаться и передавать друг другу информацию запрещается. Предложите стратегию, которая с вероятностью не меньшей 0.3 (в предположении, что все 100! способов распределения карточек по ящикам равновероятны) приведет к выигрышу всей команды. Команда выигрывает, если все 100 участников верно определили ящик с карточкой своего номера.

Стратегия. Каждый человек первым открывает ящик под его номером, вторым – под номером, который указан на карточке, лежащей в ящике, открытом перед этим и т.д. Среднее число циклов длины r в случайной перестановке – есть 1/r (покажите, используя, например, задачу "про предельные меры"). Тогда среднее число циклов длины большей n/2 есть $\sum\limits_{i=n/2}^{n}\frac{1}{i}$. Это и есть вероятность существования цикла длины большей n/2. Поэтому вероятность успеха команды – есть $1-\sum\limits_{i=51}^{100}\frac{1}{i}\approx 0,31$. Если же просто произвольно открывать ящики, то вероятность успеха будет $2^{-100}\approx 8\cdot 10^{-31}$. В случае, когда карточки разложены не случайным образом, то следует сделать случайной нумерацию ящиков, и далее следовать старой стратегии.

Решение. To be done □

Задача 166 (предельные меры; А.М. Вершик и др., 1977)**.** В качестве множества элементарных исходов рассматривается группа всевозможных подстановок (перестановок) S_n (симметрическая группа), $n \gg 1$. В этой группе $n!$ элементов. Припишем каждой подстановке одинаковую вероятность $1/n!$.

А)* Покажите, что математическое ожидание числа циклов есть $\approx \ln n$.

Б)** В каком смысле нормированные длины циклов случайной подстановки убывают со скоростью геометрической прогрессии со знаменателем e^{-1}.

В)** Положим $\rho_n(a) = |\{g \in S_n : n_{\max}(g) \leq an\}|/n!$, где $n_{\max}(g)$ - длина максимально цикла в подстановке g. Покажите, что $\rho_n(a)$ удовлетворяет *уравнению Дикмана – Гончарова* (40-ые годы XX века):

$$\rho_n(a) = \int_0^a \rho_n\left(\frac{a}{1-t}\right) dt.$$

Г)*** Покажите, что начиная с некоторого большого числа N 99% натуральных чисел n, больших, чем N обладают свойством

$$n^{0.99} < p_1 \cdot \ldots \cdot p_{11}.$$

Иначе говоря, у основной части (99%) натуральных чисел основная часть (99%) числа есть произведение наибольших простых делителей. Число 11 возникло из-за того, что мы выбрали 99% и 99%.

Решение. To be done □

12. Раздел Перколяция

Задача 167 (перколяция)**.** В квадратном пруду (со стороной равной 1) выросли (случайным образом) $N \gg 1$ цветков лотоса, имеющих форму круга радиуса $r > 0$. Назовем r_N - *радиусом перколяции*, если с вероятностью не меньшей 0.99 не любящий воду жук сможет переползти по цветкам лотоса с северного берега на южный, не замочившись.

Покажите, что $r_N \sim C\Big/\sqrt{N}$. Оцените C.

Решение. To be done □