# Вероятностные доказательства

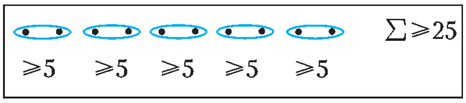
А.ШЕНЬ

## Неравенства и оценки

Начнем с совсем простой задачи.

**1.** *На доске написаны 10 чисел. Сумма любых двух из них не меньше 5. Докажите, что сумма всех чисел не меньше 25.*

▹ Тут даже и доказывать-то особо нечего: разобьем числа на пять пар. В каждой паре сумма не меньше 5, всего получается 25. ◃



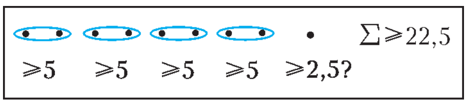

Теперь немного более сложная задача.

**2.** *На доске написаны 9 чисел. Сумма любых двух из них не меньше 5. Докажите, что сумма всех чисел не меньше 22,5.*

(Если все числа равны 2,5, то и получится 22,5.)

▹ Рассуждая аналогично предыдущей задаче, можно выбрать четыре пары, и одно число останется. В каждой паре сумма не меньше 5, всего 20. Если бы мы знали, что оставшееся без пары число не меньше 2,5, то мы получили бы требуемое. Но гарантировать это мы не можем. Скажем, если одно из чисел равно нулю, а остальные равны 5, то сумма любых двух будет не меньше 5 (а именно, 5 или 10). При этом непарным может оказаться ноль.

Наверное, вы уже сообразили, как завершить рассуждение. Разбиение на пары в наших руках, и мы можем оставить непарным любое из чисел. В любой паре хотя бы одно из чисел равно 2,5 или больше. Нам

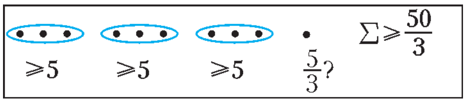



достаточно знать, что хотя бы одно из чисел таково – можно оставить непарным именно это число, и все получится. ◃

Усложним задачу еще немного.

**3.** *На доске написаны 10 чисел. Сумма любых трех из них не меньше 5. Докажите, что сумма всех чисел не меньше 50/3.*

▹ Используем тот же прием: в любой тройке чисел большее равно 5/3 или больше (иначе все три числа меньше 5/3, и сумма меньше 5). Возьмем одно из таких чисел (не меньших 5/3), останутся три тройки, в каждой из которых сумма не меньше 5. Всего получается как минимум $3\cdot 5+(5/3)=50/3$. ◃




В двух предыдущих задачах оставалось одно число. А что будет, если останется два?

**4.** *На доске написаны 11 чисел. Сумма любых трех из них не меньше 5. Докажите, что сумма всех чисел не меньше 55/3.*

Можно решить эту задачу аналогично предыдущим (см. задачу 14 и указание к ней). Но мы приведем другое решение, которое обобщается на любые количества.

▹ Обозначим наши числа $a_1,\ldots,a_{11}$ и запишем их по кругу. Сумма любой тройки соседей не меньше 5 (по условию):

$$a_1+a_2+a_3\ge 5,$$
$$a_2+a_3+a_4\ge 5,$$
$$\ldots$$
$$a_9+a_{10}+a_{11}\ge 5,$$
$$a_{10}+a_{11}+a_1\ge 5,$$
$$a_{11}+a_1+a_2\ge 5.$$

Сложим эти одиннадцать неравенств. Каждое число встречается три раза (входит в три тройки), поэтому получится

$$3\left(a_1+a_2+\ldots+a_{11}\right)\ge 55,$$

откуда и следует требуемое. ◃

## Вероятности и ожидания

Сейчас мы рассмотрим другое решение последней задачи – на языке теории вероятностей.

Пусть в ящике (или, как любят говорить в теории вероятностей, в урне) лежат 11 шаров, на которых написаны числа $a_1,a_2,\ldots,a_{11}$. Мы вынимаем один из шаров не глядя, и нам дают столько рублей, сколько на нем написано.

Сколько мы будем выигрывать в среднем за игру, если играть много раз? Каждый раз вынутый шар возвращают в урну, и ее содержимое перемешивают. Несложно сообразить, что ответ тут – среднее арифметическое всех чисел,

$$S = \frac{a_1 + a_2 + \ldots + a_{11}}{11}.$$

В самом деле, пусть мы играем большое количество раз. Поскольку все шары на ощупь одинаковые и их перемешивают, то они будут попадаться нам одинаково часто. В $1/11$ доле всех игр нам попадется шар $a_1$, в $1/11$ доле – шар $a_2$ и так далее. Всего мы выиграем (за $N$ игр)

$$\frac{N}{11}a_1 + \frac{N}{11}a_2 + \ldots + \frac{N}{11}a_{11},$$

и в среднем на одну игру придется как раз $S$. В теории вероятностей средний выигрыш за большое число игр называют *математическим ожиданием* выигрыша, так что в данной игре математическое ожидание выигрыша равно среднему арифметическому всех чисел.

А что будет, если разрешить команде из трех человек *А*, *Б* и *В* вынимать по шару? Мы имеем ввиду, что шары вынимают без возвращения, так что в урне остается 8 шаров. Каково будет математическое ожидание выигрыша команды (средний выигрыш в расчете на одну игру)?

Ясно, что по-прежнему каждый из игроков будет вынимать все шары одинаково часто (ведь тот факт, что кто-то до тебя не глядя вынул шар, не меняет равенства шансов). Значит, каждый из игроков будет в среднем выигрывать $S$, а вся команда – в среднем $3S$.

Теперь вспомним условие задачи: сумма любых трех чисел не меньше 5. Значит, в каждой игре команда выигрывает не меньше 5, и средний выигрыш тоже не меньше 5. Таким образом,

$$3S = 3 \times \frac{a_1 + a_2 + \ldots + a_{11}}{11} \geq 5,$$

поэтому $a_1 + a_2 + \ldots + a_{11} \geq 55/3$, что и требовалось доказать.

Это рассуждение выглядит нестрогим: что значит «большое число игр», «в среднем одинаково часто» и т.д.? Строгое обоснование дает математическая теория вероятностей, но оно выходит за рамки этой статьи. Мы, оставаясь на уровне правдоподобных нестрогих рассуждений, рассмотрим еще несколько подобных примеров.

## Случайные расстановки

**5.** *В клетках шахматной доски $8 \times 8$ стоят плюсы и минусы, причем плюсов столько же, сколько минусов (по 32). Докажите, что можно так расставить 8 ладей, не бьющих друг друга, чтобы не меньше 4 ладей стояли на плюсах.*

▷ Будем ставить ладьи (пронумеруем их от 1 до 8) на доску случайным образом, считая равновероятными все расстановки, удовлетворяющие правилам (ладьи не бьют друг друга: на каждой вертикали и горизонтали их по одной).

Выберем какую-то одну из ладей. На какую клетку у нее больше шансов попасть: на a1 или, скажем, на e1? Одинаково, потому что в любой расстановке можно переставить вертикали a и e (вместе с ладьями, на них стоящими), и расстановки одного типа соответствуют расстановкам другого. По тем же причинам шансы попасть на e1 такие же, как, скажем, на e4 (переставляем первую и четвертую горизонтали). Аналогичные рассуждения показывают, что вообще у этой ладьи равные шансы попасть во все клетки доски.

Следовательно, у нее равные шансы попасть на плюс и на минус (по условию ровно половина клеток занята плюсами и половина минусами). Если за плюс ей дают рубль, а за минус отбирают, то в среднем ладья будет оставаться «при своих».

Теперь внимание: раз каждая ладья остается в среднем при своих, то и вся команда остается в среднем при своих. А если бы утверждение задачи было неверно (всегда больше ладей на минусах, чем на плюсах), то команда бы всегда проигрывала, а потому и в среднем бы проигрывала – противоречие. ◁

## Случайные точки

**6.** *Известно, что более половины поверхности Земли занимают океаны. Используя из географии только этот факт, докажите, что можно найти две диаметрально противоположные точки, обе попавшие в океан.*

▷ Будем выбирать случайную точку поверхности Земли (или глобуса). Если делать это много раз, то больше чем в половине случаев выбранная точка $A$ будет попадать в океан (она равномерно распределена по всей поверхности Земли, а океаны занимают больше половины).

Противоположная ей точка $B$ тоже равномерно распределена по всей поверхности Земли и потому более чем в половине случаев попадает в океан. Значит, эти случаи ($A$ в океане и $B$ в океане) иногда должны происходить одновременно, что и требовалось доказать. ◁

Если это рассуждение кажется вам подозрительным (что значит «случайная равномерно распределенная по поверхности Земли точка»? почему доля случаев, в которых она попадает в океан, определяется долей океанов по площади?) – это не зря. Строгое его обоснование снова выходит за пределы этой статьи.

Однако по существу то же самое рассуждение можно изложить и без вероятностей. Пусть $S$ – часть Земли, занятая океанами. Рассмотрим симметричную ей относительно центра Земли часть $S'$. (Она состоит из точек, у которых диаметрально противоположные попадают в океан.) Площади $S$ и $S'$ одинаковы (симметрия сохраняет площади). Если утверждение задачи было неверно, и общих точек у $S$ и $S'$ нет, то получается противоречие, две пересекающиеся части сферы содержат более 50% площади каждая.

В следующей задаче такой простой перевод рассуждения на язык площадей невозможен.

**7.** *На белой сфере есть черное пятно, занимающее не более 10% ее площади. Покажите, что можно вписать куб в сферу таким образом, чтобы ни одна его вершина не попала в это пятно.*

▷ Возьмем куб $ABCDA'B'C'D'$ (надлежащего размера) и будем вписывать его в сферу случайным образом. При большом количестве испытаний доля случаев, когда вершина $A$ попадает в пятно (вероятность события «$A$ черная») равна $1/10$, поскольку вершина $A$ равномерно распределена по всей сфере.

По тем же причинам вершина $B$ оказывается черной в 10% случаев, и то же самое можно сказать про любую из 8 вершин. Значит, остается как минимум 20% = 100% − 80% случаев, в которых все вершины белые (может быть, меньше за счет случаев, когда несколько вершин черные). Поэтому куб можно вписать требуемым способом. ◁

На самом деле тут мы уже серьезно жульничаем и заметаем под ковер основную часть рассуждения: почему слова «случайно вписанный куб» имеют смысл и почему вершина равномерно распределена по сфере. Интуитивно это выглядит правдоподобно и действительно может быть строго обосновано (хотя я не знаю никакого решения этой задачи в рамках школьной программы).

## Случайные раскраски

Рассмотрим произвольный *граф* – несколько точек (*вершин*), соединенных линиями (*ребрами*). Раскрасим вершины графа в два цвета (скажем, черный и белый). Ребро назовем *разноцветным*, если оно (как вы уже догадались) соединяет вершины разных цветов.

Количество разноцветных ребер зависит от раскраски. Если мы хотим сделать это число минимальным, нет ничего проще – достаточно покрасить все вершины в один и тот же цвет. А что будет, если мы хотим сделать его максимальным (для данного графа)? Ответ зависит от графа, но оказывается, что всегда можно гарантировать как минимум половину разноцветных ребер.

**8.** *Докажите, что в произвольном графе можно раскрасить вершины в два цвета таким образом, чтобы как минимум половина ребер оказались разноцветными.*

▷ Будем раскрашивать вершины случайным образом (для графа с $n$ вершинами возможно $2^n$ раскрасок, и мы считаем все их равновероятными). Для каждой раскраски посчитаем, сколько ребер окажутся разноцветными. Это, как говорят, «случайная величина» – она зависит от выбора раскраски.

Каково ее математическое ожидание (среднее значение при большом числе испытаний)? Другими словами, пусть в каждой вершине графа бросают монету (независимо от других вершин, так что все комбинации равновероятны), а на каждом ребре дежурит игрок, который получает рубль, если ребро окажется разноцветным (монеты в его концах выпадут по-разному). Сколько в среднем будут зарабатывать игроки за игру?

Каждый игрок в среднем выигрывает в половине игр (есть четыре варианта комбинации цветов на его концах, и в двух из них ребро разноцветное). Значит, каждый игрок в среднем выигрывает полтинник за игру, а вся команда в среднем выигрывает $V/2$ рублей, где $V$ – число ребер.

Отсюда очевидно следует, что есть раскраски, когда не менее половины ребер разноцветные (иначе бы в каждой игре команда выигрывала меньше $V/2$, и среднее тоже было бы меньше $V/2$). ◁

На самом деле это утверждение имеет простое конструктивное доказательство (см. задачу 19).

## Случайные проекции

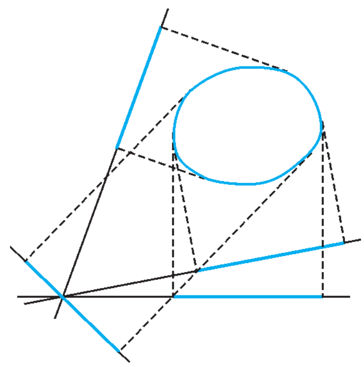
В этом разделе мы применим вероятностные соображения для решения такой задачи:

**9.** *На плоскости нарисована выпуклая фигура, ограниченная кривой длины L. Докажите, что ее диаметр, т.е. максимальное расстояние между двумя ее точками, не меньше* $L/\pi$.

Появление здесь числа $\pi$ не случайно: неравенство задачи обращается в равенство для случая окружности диаметра $D = L/\pi$.

▷ Мы докажем более сильное утверждение: средняя длина проекции нашей фигуры на случайное направление равна $L/\pi$. (Мы проводим прямую в случайно выбранном направлении, и потом из каждой точки фигуры опускаем перпендикуляр на эту прямую. Получается отрезок – проекция фигуры на прямую.)

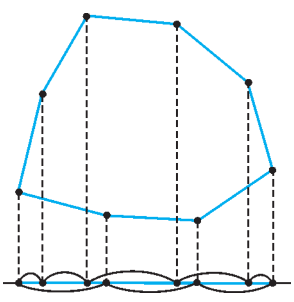
Из этого утверждения следует, что найдется направление, проекция на которое имеет длину не меньше $L/\pi$ (среднее значение не может быть больше всех). А тогда и точки фигуры, которые попадают в концы проекции, находятся на таком же или большем расстоянии (наклонная длиннее перпендикуляра).



Как же доказать обещанное? Рассмотрим кривую, ограничивающую фигуру. Проекция фигуры – это все равно что проекция ее границы. Разобьем эту кривую на множество очень маленьких частей – настолько маленьких, что каждую из них можно считать отрезком. Пусть все эти отрезки будут равной длины.

Проекция всей границы складывается из проекций этих отрезков. Точнее, удвоенная проекция всей границы равна сумме проекций отрезков (так как они покрывают ее два раза – туда и обратно).

Это верно для любого направления проектирования – значит, и *средняя* длина проекции фигуры равна сумме средних длин проекций каждого из отрезков, составляющих ее границу. Но отрезки эти равной длины, и потому средние длины их проекций одинаковы (поскольку мы проектируем на случайное направление, то направление самого отрезка роли не играет). Значит, средняя длина проекции кривой



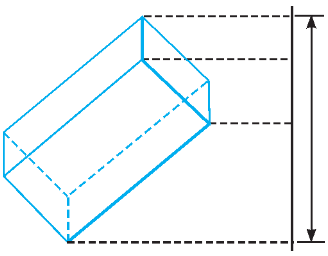
зависит только от числа этих отрезков, т.е. пропорциональна длине кривой, и коэффициент пропорциональности можно вычислять на примере окружности. ◁

## Две коробки

**10.** *В московском метро можно провозить коробки, у которых сумма измерений (длины, ширины и высоты) не превосходит некоторой границы. Возникает вопрос: можно ли перехитрить правила, поместив одну коробку внутрь другой? Другими словами, пусть один прямоугольный параллелепипед целиком содержится внутри другого. Может ли сумма измерений внутреннего быть больше суммы измерений внешнего?*

▷ Оказывается, что нет, но доказать это не так просто. Есть несколько красивых доказательств, и одно из них использует вероятностные соображения (примерно такие же, как в предыдущем разделе).

Будем проектировать коробку на прямую случайно выбранного направления (в пространстве). Из картинки видно, что проекция коробки складывается из проекции трех отрезков, идущих по ее высоте, длине и ширине.



Следовательно, средняя длина проекции (математическое ожидание) равно сумме средних длин проекций этих отрезков. А средняя длина проекции отрезка пропорциональна его собственной длине (с каким-то коэффициентом, нам сейчас не важно – с каким). Поэтому средняя длина проекции коробки пропорциональна сумме ее измерений.

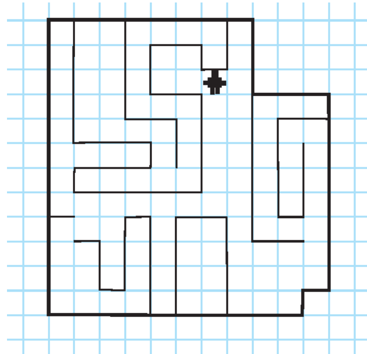
Теперь представим себе две коробки, одну внутри другой. На какую прямую их ни проектируй, проекция внутренней коробки всегда находится внутри проекции внешней и потому имеет меньшую или равную длину. Значит и средние длины проекций находятся в том же соотношении, а они пропорциональны сумме измерений. ◁

## Случайные блуждания

**11.** *На листке из тетрадки в клеточку нарисован лабиринт и в нем робот, который по командам «вверх», «вниз», «влево» и «вправо» ходит из клетки в клетку. Стены лабиринта идут по сторонам клеток, выхода из лабиринта нет. Если, начав двигаться по команде, робот упирается в стену, то он остается на месте. Мы не знаем, какой лабиринт нарисован, но знаем размер листка, измеренный в клеточках. Докажите, что можно составить последовательность команд, которая гарантирует, что робот побывает во всех клетках доступной ему части лабиринта.*

▷ Сразу же признаемся, что не только наше рассуждение будет чистым доказательством существования (без предъявления способа построения такой последовательности), но и длина такой последовательности будет астрономической.

Будем давать роботу случайные команды и смотреть,



с какой вероятностью он обойдет лабиринт. Пусть в лабиринте есть замкнутый обход из не более чем $N$ шагов. Тогда при любом начальном положении робота есть некоторая вероятность, что при подаче ему $N$ случайных команд он пойдет как раз по этому пути и обойдет лабиринт. Она не меньше $4^{-N}$ , а вероятность не обойти лабиринт не больше $\left(1-4^{-N}\right)$. Процесс можно повторять: если мы дадим $2N$ случайных команд, то вероятность *не* обойти лабиринт не больше $\left(1-4^{-N}\right)^2$ (неудача на первом этапе происходит в доле $\left(1-4^{-N}\right)$ случаев или меньше, и в каждом из них вероятность сокращается во столько же раз за счет второго этапа).

Вообще, после $Nk$ случайных команд вероятность неудачи (в данном лабиринте с данным начальным положением) не больше $\left(1-4^{-N}\right)^k$.

Хотя основание степени и близко к единице, но все равно это число можно сделать сколь угодно малым, увеличивая $k$. Заметим, что для всех лабиринтов на доске данного размера можно выбрать одно общее $N$. После этого возьмем $k$ таким, чтобы $\left(1-4^{-N}\right)^k$ стало меньше $1/M$, где $M$ – число разных способов нарисовать лабиринт и робота внутри него (на листке бумаги данного размера их конечное число). Тогда для каждого из $M$ вариантов расположения вероятность неудачи меньше $1/M$, и потому с положительной вероятностью последовательность команд длины $Nk$ окажется удачной для всех вариантов. Значит, такая удачная последовательность существует, что и требовалось доказать. ◁

## Несравнимые множества

В этом разделе мы предполагаем знакомство с элементами комбинаторики (множества, подмножества, число сочетаний).

**12.** *Пусть $A$ – множество из $2n$ элементов. Мы хотим выбрать несколько подмножеств множества $A$, причем требуется, чтобы они были несравнимы: ни одно из них не должно быть подмножеством другого. Какое максимальное число подмножеств можно выбрать?*

▷ Заметим, что требование выполняется, если все выбранные подмножества одного размера (и разные). Чтобы их число было максимальным, надо взять размер $n$, поскольку число сочетаний $C_{2n}^k$ при данном $n$

максимально при $k = n$. Это следует из формулы $C_v^u = v!/u(v-u)!$. Но можно это объяснить и без формул: из $k$-элементного подмножества можно получить $(k+1)$-элементное, добавив любой из оставшихся $2n - k$ элементов, при этом каждое $(k+1)$-элементное множество получится $k+1$ раз, так что

$$\frac{C_{2n}^{k+1}}{C_{2n}^{k}} = \frac{2n-k}{k+1},$$

и это отношение больше 1, когда $2n - k > k + 1$, т.е. $k < n - 1/2$.

Таким образом, если брать подмножества одинакового размера, то наилучший вариант получится, если взять $C_{2n}^{n}$ подмножеств размера $n$. Остается доказать – и это самое трудное, – что даже если брать множества разных размеров, то ничего лучшего не найдется.

Рассмотрим следующий случайный процесс: начав с пустого множества, мы добавляем элементы множества $A$ в случайном порядке, пока не получится все множества $A$. Пусть $X$ – некоторое подмножество множества $A$. Какова вероятность того, что мы в ходе этого процесса пройдем через $X$ (в какой-то момент будут добавлены все элементы $X$ и только они)?

Легко сообразить, что она равна $1/C_{2n}^{k}$, где $k$ – число элементов в $X$. В самом деле, в ходе добавления элементов мы пройдем ровно через одно $k$-элементное множество (после $k$ шагов), и все $k$-элементные множества равновероятны (симметрия). Таких множеств $C_{2n}^{k}$, значит, каждое из них появляется в $1/C_{2n}^{k}$ доле случаев (имеет вероятность $1/C_{2n}^{k}$).

Пусть теперь мы выбрали какое-то количество подмножеств (не обязательно одинакового размера), соблюдая требования задачи (никакие два выбранных подмножества не сравнимы). Для каждого из выбранных подмножеств вероятность пройти через него равна $1/C_{2n}^{k}$, что, как мы знаем, не меньше $1/C_{2n}^{n}$. С другой стороны, события эти, как говорят, «несовместны»: не может быть так, что в ходе добавления элементов мы сначала получим одно выбранное подмножество, а потом другое (они ведь несравнимы). Поэтому сумма вероятностей этих событий не больше 1, и потому количество событий, т.е. количество выбранных подмножеств, не больше $C_{2n}^{n}$, что и требовалось доказать. ◁

## Еще несколько задач

**13.** Среди любых трех участников математического кружка есть хотя бы одна девочка. Сколько мальчиков может быть в этом кружке?

**14.** Решите задачу 4 аналогично предыдущим (задачи 1, 2).

*Указание.* Докажите, что если сумма трех чисел равна 5, то сумма некоторых двух из них не меньше $10/3$. То же самое верно, если сумма трех чисел больше 5.

**15.** Сделайте рассуждение раздела «Вероятности и ожидания» строгим, рассмотрев случай, когда все комбинации из трех шаров были вытащены по одному разу.

**16.** Покажите, что можно найти 8 расстановок ладей на шахматной доске, которые вместе используют все 64 клетки доски (по одному разу каждую). Используя этот факт, дайте другое решение задачи 5.

**17.** Не используя вероятностей, докажите такой ослабленный вариант утверждения задачи 7: если черное пятно занимает 10% площади сферы, то в нее можно вписать *прямоугольный параллелепипед*, у которого все вершины попадают в белые точки.

*Указание.* Проведем координатные плоскости, которые делят сферу на 8 равных частей. Будем искать параллелепипед, стороны которого параллельны осям координат. Он задается одной из своих вершин, и она не должна попасть в восемь множеств площади $1/10$.

**18.** Докажите, что для данного графа существует раскраска, делающая все его ребра разноцветными, если и только если в этом графе нет циклов нечетной длины (нельзя вернуться в исходную вершину, нечетное число раз пройдя по ребрам).

**19.** Придумайте конструктивное решение задачи 8 (способ раскраски, который гарантирует, что разноцветных ребер будет не меньше половины).

*Указание.* Красим вершины по очереди «жадным алгоритмом» – каждая следующая красится так, чтобы число разноцветных ребер максимально увеличивалось.

**20.** Вершины графа с $V$ ребрами разрешим красить в три цвета. Докажите. что можно найти раскраску с $2V/3$ или более разноцветными ребрами.

**21.** На листок в линейку случайно бросают иглу, длина которой равна расстоянию между линейками. Докажите, что она пересечет одну из линеек с вероятностью $2/\pi$.

*Указание.* Математическое ожидание числа пересечений зависит (пропорционально) от длины иглы, но не от ее формы, а коэффициент пропорциональности можно найти на примере окружности.

**22.** Несамопересекающаяся кривая длины 22 находится внутри круга радиуса 1. Докажите, что найдется прямая, имеющая с этой кривой по крайней мере 8 общих точек.

**23.** На кальке нарисована фигура площади 2,5. Докажите, что кальку можно так положить на клетчатую бумагу (со стороной клетки 1), что фигура покроет не менее трех узлов сетки. Докажите, что можно положить кальку так, чтобы фигура покрыла не более двух узлов сетки.

**24.** Рассмотрим $R$-окрестность прямоугольной коробки (все точки, которые находятся на расстоянии не более $R$ хотя бы от одной из точек коробки; внутренность коробки также включаем в ее окрестность).

а) Покажите, что объем этой окрестности можно найти по формуле

$$V + S \cdot R + L \cdot \pi R^2 + \frac{4}{3}\pi R^3 \; 2/\pi,$$

где $V$ – объем коробки, $S$ – площадь ее поверхности, $L$ – сумма ее измерений.

б) Если одна коробка находится внутри другой, то и $R$-окрестность первой находится внутри $R$-окрестности второй. Сравнив их объемы при больших $R$, покажите, что сумма измерений внутренней коробки не больше суммы измерений внешней (задача 10).

**25.** Решите задачу 11, не используя вероятностных соображений.

*Указание.* Рассмотрим все возможные лабиринты и расположения робота по очереди. Выберем какой-то первый вариант и напишем соответствующую последовательность. (Если этот вариант осуществится на самом деле, то нам повезло.) Для второго варианта эта последовательность (вообще говоря) не годится, но тем не менее посмотрим, куда она приводит, и допишем то, что нужно, чтобы обойти лабиринт второго типа. И так далее.

**26.** Какое максимальное число несравнимых подмножеств можно выбрать в $(2n+1)$-элементном множестве?