IX ЗАОЧНАЯ ИНТЕРНЕТ-ОЛИМПИАДА ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ И СТАТИСТИКЕ

Решения и ответы

I. Эссе

1. Дефицит бумаги. Однажды в одном крупном городе управление образования обратилось к математикам с просьбой рассчитать, сколько нужно иметь запасных бланков ЕГЭ, чтобы хватило всем выпускникам города. Математики сделали расчет и сказали, что нужно иметь 10 тысяч запасных бланков, и тогда всем выпускникам хватит почти наверняка (с вероятностью не менее 0,99). Управление поблагодарило математиков, заготовило 10 тысяч запасных бланков и разделило их поровну между всеми пятью городскими пунктами проведения экзамена, поскольку на каждом пункте ожидалось примерно поровну выпускников. Велико же было удивление начальника управления, когда на трех пунктах бумаги не хватило, а на двух – бумага осталась, и немало. Сначала даже решили, что математики ошиблись в расчетах, однако вскоре стало ясно, что математики все сделали верно – после того, как бумагу оттуда, где она осталась, перевезли туда, где ее не хватило, остался некоторый запас. В чем же была ошибка, и как нужно было действовать, чтобы не попасть в такую неприятную ситуацию?

Возможная идея эссе. Скорее всего, математики опирались на оценку среднего m и на оценку дисперсии σ^2 случайной величины X «Необходимое число запасных бланков для города». Есть причины считать, что распределение этой случайной величины близко к нормальному, и поэтому несложно оценить нужный запас, чтобы его хватило с любой наперед заданной вероятностью, скажем 0,99. А именно, требуется $m+k\sigma$ запасных бланков, где k – некоторый коэффициент, зависящий от выбранной вероятности 0,99.

Но на каждом из пунктов приема экзамена среднее число запасных бланков свое и стандартное отклонение тоже свое. По условию, на всех пунктах выпускников примерно поровну, следовательно, можно считать, что среднее и дисперсия числа бланков на всех пунктах примерно одинаковы и равны a и s^2. И поэтому на каждый пункт требуется $a+ks$ запасных бланков.

При этом $m=5a$ и, в силу независимости пунктов количества бумаги на разных пунктах $\sigma^2=5s^2$, откуда $s=\frac{\sigma}{\sqrt{5}}$. Чиновники же передали на каждый пункт

$$\frac{m+k\sigma}{5}=a+k\frac{s}{\sqrt{5}}$$

листов, что, безусловно, меньше, чем $a+ks$.

Например, если $a=1700$, $k=3$ и $s=223$ для каждого пункта, то $a+ks=1700+3\cdot 223=2369$ листов. А на весь город нужно $5\cdot 2369=11845$ листов.

Математиков же просили сделать расчет для города. Они нашли, что $m=5a=8500$, $\sigma=\sqrt{5}s=498,6$ и дали результат для всего города:

$m+k\sigma=8500+3\cdot 498,6=9995,8$ листов и округлили вверх до 10000 листов.

На весь город бумаги хватило, а на часть пунктов – нет.

Нужно было грамотно ставить задачу математикам.

2. Кубик в форме шара. Занятная вариация обычного шестигранного кубика, который на самом деле шарик, или шарика, который на самом деле – кубик (см. рисунок). Шарик свободно катается, но, остановившись, всегда встаёт точно одной из шести нарисованных «граней» кверху. При этом все шесть фиксированных положений равновероятны (вероятность любого другого промежуточного положения нулевая). Представьте, что вам дано задание изготовить такой круглый кубик. Как бы вы подошли к делу? Придумайте и подробно опишите хотя бы один способ изготовить такое невозможное чудо.

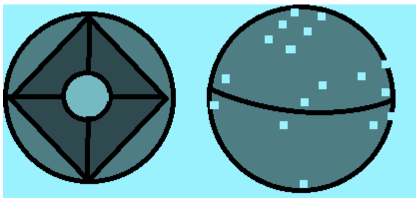
Возможная идея эссе. Если внутри шара находится полость в виде правильного октаэдра, а в ней – тяжелый шарик, то шарик начинает вести себя, как игральная кость. Автор одного из эссе предложил заполнить полость вязкой жидкостью, например, маслом, чтобы шарик не громыхал при перекатывании. Другой автор предложил сделать грани октаэдра не плоскими, а несколько выгнутыми к центру.

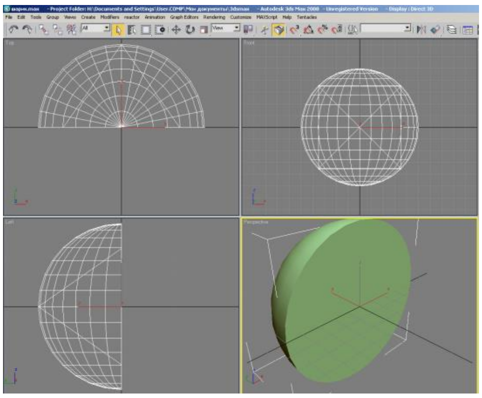
Другие предложения касались магнитных пластинок. Но в этом случае шарик следует бросать только на железный лист.

Было несколько предложений, связанных с помещением кубика внутрь шарика или просто размещения шести грузиков, но, конечно, это не приведет к эффекту, так как симметрично расположенные грузики будут находиться в равновесии постоянно.

Мы публикуем здесь лучшее из присланных эссе по поводу шарокубика. Автор Максим Карпенко (6 класс. Москва, г.Зеленоград). Эссе публикуется с незначительными редакторскими правками.

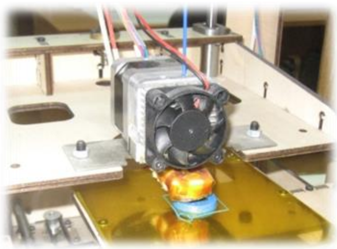
Эссе Максима Карпенко. Сначала была идея сделать шарик из двух половинок, просверлить в нём дырочки и вставить по грузику в место, где должны быть метки (точки). После размышлений я пришел к выводу, что у такого шарика будут много точек равновесия, и он будет не всегда останавливаться цифрой наверху.





Потом появилась идея сделать шарик с полостью внутри, в которой будет перемещаться грузик и застревать в таких положениях, чтоб в этот момент сверху была цифра. Оказалось, что полость должна быть в форме октаэдра[1]. Идея понравилась папе, и мы решили проверить её на практике, сделав шарик на 3D принтере и оставив внутри него полый октаэдр в котором помещается железный шарик.

С исполнением идеи помогал папа, он нарисовал модель половинки шарика в программе 3dsmax и затем с помощью 3D принтера сделал шарик.



Так как шарик делали первый раз и не было возможности исправить некоторые ошибки, то он получился не совсем идеальный. Хотелось бы попробовать разные соотношения размеров шарика и грузика, или сделать шарик из свинца, а ещё попробовать ртуть. Выяснилось, что надо либо лучше калибровать принтер, либо в проекте учитывать неточность его работы

[1] Октаэдр – восьмигранник с шестью симметрично расположенными вершинами.

по разным осям. Так получилось, что шарик вышел немного сплюснутым с полюсов – примерно на 1 мм при диаметре 25 мм. (4%).

На полюсах потом были изображены цифры 1 и 6. Хочу заметить, что для обычного игрального кубика такая погрешность незаметна, а для этого кубика она может изменить число, которое должно выпасть. К тому же из-за того, что шарик делали на принтере, его поверхность была немного ребристой и иногда он не точно останавливается на цифре сверху, а рядом с ней (внутренние грани тетраэдра практически невозможно дорабатывать вручную). Также не был решен вопрос точного совмещения двух половинок. Ещё при проведении пробных бросаний выяснилось, что на обычном столе такой шарик бросать неудобно, он часто скатывается и падает на пол. Нужна большая чаша, пиала или блюдо без острых бортиков.

Вот она, долгожданная статистика этого эксперимента для 50 бросаний.

1	2	3	4	5	6	Промеж. положение
12 раз (24%)	5 раз(10%)	4 раза(8%)	10 раз(20%)	5 раз(10%)	11 раз(22%)	3 раза(6%)


Видимо для массового производства 3D принтер не очень удобен (печать шарика заняла около часа), и надо отливать его из пластика.

3. Закон Ципфа. (Zipf's law[2]). Существует множество видов данных, которые подчиняются удивительной закономерности – закону Ципфа. Возьмем, к примеру, население городов некоторой страны. При некоторых естественных условиях формирования городов численность их населения приходит к состоянию, когда население крупнейшего города примерно вдвое больше, чем население второго по численности, в три раза больше, чем население третьего и т.д.

Закон Ципфа встречается в лингвистике: в естественном языке (русский, английский и т.п.) самое распространенное слово встречается примерно вдвое чаще, чем второе по частоте, втрое чаще, чем третье по частоте и т.п.

Математически закон можно описать так: пусть дан набор количественных характеристик некоторого множества (население городов, частота слов, доходы людей и т.п.). Упорядочим числа от наибольшего к наименьшему:

$$A_1 > A_2 > A_3 > \ldots > A_k > \ldots > A_n.$$

Номер числа в этой последовательности называют рангом. Закон Ципфа выполняется для этих данных, если $A_1 \approx kA_k$. Графически это означает, что точки с координатами $(k; A_k)$ лежат вблизи гиперболы $y = \dfrac{A_1}{x}$.

[2] Закон носит имя американского лингвиста Джорджа Кингсли Ципфа, хотя впервые был описан задолго до Ципфа французским стенографистом Жаном-Батистом Эсту в работе «Диапазон стенографии» в 1908 году. Обоснования закона Ципфа появились недавно, уже в XXI веке, но до сих пор неясно, опирается ли закон Ципфа больше на природу величин, которые ему подчиняются, или на свойства случайных последовательностей.

В качестве примера мы взяли численность населения шести крупнейших городов США: Нью-Йорка (ранг 1), Лос-Анджелеса (2), Чикаго (3), Хьюстона (4), Филадельфии (5) и Финикса (6). На диаграмме (рис. 1) показана численность этих городов по рангам (точки). Линия – гипербола $y = \frac{A_1}{x}$, где $A_1 = 8405837$ – численность Нью-Йорка, а x – ранг города от 1 до 6. Для этой совокупности закон Ципфа, похоже, выполняется.

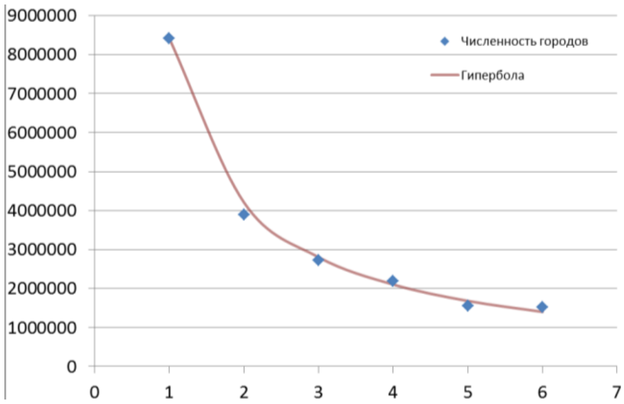


Рис. 1

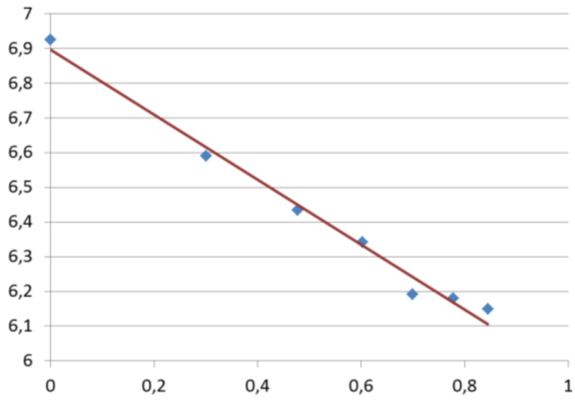


Рис. 2

Тот, кто знаком с логарифмами, может получить еще более удобную картинку, прологарифмировав обе величины: $\lg y = \lg A_1 - \lg x$. Переходя к новым переменным $z = \lg y$ и $u = \lg x$, мы получаем точки вблизи прямой $z = \lg A_1 - u$. Такая интерпретация удобна, потому что человеческий глаз чутко реагирует на отклонение от прямой. Легко устроить «глазомерную» проверку закона Ципфа: если для некоторых данных окажется, что точки $(\lg k; \lg A_k)$ расположены вблизи прямой, можно предположить, что эти данные подчиняются закону Ципфа. На рис.2 пример про города США показан в логарифмическом виде (по оси абсцисс десятичный логарифм ранга, по оси ординат – десятичный логарифм численности населения).

Интересно, выполняется ли закон Ципфа для городов России? Для городов мира? Для объема воды в крупнейших озерах? Для высот высочайших гор мира? Для числа пассажиров городских метрополитенов? Нужные данные можно найти в интернете. Например, некоторые полезные таблицы размещены на странице

http://ptlab.mccme.ru/node/350 .

Может быть, вам удастся найти какой-нибудь другой вид данных, который подчиняется закону Ципфа? У вас есть шанс стать первооткрывателем.

Комментарий. Возможных идей множество. К сожалению, в большинстве авторы ограничивались пассивной проверкой закона Ципфа для некоторых данных, при этом вольно трактуя результаты. Интересно эссе Анастасии Уваровой (10 класс, Москва), где выдвинута занятная гипотеза об антропогенности данных, для которых закон выполняется (Анастасия пользуется оборотом «человеческие объекты»). Мы публикуем здесь эссе Алии Трапезниковой (11 класс, Казань). Это эссе привлекло нас тем, что, исследуя разные данные на согласованность, Алия обнаружила обобщенный закон Ципфа, согласно которому указанной закономерности подчиняются не исходные данные, а некоторые их степени. Другая интересная идея – переход от совокупностей данных к их подмножествам, при этом исследуется вопрос об изменении характера зависимости. Очевидным недостатком работы является отсутствие попыток отбросить данные, плохо согласующиеся с исследуемым законом и отсутствие попыток понять, как природа данных может влиять на их принадлежность исследуемому закону.

Эссе Алии Трапезниковой. Для начала мы проверили, выполняется ли закон Ципфа для разных типов данных: списка богатейших людей мира, рейтинга самых популярных сайтов Рунета, списка самых кассовых фильмов, населения стран мира и пр. Для этого построили точки $(\lg N; \lg w)$, где w – численная характеристика объекта (размер состояния, количество пользователей и т.п.), N – его номер в рейтинге (рис. 1).

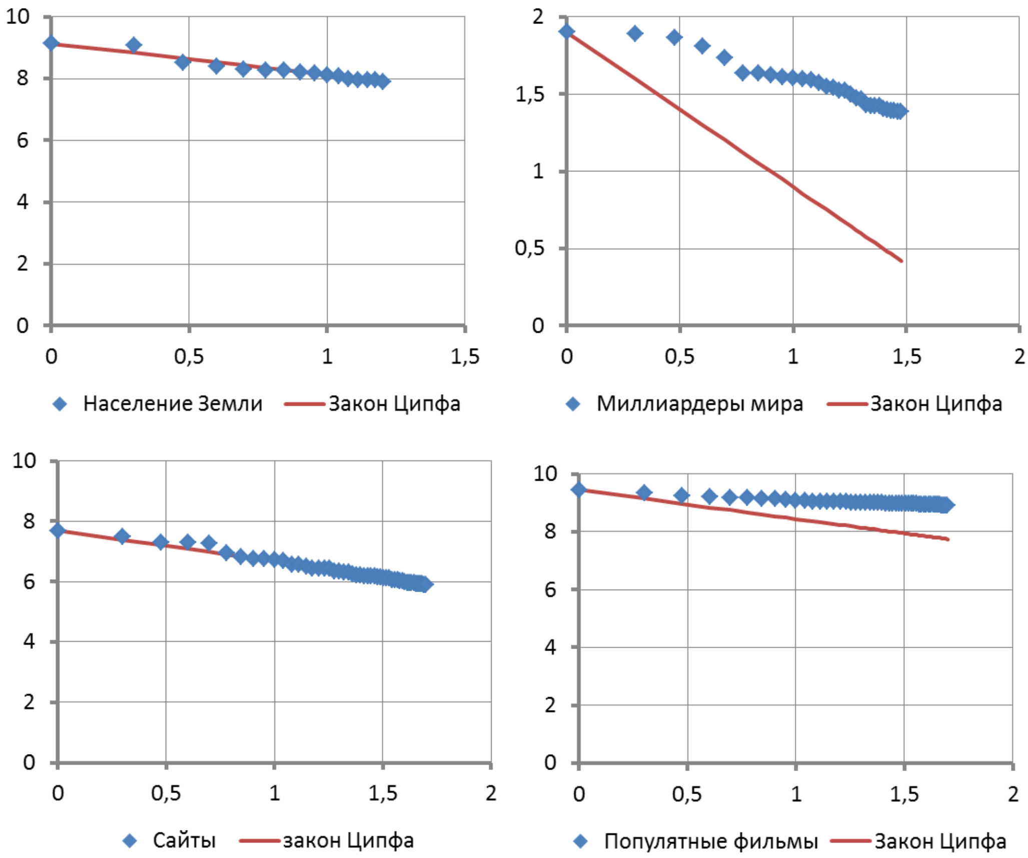


Рис 1. Данные, для которых проводились исследования

Мы заметили, что полученные точки группируются вблизи прямой, но эта прямая не всегда совпадает с ожидаемой $y = \lg A - x$, где A – численная характеристика первого в рейтинге объекта.

Тогда мы решили проверить, как связаны между собой прямые, вблизи которых группируются точки $(\lg N; \lg w)$ для разных рейтингов одного типа данных; например, для населения городов мира, городов разных стран и республик; для мировых кассовых сборов фильмов за всю историю кинематографа, фильмов последнего десятилетия, мультфильмов. Оказалось, что для кассовых сборов разных выборок фильмов точки $(\lg N; \lg w)$ (где N - номер фильма в рейтинге выборки, w – сумма сборов) лежат вблизи прямой $y = ax + b$, где $a \approx -0,3$ (рис. 2).

Похожая картина получилась для численности населения городов в разных странах ($a \approx -0,8$), но прямая для городов мира выбивается из общего фона, её угловой коэффициент примерно –0,38 – заметно больше других (рис.3)).

Очевидно, закон Ципфа не может одновременно выполняться для населения городов мира и городов каждой из стран, рейтинга популярности фильмов за последние десять лет и за последний год, и т.д. Мы задумались, как дожен меняться угловой коэффициент прямой, вблизи которой группируются точки $(\lg N; \lg w)$), если рассматривать всё более мелкие группы (мир, часть света, страну, регион, …).

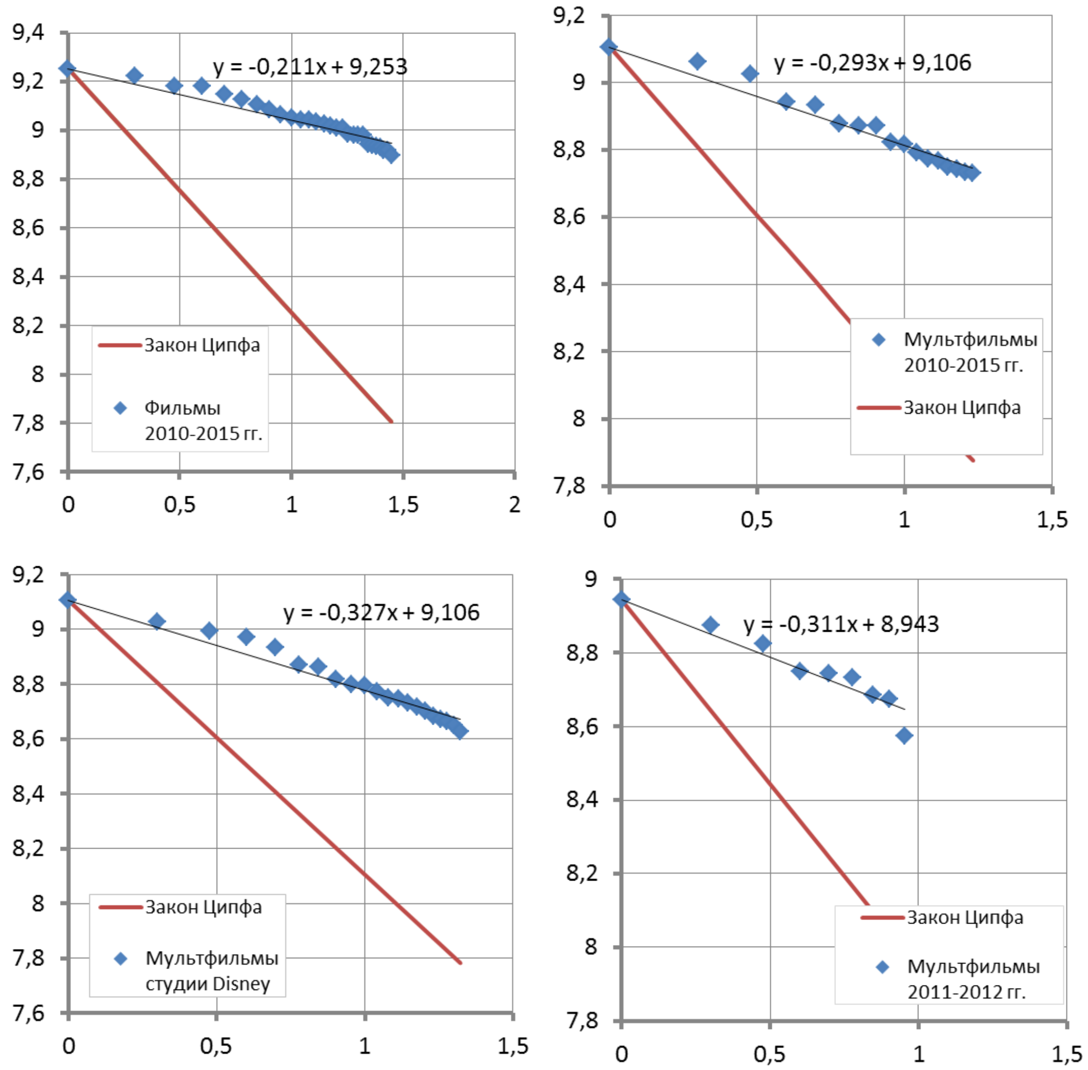


Рис. 2. Кассовые сборы популярных фильмов

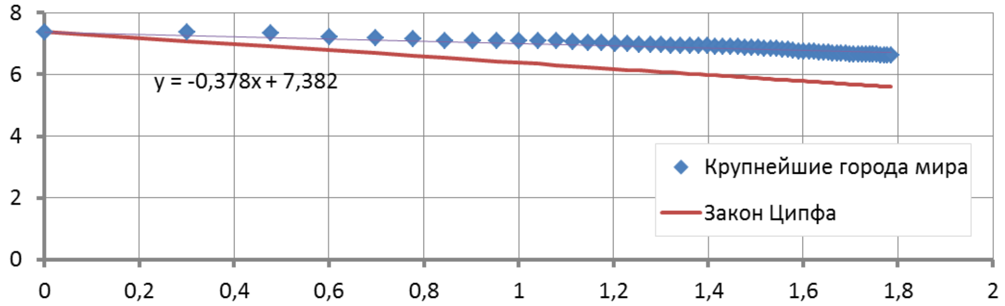


Рис. 3. Население городов стран мира

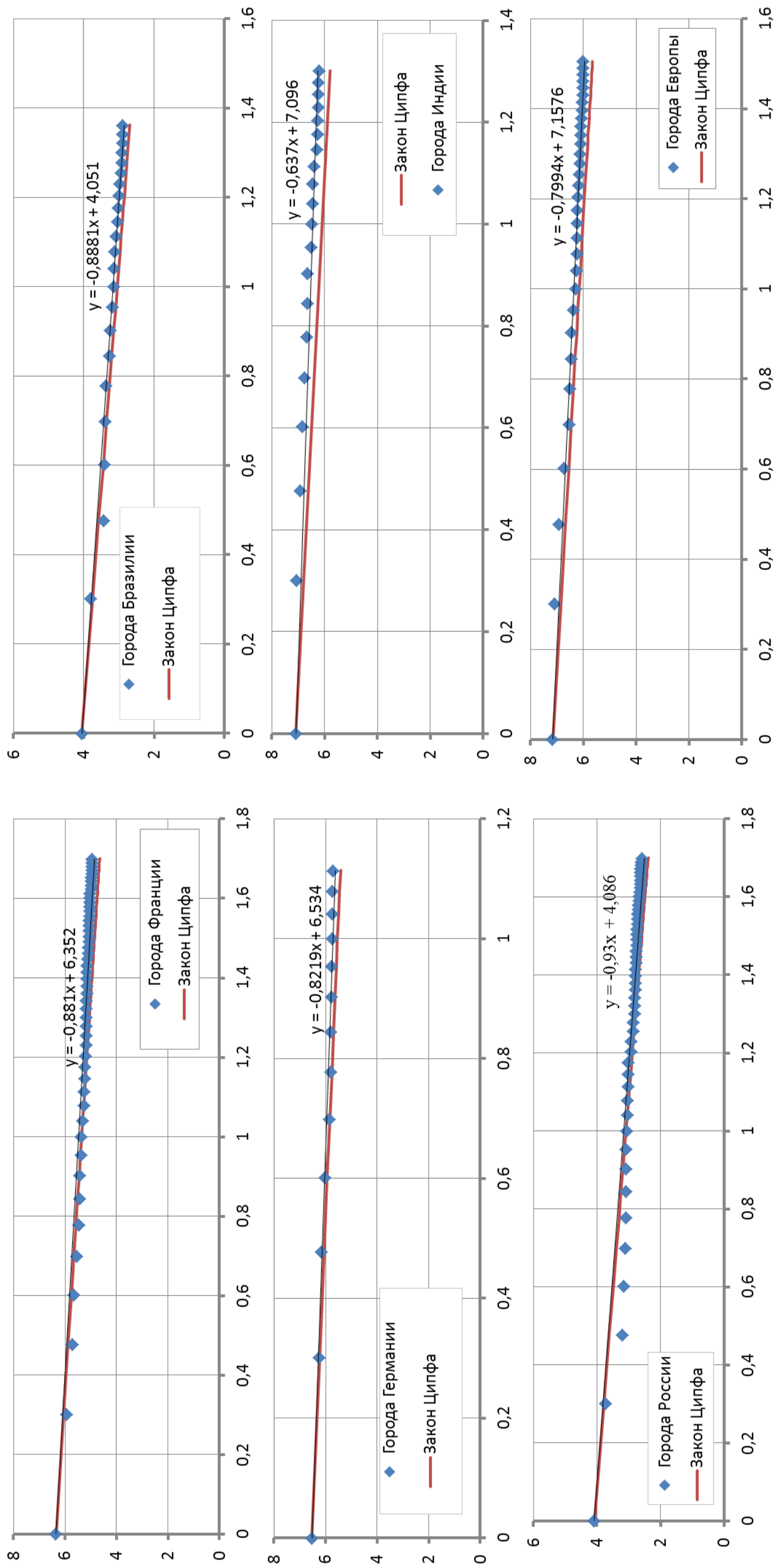


Рис. 4. Цепочка «Мир – Европа – Россия – Татарстан и Краснодарский край»

Пусть для некоторого вида данных (например, для населения городов мира) точки $(\lg N; \lg w)$ группируются вблизи прямой $y = ax + b$, то есть $w \approx N^a \cdot 10^b$. Тогда, если мы будем рассматривать только некоторые (выбранные случайным образом или по некоторому естественному признаку) данные того же вида (например, только европейские города) точки $(\lg N; \lg w)$ должны группироваться вблизи прямой с угловым коэффициентом, меньшим a. Действительно: если первый в «частном» рейтинге объект имеет номер n в «общем» рейтинге, то второй в «частном» рейтинге, скорее всего, будет иметь номер чуть меньше $2n$, третий – меньше $3n$, и так далее.

Эти теоретические рассуждения подтвердились при проверке: если для городов мира угловой коэффициент а оказался примерно $-0,38$, то для городов Европы он стал примерно $-0,80$, для России он оказался около $-0,90$, для республики Татарстан – около $-1,44$, и для городов Краснодарского края – примерно $-1,13$ (рис.4).

После этого мы проверили, уменьшается ли угловой коэффициент a, если мы рассматривать списки самых кассовых фильмов, самых кассовых мультфильмов, самых кассовых мультфильмов студии «Disney». Мы увидели, что и для такой цепочки включений верно, что коэффициент a от общего к частному уменьшается (рис. 5). Такие же результаты мы получились для списков богатейших людей мира, России, Америки и пр.

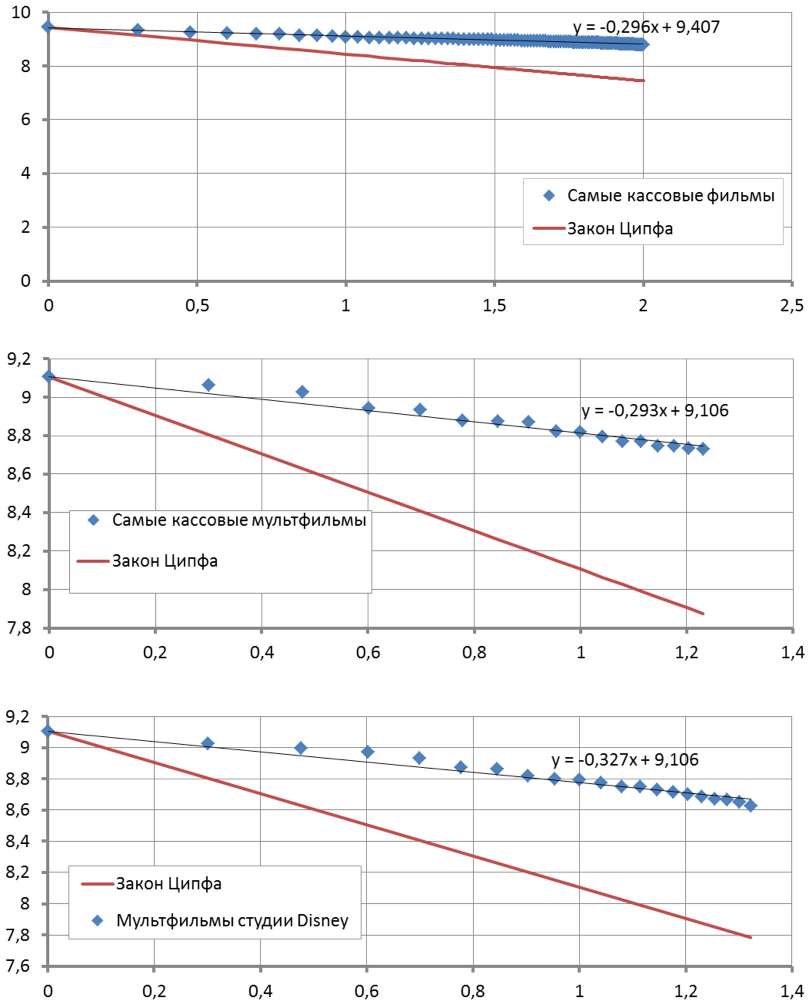


Рисунок 5. Цепочка рейтингов Фильмы – Мультфильмы – Мультфильмы Disney

После этого мы проверили, насколько сильно различаются коэффициенты a, если рассматривать численность городов в разных странах, в разных регионах страны, кассовые сборы фильмов и мультфильмов разных студий, разных лет (рис.5). Оказалось, что эти угловые коэффициенты изменяются в достаточно узких промежутках.

Таким образом, мы заметили, что закон Ципфа в чистом виде выполняется для очень небольшого количества данных. Для большинства исследованных нами данных выполняется другая закономерность: точки $(\lg N; \lg w)$ группируются возле $y = ax + b$, причём коэффициент a тем меньше, чем более узкую выборку мы исследуем. То есть $\lg w \approx a \cdot \lg N + b$, откуда $w \approx N^a \cdot 10^b$. При этом прямую $y = ax + b$ мы строили так, чтобы она проходила через точку $(0; \lg A)$ (где A – числовая характеристика первого объекта первого ранга), то есть $\lg A = 0 \cdot a + b$, то есть $A = 10^b$. Отсюда $w \approx N^a \cdot A$.

В частности, при $a = -1$ получаем исходный вид закона Ципфа. Найденную закономерность можно назвать расширенным или обобщенным законом Ципфа.

Используемые таблицы данных

А) Население стран мира:
https://ru.wikipedia.org/wiki/%D0%9D%D0%B0%D1%81%D0%B5%D0%BB%D0%B5%D0%BD%D0%B8%D0%B5_%D0%97%D0%B5%D0%BC%D0%BB%D0%B8

Б) Миллиардеры мира:
https://ru.wikipedia.org/wiki/%D0%A1%D0%BF%D0%B8%D1%81%D0%BE%D0%BA_%D0%BC%D0%B8%D0%BB%D0%BB%D0%B8%D0%B0%D1%80%D0%B4%D0%B5%D1%80%D0%BE%D0%B2_(2015)

В) Самые популярные сайты: http://top1000-ru.hotlog.ru/?page=1

Г) Список самых кассовых фильмов:
https://ru.wikipedia.org/wiki/%D0%A1%D0%BF%D0%B8%D1%81%D0%BE%D0%BA_%D1%81%D0%B0%D0%BC%D1%8B%D1%85_%D0%BA%D0%B0%D1%81%D1%81%D0%BE%D0%B2%D1%8B%D1%85_%D1%84%D0%B8%D0%BB%D1%8C%D0%BC%D0%BE%D0%B2

II. Задачи

1. Отопительный сезон. Городской муниципалитет Затонска принял правило: отопление в домах следует включать не раньше 26 октября, но только если средняя температура в течение трех предыдущих дней ниже $8°C$. В городе два района – Прибрежный и Заречный.

В Прибрежном районе правило поняли так: если три дня подряд средняя дневная температура каждый день ниже $8°C$, то на четвертый день нужно включить отопление, если этот день случился 26 октября или позже.

В Заречном районе правило поняли иначе: если средняя температура за трехдневный период ниже $8°C$, то на четвертый день нужно включить отопление, если этот день не раньше 26 октября.

В таблице показана средняя дневная температура за несколько дней октября.

Дата	22 окт	23 окт	24 окт	25 окт	26 окт	27 окт	28 окт	29 окт	30 окт
Ср. темп. $°C$	3	9	10	6	3	7	10	6	8

а) **(от 6 класса. 1 балл).** Какого числа отопление включили в Прибрежном районе? Какого числа отопление включили в Заречном районе?

б) **(от 6 класса. 1 балл).** Докажите, что какие бы ни случились дни в октябре, в Заречном районе отопление включат не позже, чем в Прибрежном.

Решение.

а) В Заречном районе вычисляют среднюю температуру за трехдневный период. В период 23 – 25 октября средняя температура $8{,}33°C$, поэтому 26 октября отопление включать нельзя. Зато за период с 24 по 26 октября средняя температура ниже $8°C$, значит, отопление нужно включить 27 октября. В Прибрежном районе не вычисляют среднюю температуру за три дня, а смотрят на температуру в каждый отдельный день. Первый раз три дня подряд средняя дневная температура ниже $8°C$ случилась 25, 26 и 27 октября. Значит, в Прибрежном районе тепло дадут только 28 октября.

б) Если три дня подряд средняя дневная температура за каждый день ниже пороговой $8°C$, то средняя за все три дня также ниже. Обратное неверно, поэтому в Прибрежном районе отопление включат не раньше, чем в Заречном, а возможно – позже.

Ответ: а) в Заречном районе 27 октября, в Прибрежном – 28 октября.

2. Отметка за полугодие. К концу полугодия у Василия Петрова в журнале стояли такие отметки по математике: 4, 1, 2, 5, 2. Перед тем как выставить полугодовую отметку, учитель математики сказал Васе:

– Вася, ты можешь выбрать метод, как вывести твою отметку за полугодие. Предлагаю два варианта. Метод А: среднее арифметическое текущих отметок с округлением до целого. Метод Б: медиана текущих отметок.

Лучший метод для Васи – это такой метод, который даст Васе в полугодии наибольшую отметку.

а) (от 6 класса. 1 балл). Какой метод для Васи лучший?

б) (от 6 класса. 2 балла). Потом учитель подумал и добавил:

– Имей в виду, Василий, что если ты сумеешь выбрать лучший для себя метод, то я поставлю тебе в журнал еще две пятерки, а уже потом выведу отметку за полугодие.

Докажите, что в этих условиях метод А не является для Васи лучшим.

Решение.

а) Среднее текущих отметок 2,8 (округляется до 3), а медиана отметок равна 2. Лучше выбрать метод А.

б) Предположим, что метод А лучший. Следовательно, учитель должен поставить две пятерки. Но тогда среднее после округления остается 3, а медиана поднимается до 4. Следовательно, теперь метод А дает меньше, чем метод Б, поэтому метод А – не лучший.

3. Переводчик (от 6 класса. 1 балл). Василий Петров выполняет задание по английскому языку. В этом задании есть 10 английских выражений и их переводы на русский в случайном порядке. Нужно установить верные соответствия между выражениями и их переводами. За каждое правильно установленное соответствие даётся 1 балл. Таким образом, можно получить от 0 до 10 баллов. Вася ничего не знает, поэтому выбирает варианты наугад. Найдите вероятность того, что он получит ровно 9 баллов.

Решение. Не может случиться так, что угадано ровно девять верных ответов, потому что в этом случае десятый ответ тоже будет верным. Значит, искомая вероятность равна нулю.
Ответ: 0.

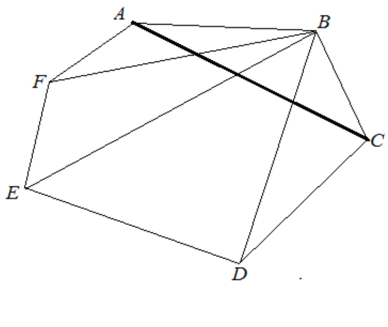
4. Бракованные монеты (от 6 класса. 1 балл). К юбилею Санкт-Петербургских математических олимпиад монетный двор отчеканил три юбилейные монеты. Одна монета получилась правильно, у второй монеты на обеих сторонах оказалось два орла, а у третьей обе стороны – решки. Директор монетного двора не глядя выбрал одну из этих трёх монет и бросил её наудачу. Выпал орёл. Чему равна вероятность того, что на второй стороне этой монеты тоже орёл?

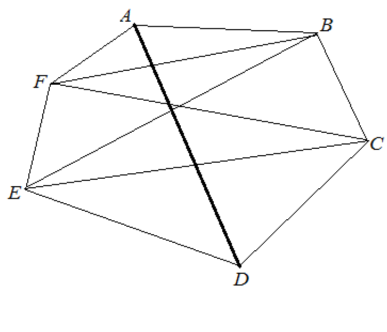
Решение. Известно, что выпал орел. Значит, была выбрана либо первая монета, либо вторая. Всего на этих монетах орел изображен три раза. В одном из этих трех равновозможных случаев на противоположной стороне монеты решка, а в двух других случаях – орлы.

Ответ: $\frac{2}{3}$.

5. Пересекающиеся диагонали (от 7 класса. 2 балла). В выпуклом шестиугольнике независимо друг от друга выбраны две случайные диагонали. Найдите вероятность того, что эти диагонали пересекаются внутри шестиугольника (внутри – то есть не в вершине).

Решение. Всего у шестиугольника 9 диагоналей и, значит, $\frac{9\cdot 8}{2}=36$ пар диагоналей. Осталось понять, сколько пар диагоналей пересекается внутри шестиугольника. Диагонали у шестиугольника бывают двух видов: главные (соединяют противоположные вершины) и неглавные. Например, диагональ AC – неглавная. Всего неглавных диагоналей 6, и каждая пересекает 3 другие диагонали (см. рис.). Главных диагоналей (таких, как AD) всего 3 штуки, и каждая пересекает 4 другие диагонали. Всего получается $\frac{6\cdot 3+3\cdot 4}{2}=15$ точек пересечения (деление пополам нужно, поскольку выражение в числителе учитывает каждую точку пересечения дважды).





Сколько внутри точек пересечения, столько и пар диагоналей, пересекающихся внутри. Следовательно, вероятность того, что две случайно выбранные диагонали пересекаются внутри, равна

$$\frac{15}{36}=\frac{5}{12}.$$

Ответ: $\frac{5}{12}$.

6. Три мишени. Стрелок стреляет по трём мишеням до тех пор, пока не собьёт все. Вероятность попадания при одном выстреле равна p.

а) (от 7 класса. 2 балла). Найдите вероятность того, что потребуется ровно 5 выстрелов.
б) (от 8 класса. 2 балла). Найдите математическое ожидание числа выстрелов.

Решение.

а) Искомое событие наступает, только если четыре первых выстрела дали две сбитые мишени, а последний выстрел был удачным. Вероятность этого равна

$$C_4^2 p^2 (1-p)^2 \cdot p = 6p^3 (1-p)^2.$$

б) Заметим, что при вероятности попадания p среднее число выстрелов, нужных для того, чтобы сбить первую мишень, равно $1/p$. Это интуитивно понятно, а строго следует из свойств геометрического распределения. Значит, еще $1/p$ выстрелов в среднем потребуется на вторую мишень и столько же – на третью.

Ответ: а) $6p^3(1-p)^2$; б) $\frac{3}{p}$.

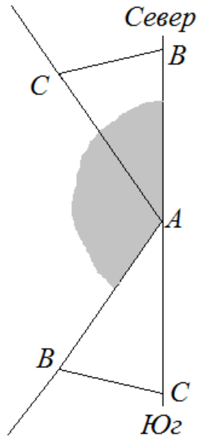
7. **Паросочетания (от 8 класса. 2 балла)**. На соревнования приехали 10 теннисисток, из них 4 из России. По правилам для проведения первого тура теннисистки разбиваются на пары случайным образом. Найдите вероятность того, что в первом туре все россиянки будут играть только с россиянками.

Решение. Возьмем какую-нибудь одну россиянку. В пару к ней попадет другая россиянка с вероятностью $\frac{3}{9}=\frac{1}{3}$. При этом условии рассмотрим двух оставшихся россиянок. Рассуждая так же, получаем, что они окажутся в одной паре с вероятностью $\frac{1}{7}$. Следовательно, искомая вероятность равна $\frac{1}{3}\cdot\frac{1}{7}=\frac{1}{21}$.

Ответ: $\frac{1}{21}$ или прибл. 0,048.

8. (от 9 класса. 2 балла). В треугольнике ABC угол A равен 40°. Треугольник случайным образом бросают на стол. Найдите вероятность того, что вершина A окажется восточнее двух других вершин.

Решение. Проведем через вершину A прямую, идущую строго с юга на север. Событие «Вершина A восточнее двух других» осуществляется тогда и только тогда, когда вершины B и C расположены в западной полуплоскости от проведенной прямой. На рисунке показаны два крайних положения треугольника, при которых вершины B и C не расположены в восточной полуплоскости. Видно, что луч AB должен проходить внутри закрашенного угла, который является внешним к углу A треугольника. Следовательно, вероятность этого события равна $\dfrac{180°-40°}{360°}=\dfrac{7}{18}$.



Ответ: $\dfrac{7}{18}$.

9. Стальные двери (от 9 класса. 2 балла). На заводе имени матроса Железняка изготавливают прямоугольники длиной 2 м и шириной 1 м. Длину отмеряет рабочий Иванов, а ширину, независимо от Иванова, отмеряет рабочий Петров. Средняя ошибка у обоих нулевая, но Иванов допускает стандартную ошибку измерения (стандартное отклонение длины) 3 мм, а Петров допускает стандартную ошибку 2 мм.

а) (от 9 класса. 1 балл). Найдите математическое ожидание площади получившегося прямоугольника.

б) (от 9 класса. 2 балла). Найдите стандартное отклонение площади получившегося прямоугольника в квадратных сантиметрах.

Решение.

а) Пусть X – ширина, а Y – длина вырезанного прямоугольника в метрах. По условию $\mathrm{E}X=2,\ \mathrm{E}Y=1$. Поскольку измерения независимы, $\mathrm{E}(XY)=\mathrm{E}X\cdot\mathrm{E}Y=2$ (кв.м.).

б) Из условия следует, что $\mathrm{D}X=0{,}003^2=9\cdot10^{-6}$ и $\mathrm{D}Y=4\cdot10^{-6}$.

$$\mathrm{D}(XY)=\mathrm{E}(XY)^2-\mathrm{E}^2(XY)=\mathrm{E}X^2\cdot\mathrm{E}Y^2-4.$$

Воспользуемся равенством $\mathrm{D}X=\mathrm{E}X^2-\mathrm{E}^2X$, выразим из него $\mathrm{E}X^2$ и подставим в полученное выражение. Аналогично поступим с $\mathrm{E}Y^2$:

$$\mathrm{D}(XY)=\left(\mathrm{D}X+\mathrm{E}^2X\right)\left(\mathrm{D}Y+\mathrm{E}^2Y\right)-4=\left(4\cdot10^{-6}+1\right)\left(9\cdot10^{-6}+4\right)-4=36\cdot10^{-12}+25\cdot10^{-6}.$$

По сравнению со вторым слагаемым первое настолько мало, что его можно отбросить. Получаем: $\sqrt{\mathrm{D}(XY)}\approx\sqrt{25\cdot10^{-6}}=0{,}005$ (кв.м), что дает 50 кв.см.

Ответ: а) 2 кв.м; б) прибл. 50 кв.см.

10. Диски (от 9 класса. 3 балла). На знакомом нам заводе вырезают металлические диски диаметром 1 м. Известно, что диск диаметром ровно 1 м весит ровно 100 кг. При изготовлении возникает ошибка измерения, и поэтому стандартное отклонение радиуса составляет 10 мм. Инженер Сидоров считает, что стопка из 100 дисков в среднем будет весить 10000 кг. На сколько ошибается инженер Сидоров?

Решение. По условию $\mathrm{E}R=0{,}5$ м, $\mathrm{D}R=10^{-4}$ (кв.м). Найдем математическое ожидание площади одного диска:

$$\mathrm{E}S=\mathrm{E}\left(\pi R^2\right)=\pi\mathrm{E}R^2=\pi\left(DR+\mathrm{E}^2R\right)=\pi\left(10^{-4}+0{,}25\right)=0{,}2501\pi.$$

Значит, математическое ожидание массы диска равно $\dfrac{0{,}2501\pi}{0{,}25\pi}\cdot100=100{,}04$ кг. Следовательно, стопка из 100 дисков в среднем будет весить 10004 кг.

Ответ: на 4 кг.

11. Пересекающиеся диагонали 2 (от 9 класса. 3 балла). В выпуклом многоугольнике, в котором нечетное число вершин, равное $2n+1$, выбирают независимо друг от друга две случайные диагонали. Найдите вероятность того, что эти диагонали пересекаются внутри многоугольника.

Решение. Всего в $2n+1$- угольнике

$$C_{2n+1}^{2}-(2n+1)=\frac{(2n+1)\cdot 2n}{2}-(2n+1)=2n^2-n-1$$

диагоналей, и, следовательно, пар диагоналей всего

$$C_{2n^2-n-1}^{2}=\frac{\left(2n^2-n-1\right)\left(2n^2-n-2\right)}{2}.$$

Две пересекающиеся диагонали однозначно определяются выбором четырех вершин, значит, способов выбрать две пересекающиеся диагонали

$$C_{2n+1}^{4}=\frac{(2n+1)2n(2n-1)(2n-2)}{24}=\frac{(2n+1)n(2n-1)(n-1)}{6}.$$

Диагонали выбираются случайным образом, поэтому все комбинации равновозможны. Поэтому вероятность того, что две случайные диагонали пересекаются, равна

$$\frac{C_{2n+1}^{4}}{C_{2n^2-n-1}^{2}}=\frac{(2n+1)n(2n-1)(n-1)\cdot 2}{6\cdot\left(2n^2-n-1\right)\left(2n^2-n-2\right)}=\frac{n(2n-1)}{3\left(2n^2-n-2\right)},$$

Ответ: $\dfrac{2n^2-n}{3\left(2n^2-n-2\right)}$.

Комментарий. С ростом *n* полученная вероятность приближается к 1/3. Возникает предположение, что если взять эллипс или окружность и поставить вопрос о вероятности пересечения двух случайных хорд, ответ должен быть 1/3. К сожалению, не все так просто. Важно, как именно выбраны случайные хорды, ведь хорд у окружности бесконечно много, в отличие от диагоналей многоугольника. Можно поступить, например, так: каким-то случайным образом выберем на окружности четыре точки *A*, *B*, *C* и *D*, независимо одна от другой. Какова вероятность того, что хорды *AB* и *CD* пересекаются? И вот здесь ответ действительно 1/3. Можете попробовать это доказать.

12. Сногсшибательная новость. На конференцию приехали 18 ученых, из которых ровно 10 знают сногсшибательную новость. Во время перерыва (кофе-брейка) все ученые разбиваются на случайные пары, и в каждой паре каждый, кто знает новость, рассказывает эту новость другому, если тот её ещё не знал.

а) (от 6 класса. 1 балл). Найдите вероятность того, что после кофе-брейка число ученых, знающих новость, будет равно 13.

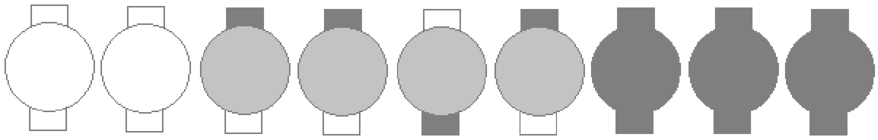
б) **(от 10 класса. 4 балла).** Найдите вероятность того, что после кофе-брейка число, ученых, знающих новость, будет равно 14.

в) **(от 9 класса. 3 балла).** Обозначим буквой X количество ученых, которые знают сногсшибательную новость после кофе-брейка. Найдите математическое ожидание X.

Решение. а) После кофе-брейка новость будут знать все, кто попал в пару, где есть хотя бы один «знайка» (тот, кто знал новость раньше). Поскольку эти ученые образуют целое число пар, их число четно. Поэтому вероятность того, что их будет 13 человек, равна нулю.

б) Задача напоминает задачу 7 про теннисисток, поскольку дело сводится к исчислению пар, состоящих из «знайки» и «незнайки». Для простоты будем считать, что ученые расселись за двухместными столиками. Стулья, на которых сидят знайки, закрасим серым цветом. Стулья незнаек оставим белыми.

Всего существует $C_{18}^{10} = C_{18}^{8} = 43758$ способов рассадить 10 знаек и 8 незнаек на 18 стульев, стоящих около двухместных столиков. Нас интересуют такие рассадки, при которых ровно 4 знайки оказались за одним столиком с незнайками, чтобы передать им свое знание и довести число знающих до 14 (см. рис.)



Будем рассуждать, как в задаче 7: выберем из девяти столиков четыре для пар «знайка-незнайка» (на рисунке эти столики покрыты серыми скатертями). Из оставшихся 5 столиков выберем 2 столика для незнаек (белые скатерти). Заметим, что пары «знайка-незнайка» можно посадить 2^4 способами, заставляя знайку и незнайку за каждым столиком меняться местами. На рисунке показана одна из таких возможных рассадок, а общее число нужных рассадок равно $C_9^4 \cdot C_5^2 \cdot 2^4 = 126 \cdot 10 \cdot 16 = 20160$.

Таким образом, вероятность того, что число знающих новость станет ровно 14, равна

$$\frac{C_9^4 \cdot C_5^2 \cdot 2^4}{C_{18}^8} = \frac{20160}{43758} = \frac{1120}{2431} \approx 0,461.$$

в) Можно решить задачу пункта б) в общем виде, тем самым получив распределение величины X, а затем найти математическое ожидание $\mathrm{E}X$. Но мы поступим проще, применив индикаторы. Занумеруем каким-нибудь способом незнаек числами от 1 до 8 и введем индикаторы I_k по числу незнаек. $I_k = 1$, если во время кофе-брейка незнайка номер k стал знайкой, и $I_k = 0$, если он так и не узнал новость. Очевидно,

$$X = 10 + I_1 + I_2 + \ldots + I_8.$$

Распределения всех индикаторов одинаковы:

$$I_k \sim \begin{pmatrix} 0 & 1 \\ \frac{7}{17} & \frac{10}{17} \end{pmatrix},$$

поскольку в пару к любому незнайке может подсесть любой из оставшихся 17 ученых, и ровно 10 из них знайки. Следовательно, $\mathrm{E}I_k = \frac{10}{17}$. Тогда

$$\mathrm{E}X = 10 + \mathrm{E}I_1 + \mathrm{E}I_2 + \ldots + \mathrm{E}I_8 = 10 + 8 \cdot \frac{10}{17} = 14\frac{12}{17} \approx 14,7.$$

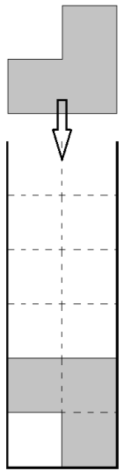
Ответ: а) 0; б) прибл. 0,461; в) прибл. 14,7.

13. Маленький тетрис. Высокий прямоугольник ширины 2 открыт сверху, и в него падают в случайной ориентации Г-тримино (см. рисунок).

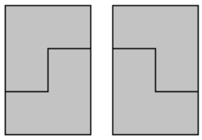
а) (от 9 класса. 3 балла). Упало k тримино. Найдите математическое ожидание высоты получившегося многоугольника.

б) (от 10 класса. 6 баллов). Упало 7 тримино. Найдите вероятность того, что сложенная из тримино фигура будет иметь высоту 12.

Решение. а) Случайную величину «высота получившегося многоугольника» обозначим X. Очевидно, $X = 2k - (I_2 + I_3 + \ldots + I_k)$, где I_j – индикатор события «тримино с номерами j и $j-1$ образовали блок высоты 3». Это может быть только в двух случаях, показанных на рисунке слева. Вероятность такого равна $\frac{1}{8}$. Таким образом $\mathrm{E}\, I_j = \frac{1}{8}$. Поэтому $\mathrm{E}\, X = 2k - \frac{1}{8}(k-1) = \frac{15k+1}{8}$.



б) Обозначим $u_{k,n}$ и $d_{k,n}$ вероятности того, что фигура из k упавших тримино достигнет высоты n, и при этом верхнее тримино будет расположено выступом вверх (вниз). Сумму этих вероятностей, то есть вероятность того, что высота многоугольника из k тримино равна n, обозначим $p_{k,n}$. Тогда



$$u_{k,n} = \frac{1}{2} p_{k-1,n-2},\quad d_{k,n} = \frac{1}{4} u_{k-1,n-1} + \frac{1}{4} u_{k-1,n-2} + \frac{1}{2} d_{k-1,n-2}.$$

Полная вероятность равна

$$p_{k,n} = u_{k,n} + d_{k,n} = \frac{1}{2} p_{k-1,n-2} + \frac{1}{2} d_{k-1,n-2} + \frac{1}{4} u_{k-1,n-2} + \frac{1}{4} u_{k-1,n-1} = p_{k-1,n-2} - \frac{1}{4} u_{k-1,n-2} + \frac{1}{4} u_{k-1,n-1},$$

откуда $p_{k,n} = p_{k-1,n-2} + \frac{1}{8} p_{k-2,n-3} - \frac{1}{8} p_{k-2,n-4}$.

Для начала рекурсии положим: $p_{0,0} = p_{1,2} = 1$, кроме того $p_{k,n} = 0$ при $n > 2k$. Получаем: $p_{2,3} = p_{1,1} + \frac{1}{8} p_{0,0} - \frac{1}{8} p_{0,-1} = 0 + \frac{1}{8} - 0 = \frac{1}{8}$, $p_{2,4} = p_{1,2} + \frac{1}{8} p_{0,1} - \frac{1}{8} p_{0,0} = 1 + 0 - \frac{1}{8} = \frac{7}{8}$ и так далее. Расчет можно провести, например, с помощью Excel (см. пример на рисунке). Для удобства можно ввести $n = -1$, положив соответствующие вероятности равными нулю. На рисунке этому соответствует пустой столбец C. Тогда останется заполнить соответствующей формулой только ячейки, начиная с $k = 2, n = 3$ (G5 в нашем примере).

P10 fx =N9+M8/8-L8/8

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1			Высота получившегося многоугольника n															
2			-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
3	Число тримино k	0		1														
4		1				1												
5		2					0,125	0,875	0	0	0	0	0	0	0	0	0	0
6		3					0	0	0,25	0,75	0	0	0	0	0	0	0	0
7		4					0	0	0	0,016	0,344	0,641	0	0	0	0	0	0
8		5					0	0	0	0	0	0,047	0,406	0,547	0	0	0	0
9		6					0	0	0	0	0	0	0,002	0,088	0,443	0,467	0	0
10		7					0	0	0	0	0	0	0	0	0,008	0,133	0,461	0,398
11		8					0	0	0	0	0	0	0	0	0	2E-04	0,019	0,177
12		9					0	0	0	0	0	0	0	0	0	0	0	0,001
13		10					0	0	0	0	0	0	0	0	0	0	0	0
14		11					0	0	0	0	0	0	0	0	0	0	0	0

На пересечении 7 строки и 12 столбца находим: $p_{7,12} \approx 0{,}133$.

Ответ: а) $\frac{15k+1}{8}$; б) прибл. 0,133.

14. Запутанное дело (от 10 класса. 4 балла). Расследуя одно дело, следователь Башковицкий обнаружил, что ключевой свидетель – тот из семьи Петровых, кто в тот роковой день пришёл домой прежде прочих. Расследование выявило следующие факты.

1. Соседка Марья Кузьминична хотела одолжить у Петровых соли, звонила им в дверь, но никто не открыл. Во сколько? Да кто ж знает? Темно уж было…

2. Галина Ефимовна Петрова, придя вечером домой, обнаружила обоих детей на кухне, а мужа на диване – у него болела голова.

3. Муж Анатолий Иванович заявил, что как пришёл, сразу лёг на диван и задремал, никого не видел, ничего не слышал, соседка точно не приходила – звонок бы его разбудил.

4. Дочь Светлана сказала, что, вернувшись домой, сразу ушла к себе в комнату, про отца ничего не знает, но в прихожей, как всегда, споткнулась о Димкин ботинок.

5. Дмитрий когда пришёл – не помнит, отца не видел, а как Светка ругалась из-за ботинка – слышал.

– Ага, – задумался Башковицкий. – Какова же вероятность того, что Дмитрий вернулся домой раньше отца?

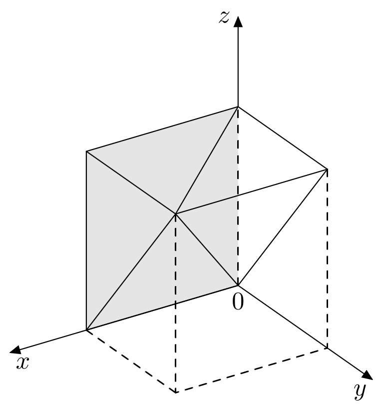
Решение. Относительно Дмитрия, Светланы и Анатолия определённости нет, известно лишь, что все трое пришли домой между визитом Марьи Кузьминичны и приходом Галины Ефимовны – именно эти два события ограничивают период времени, интересующий Башковицкого. Будем считать, что вероятность прихода члена семьи домой в какой-то промежуток времени пропорциональна длине этого промежутка.

Введём обозначения. Пусть от визита Марьи Кузьминичны до прихода Галины Ефимовны прошло время m, до прихода Анатолия – время x, до прихода Дмитрия – время y, а до прихода Светланы – z.

Пользуясь свидетельскими показаниями, составим систему:

$$\begin{cases} 0 < x < m, \\ 0 < y < z < m, \end{cases}$$

которая в системе координат $Oxyz$ дает треугольную призму F с вершинами $(0;0;0)$, $(0;0;m)$, $(0;m;m)$, $(m;0;0)$, $(m;0;m)$ и $(m;m;m)$ (см. рисунок). При этих условиях требуется найти вероятность события $y < x$. Внутри призмы F это неравенство определяет четырёхугольную пирамиду G с вершинами $(0;0;0)$, $(m;0;0)$, $(m;0;m)$, $(0;0;m)$ и $(m;m;m)$.



Искомая вероятность равна $\frac{V_G}{V_F} = \frac{m^3}{3} : \frac{m^3}{2} = \frac{2}{3}$.

Ответ: $\frac{2}{3}$.

15. Английский клуб (от 9 класса. 6 баллов). Каждую пятницу десять джентльменов приходят в клуб, и каждый отдает швейцару свою шляпу. Каждая шляпа точно впору своему хозяину, но двух одинаковых по размеру шляп нет. Уходят джентльмены по одному в случайном порядке.

Провожая очередного джентльмена, швейцар клуба пробует надеть ему на голову первую попавшуюся шляпу. Если налезает, джентльмен уходит в этой шляпе. Если мала, то швейцар пробует следующую случайную шляпу из оставшихся. Если все оставшиеся шляпы оказались малы, швейцар говорит бедняге: «Сэр, сегодня шляпа вам не к лицу», и джентльмен отправляется домой с непокрытой головой. Найдите вероятность того, что в следующую пятницу у швейцара не останется ни одной шляпы.

Решение. Пусть джентльменов n. Занумеруем их в порядке возрастания размеров их шляп номерами от 1 до n. Шляп не останется, только если каждый забрал свою же шляпу. Обозначим вероятность этого p_n. Если первым уходит k-й джентльмен (вероятность этого $\frac{1}{n}$), то вероятность получить свою шляпу у него равна $\frac{1}{n-k+1}$ (поскольку своя шляпа – единственная среди $n-k+1$ подходящих этому джентльмену). Если это случилось, то вероятность того, что все оставшиеся заберут свои же шляпы, равна p_{n-1}. По формуле полной вероятности получаем:

$$p_n = \frac{1}{n}\cdot\frac{1}{n}p_{n-1} + \frac{1}{n}\cdot\frac{1}{n-1}p_{n-1} + \ldots + \frac{1}{n}\cdot\frac{1}{2}p_{n-1} + \frac{1}{n}\cdot\frac{1}{1}p_{n-1} = \frac{H_n}{n}p_{n-1},$$

где $H_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$ – n-е гармоническое число. Учитывая, что $p_1 = 1$ (если джентльмен только один, то он обязательно возьмёт свою шляпу), получаем:

$$p_n = \frac{H_n}{n}p_{n-1} = \frac{H_n H_{n-1}}{n(n-1)}p_{n-2} = \ldots = \frac{(H_n)!}{n!},$$

где символом $(H_n)!$ обозначено произведение всех гармонических чисел от $H_1 = 1$ до H_n.

Подставляя $n = 10$, получаем: $p_{10} = \dfrac{1\cdot\left(1+\frac{1}{2}\right)\left(1+\frac{1}{2}+\frac{1}{3}\right)\cdot\ldots\cdot\left(1+\frac{1}{2}+\ldots+\frac{1}{10}\right)}{10!} \approx 0,000516$. Вычисление можно легко провести с помощью калькулятора или компьютера.

Ответ: прибл. 0,000516.

16. Ничьи (от 9 класса. 6 баллов). Две хоккейные команды одинаковой силы договорились, что будут играть до тех пор, пока суммарный счёт не достигнет 10. Найдите математическое ожидание числа моментов, когда наступала ничья.

Решение. Если $2n$ – максимальный суммарный счёт, то игру можно рассматривать как случайное блуждание длины $2n$: на каждом шаге разрыв в счете либо увеличивается на единицу, либо уменьшается на единицу.

Пусть I_{2k} – индикатор ничьей на $2k$ шаге:

$$I_{2k} \sim \begin{pmatrix} 0 & 1 \\ 1-\dfrac{C_{2k}^k}{2^{2k}} & \dfrac{C_{2k}^k}{2^{2k}} \end{pmatrix}.$$

Случайная величина X «число моментов, когда наступала ничья» равна сумме всех индикаторов. Начало игры не будем считать «наступлением ничьей». Следовательно $X = I_2 + I_4 + \ldots + I_{2n}$. Это даёт решение:

$$\mathrm{E}\,X = \sum_{k=1}^{n}\mathrm{E}I_{2k} = \sum_{k=1}^{n}\frac{C_{2k}^k}{2^{2k}}.$$

При $n=5$ получаем: $\mathrm{E}\,X=\frac{2}{4}+\frac{6}{16}+\frac{20}{64}+\frac{70}{256}+\frac{252}{2048}\approx 1{,}707$.

Ответ: прибл. 1,707.

Замечание. Если воспользоваться формулой Стирлинга $n!\approx\sqrt{2\pi n}\left(n/e\right)^n$, то после преобразований получим:

$$\mathrm{E}\,X\approx\sum_1^n\frac{1}{\sqrt{\pi k}}=\frac{1}{\sqrt{\pi}}\sum_1^n\frac{1}{\sqrt{k}}.$$

При $n=5$ эта приближенная формула дает 1,823 (все знаки верные). Даже для такого небольшого n относительная погрешность составляет около 4%, а с ростом n будет уменьшаться.

Дальнейшее упрощение (также приближенное) можно сделать, если заменить сумму подходящим интегралом. Например, так:

$$\frac{1}{\sqrt{\pi}}\sum_1^n\frac{1}{\sqrt{k}}\approx\frac{1}{\sqrt{\pi}}\int_0^n\frac{1}{\sqrt{x+\pi/4}}\,dx=\frac{2}{\sqrt{\pi}}\left(\sqrt{n+\pi/4}-\frac{\sqrt{\pi}}{2}\right)=\sqrt{\frac{4n}{\pi}+1}-1.$$

При больших n единица под корнем практически не играет роли, поэтому ее можно отбросить, получив еще более простую формулу

$$\mathrm{E}\,X\approx 2\sqrt{\frac{n}{\pi}}-1.$$

Убедитесь с помощью компьютера, что обе эти формулы дают приемлемо малую ошибку.